Teaching Children to Attribute Second-order False Belief: A Training Study (Extended Abstract)

Burcu Arslan[1], Rineke Verbrugge[1], Niels Taatgen[1], Bart Hollebrandse[2]
[1] Institute of Artificial Intelligence, University of Groningen, P.O. Box 407,
9700 AK Groningen, The Netherlands
[2]Faculty of Arts, University of Groningen, P.O.Box 716,
9700 AS Groningen, The Netherlands
{b.arslan, l.c.verbrugge, n.a.taagen, b.hollebrandse}@rug.nl

1 INTRODUCTION

To understand that different people have different mental states, such as desires, beliefs, knowledge and intentions, which can be different from one's own, is called Theory of Mind [1]. While explicit first-order Theory of Mind (ToM) ("Mary *believes* that [there is a chocolate in the drawer]") develops around the age of four [2], second-order ToM ("Jack *thinks* that [Mary *knows* that [there is a chocolate in the fridge]]") develops between the ages of five and six [3].

One of the most applied verbal tasks for studying the development of ToM is the false belief task (FBT). The goal of the first-order FBT is to examine whether children can attribute a false belief to another person in a given story where the child knows the reality and the other person has a false belief. Similarly, the goal of the second-order FBT is to test whether children can attribute a false belief to another person who is attributing a belief to the third agent.

In previous work, Arslan, Taatgen and Verbrugge (2013) constructed a computational cognitive model in order to show the developmental transitions from first- to second-order ToM [4]. There are two main predictions of their model:

1. Children who are able to give correct answers for the first-order ToM questions but not for the second-order ToM questions do give first-order (and not zero-order) answers for the second-order ToM questions.
2. Children who are able to pass the first-order ToM task can learn to pass second-order ToM with the help of feedback.

In order to test these predictions, we conducted a training experiment with primary school children between the ages of 5 and 6 years old.

2 METHOD

2.1. Participants

A sample of 26 Dutch children was tested. The children were pre-tested to ensure that they had not yet acquired second-order false belief reasoning. In line with our pre-defined exclusion criteria, three children who gave correct answers for all of the second-order false belief questions during the pre-test were excluded from the analysis. Thus the results of 23 children were included to the analysis (15 female, M_{age}=5.8 years, SE=0.06, range: 5.1 – 6.2).

2.2. Materials

We used three different types of second-order FBTs: (a) Birthday Puppy-like stories, (b) Chocolate Bar-like stories, (c) Bake Sale-like Stories. To be able to test the first prediction of the model, we constructed the Chocolate Bar-like stories and the Bake Sale-like Stories in such a way that it is possible to distinguish the most likely order of the child's ToM reasoning from the given answer. In total, we constructed 30 different second-order false belief stories.

We also constructed four true belief stories in order to test whether children possibly use a simple strategy instead of reasoning about second-order mental state attributions. Two of the true belief stories had the same structure as the Bake Sale-like stories and the other two had the same structure as the Birthday Puppy-like stories.

In addition to testing the transfer effect by training the children with second-order FBT (near transfer), we wanted to see whether children can transfer what they learned during training to another domain in which they should apply first- and second-order reasoning (far transfer). For this purpose, we adopted Meijering and colleagues' (submitted) Marble Drop (MD) game [5]. In this game, children are instructed that they play against a computer. Both the child's and the computer's aim is to maximize their own pay-off. During the game, sometimes children are asked to make a decision that takes into account the computer's decision (first-order), and sometimes they are asked to make a prediction about the computer's decision, which takes into account their own decision at the final decision point (second-order). Figure 1 shows an example of a second-order game where the computer controls the blue trapdoors and the child controls the orange trapdoors.

Lastly, we adapted Towse, Hitch, and Hutton's (1998) Counting Span Task and used this during the pre- and post-tests to control for children's working memory score [6].

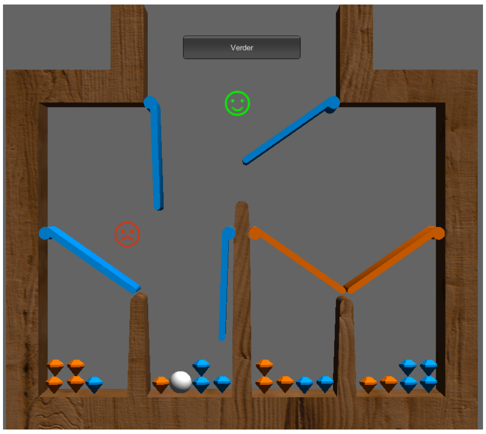

Fig.1. An example of second-order game

2.3. Procedure

Children were tested individually for about 30 minutes in their preschool in a series of four different sessions: i) pre-test, ii) first training session, iii) second training session, and iv) post-test, respectively. There was at least one week between pre- and post-tests. Children were trained in two different days in between the pre-test and the post-test. There was at least one day between the first and the second training days. Drawings of the episodes of the stories were presented one by one together with their audio recordings. The drawings remained visible throughout the story.

Pre- and post-test. Children were tested with one MD game, one Counting Span Task, and three second-order FBTs (one Bake Sale-like, one Birthday Puppy-like, and one Chocolate Bar-like) in a random order. Children did not get any feedback for the second-order FBTs in the pre-test and post-test. Because children were not trained with Chocolate Bar-like stories during the two training sessions, this type of stories was used to test the near transfer effect of the training. In the MD game, positive (green happy smiley) or negative feedback (red sad smiley) was always provided during the trials.

