Modeling Developmental Transitions in Reasoning about False Beliefs of Others

Burcu Arslan (B.Arslan@rug.nl), Niels Taatgen (N.A.Taatgen@rug.nl), Rineke Verbrugge (L.C.Verbrugge@rug.nl)
Institute of Artificial Intelligence, University of Groningen, P.O. Box 407,
9700 AK Groningen, The Netherlands

Abstract

Reasoning about false beliefs of others develops with age. We present here an ACT-R model in order to show the developmental transitions. These start from a child's reasoning from his/her own point of view (zero-order) to taking into consideration another agent's beliefs (first-order), and later to taking into consideration another agent's beliefs about again other agents' beliefs (second-order). The model is based on a combination of rule-based and simulation approaches. We modeled the gradual development of reasoning about false beliefs of others by using activation of declarative knowledge instead of utility learning. Initially, in addition to the story facts, there is only one strategy chunk, namely a zero-order reasoning chunk, in declarative memory. The model retrieves this chunk each time it has to solve a problem. Based on the feedback, the model will strengthen a successful strategy chunk, or it will add or strengthen an alternative strategy if the current one failed.

Keywords: Second-order Theory of Mind; Cognitive Development; False Belief Reasoning; Cognitive Modeling.

Introduction

The ability to understand that other 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 (ToM) (Premack & Woodruff, 1978).

Adults can recursively apply this ability. Zero-order reasoning concerns our real-life environment. For example, if Murat thinks: "Ayla wrote a novel under a pseudonym", he is applying zero-order reasoning. Now Ayla might correctly think, "Murat *knows* that I wrote a novel under a pseudonym", thus making a first-order knowledge attribution to Murat. In addition to first-order reasoning, there are more complex social situations, such as Murat (mistakenly) thinking, "Luckily, Ayla *believes* that I do not *know* that she wrote a novel under a pseudonym". Thereby, Murat makes a second-order attribution to Ayla, namely of a false belief about his own knowledge of the real world. When discussing this example, we, in turn, are performing third-order reasoning about Murat by making the attribution: "Murat *thinks* that Ayla *believes* that he does not *know* that she wrote a novel under a pseudonym".

First-order theory of mind develops between the ages of three and five (Wimmer & Perner, 1983). Interestingly, second-order theory of mind develops much later, around at the age of 5 (Miller, 2009). The reason for this gap is not entirely clear yet and attracts the curiosity of researchers who are working on ToM. Although the development of first-order ToM has been intensely investigated and documented in the literature, there are much fewer studies related to the development of higher-order ToM, but see, for example, (Perner and Wimmer, 1985; Liddle and Nettle, 2006; Flobbe et al., 2008; Arslan et al., 2012; Miller, 2012; Raijmakers et al., 2013; Hollebrandse et al., 2013).

The two main theories about ToM are *theory-theory* and *simulation theory*. In short, the theory-theorists claim that mental states are just like other entities and can be inferred from observing the behavior of others. According to Gopnik and Wellman (1992), children act like scientists by constructing theories on the basis of their observations and construct rules that can be applied for different circumstances. However, Mitchell and colleagues (2009) find this theory insufficient to explain children's gradual development: if children used a rule-based approach, the developmental trend would be sharp, and they would start to pass the false belief tasks as soon as they had built up the correct rule. In contrast, *simulation theorists* claim that people imagine themselves in the other person's situation and make decisions by running simulations. This approach appears to be supported by the discovery of mirror neurons (Keysers & Gazzola, 2006). According to simulation theorists, children find it difficult to simulate another person's mental processes, as they should not only put themselves into the other person's shoes, but after that they should inhibit and set aside their own perspective and take into account the other person's knowledge and beliefs.

By combining the two main theories of ToM, Mitchell and colleagues (2009) constructed a *hybrid theory*. According to this theory, children first do simulations. Since they cannot set aside their own perspective, they make systematic errors. After learning to perform successful simulations by setting aside their perspective, people come up with more general rules and use these rules instead of simulation.

Leslie and colleagues (2004) posit a modular approach. They suggest that there are two different mechanisms to provide social cognition, namely the theory of mind mechanism (ToMM) and the selection process (SP) mechanism. While ToMM generates and represents possible beliefs, SP inhibits the counter-beliefs and selects the appropriate belief. They state that the ToMM is innate. By contrast, SP develops with age and causes the failure of younger children on theory of mind tasks, such as the false belief task.