Training Sessions. Per training sessions, each child was trained with 6 different second-order FBTs (3 Bake Sale-like, and 3 Birthday Puppy-like) in a random order. Per story, a feedback (correct or incorrect) and an explanation were provided by the experimenter in an interactive fashion. After three second-order FBTs were presented, one true belief story was presented to check whether the children were using a simple strategy or not.

3 RESULTS

Most of the time (63%) children who gave wrong answers provided first-order answers for second-order false belief questions. Whilst 17% of answers were zero-order answers, 20% of answers were “I don’t know.” This finding confirms the first prediction of Arslan and colleagues’ (2013) model.

Figure 2.a shows children’s improvement from pre- to post-tests. There is a significant improvement from pre- to post-tests (V=27, p = .002), and from training day 1 to training day 2 (V=36, p= .03). These results confirm the second prediction of Arslan and colleagues’ (2013) model that children’s performance can be improved by providing feedback. Whilst the children gave correct answers to 76% of the true belief questions in the first training day, the correct answers increased to 91% in the second training day. This result suggests that children did not use a simple strategy (i.e. pattern recognition) instead of reasoning about the questions.

Figure 2.b shows children’s improvements in different types of second-order FBTs from pre- to post-tests. There are significant improvements from pre- to post tests in Bake Sale-like stories (V= 27, p= .008) and Chocolate Bar-like stories (V=15, p= .008), and there is a marginally significant improvement in Birthday Puppy-like stories (V=11, p= .065).

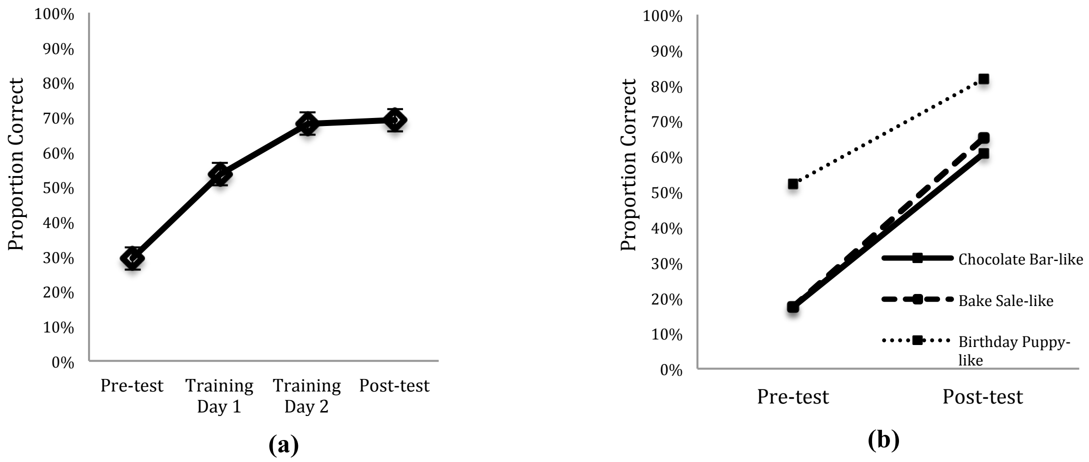

Fig. 2. **(a)** Changes in performance over the course of training. **(b)** Children’s performance from pre- to post-tests in different types of story types

Children’s MD game scores did not differ from the chance level for the first- and the second-order games during pre- and post-tests. Moreover, children’s counting span score does not predict their pre- and post-tests scores.

4 CONCLUSIONS

This study shows that children can learn to attribute second-order false belief to an agent with the help of explicit feedback with explanation. Our finding that children’s performance also improved in Chocolate Bar-like stories (near transfer) suggests that this improvement is not just limited to the stories that we used during training sessions, but that there is a more flexible improvement that they can use in a different type of situation. On the other hand, our finding that children's performance did not improve in the Marble Drop game (far transfer) suggests that children cannot transfer their knowledge in a completely different setting in which they should apply second-order reasoning.

Because we provided feedback with explanation during the training sessions, we do not know whether the improvement is because of the explanation, nor whether providing just feedback without explanation would also work. To test this, as a second condition, we will train children by providing feedback without explanation.

5. ACKNOWLEDGEMENTS

This work was supported by the Netherlands Organisation for Scientific Research (NWO) Vici grant NWO 277-80-001, awarded to Rineke Verbrugge.

6. REFERENCES

1. Premack, D., & Woodruff, G.: Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 4. (1978) 515-526.
2. Sullivan, K., Zaitchik, D., & Tager-Flusberg, H.: Preschoolers can attribute second-order beliefs. *Developmental Psychology*, 30 (3). (2008) 395-402.
3. Wimmer, H., & Perner, J.: Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition, 13. (1983) 103–128.
4. Arslan, B., Taatgen, N. A., & Verbrugge, R. : Modeling Developmental Transitions in Reasoning about False Beliefs of Others. In R. West & T. Stewart (eds.), *Proceedings of the 12th International Conference on Cognitive Modeling*, Ottawa: Carleton University. (2013) 77-82.
5. Meijering, B., Van Tijn, H., Taatgen, N.A., & Verbrugge, R.: *Reasoning about self versus others: Changing perspective is hard.* (Submitted).
6. Towse, J., Hitch, G., & Hutton, U.: A reevaluation of working memory capacity in children. *Journal of Memory and Language*, 39. (1998) 195-217.