The Second-order False Belief Task

The main idea of the false belief task (FBT) is to examine whether children can attribute a false belief to another agent in a given story where the child knows the reality and the other agent has a false belief. There are many variations of second-order false belief tasks (e.g. Flobbe, Verbrugge, Hendriks, Krämer, 2008; Perner & Wimmer, 1985; Tager-Flusberg & Sullivan, 1994). In our model we used Flobbe and colleagues' (2008) second-order 'Chocolate Bar' story, which was originally constructed as a first-order FBT by Hogrefe and Wimmer (1986). The English version of our task as follows:

'Murat and Ayla are brother and sister. Here they are in the living room. Then mother returns from shopping. Mother bought some chocolate. She gives the chocolate to Murat. Ayla doesn't get any chocolate, because she has been naughty. Murat eats some of the chocolate and puts the remainder into the drawer. He doesn't give any chocolate to Ayla. That makes Ayla angry. Now Murat goes to help mother in the kitchen. He is helping with the dishes. Ayla is alone in the living room. Murat is in the kitchen. Because she is angry with Murat, Ayla hides the chocolate. She takes the chocolate out of the drawer and puts it into the toy box. Murat is busy doing the dishes. He throws the fruit leftovers into the rubbish bin in the garden. Through the window he sees the living room. He sees how Ayla takes the chocolate out of the drawer, and puts it into the toy box. Ayla does not see Murat.'

At this point, the experimenter asks control questions in order to make sure that the participants do not have any problem with understanding the story facts. If the participants give correct answers to those questions, the experimenter continues to tell the story as follows:

'Murat has finished the dishes. He is hungry. Now he wants to eat some of his chocolate. Murat enters the living room. He says: "Hmm, I would like some chocolate."'

Now the experimenter asks the second-order false belief question: 'Where does Ayla think that Murat will look for the chocolate?' and the justification question: 'Why does she think that?' If the participant's answers to the second-order false belief question and to the justification question are correct, then the second-order false belief answer is scored as "Correct".

Empirical Work

Wimmer and Perner (1983) devised a first-order FBT for the first time and concluded that first-order false belief reasoning develops between the ages three and five. After their preliminary study, hundreds of studies replicated their results and showed that while children at the age of 4 can answer the first-order false belief question (e.g. 'Where will Murat look for the chocolate?') correctly, most children at the age of 3 generally give an answer in terms of reality without taking into account others' false beliefs. Some studies (e.g. Siegal & Beattie, 1991; Sullivan & Winner, 1993) manipulated some conditions of the classic FBT and showed that 3-year-olds can pass the test. In order to analyze the contradictory findings, Wellman and colleagues (2001) published a review article. They analyzed 178 studies about the first-order FBT. Their analysis showed that while children around the age of 3 are more than 80% incorrect, children around the age of 4 are 50% correct (see Figure 1). As can be seen from the results, 3-year-old children do not make errors randomly; instead their errors are systematic in the way that they tend to report their own knowledge instead of the other agent's belief. One of the explanations for this type of error is the children's *'reality-bias'* (Mitchell et al., 1996). According to this view, in order to give correct answers to the false belief questions, children should inhibit their own response and take into account others' perspectives. Wellman and colleagues (2001) concluded that developmental changes are not due to task manipulations but arise thanks to conceptual changes in children's understanding of other people's minds.

After Wimmer and Perner's (1983) seminal study of a first-order false belief reasoning, researchers studied the second-order false belief reasoning and their results showed that this ability does not develop before the age of 5 (Miller, 2009). In a very recent study, Arslan and colleagues (2012) also studied the development of second-order false belief reasoning with Turkish children between the ages 4 and 12. Their results are compatible with the previous studies of second-order FBT. They found that there is a big step-wise increase between the ages 4 (12% correct) and 7 (65% correct). After that, the development continues until around the age of 11 when children give 100% correct answers for the second-order false belief questions (see Figure 1).

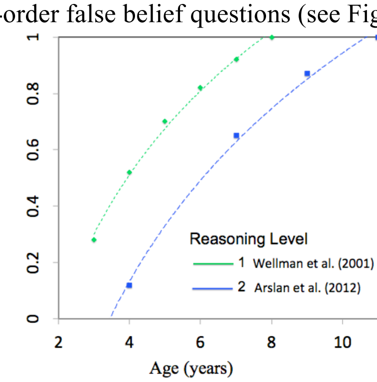

Figure 1: The development of first- and second-order false belief reasoning (adapted from Wellman et al., 2001; and Arslan et al., 2012)

Previous Models of Explicit False Belief Task

There are few computational models of false belief tasks (FBT). The earliest one is Wahl and Spada's (2000) Prolog model. They simulated children's first-order and second-order false belief reasoning by using practical syllogisms and also inference schemata. They used two different types of FBT, namely explanation tasks and prediction tasks. In the explanation tasks, people are expected to infer another person's intentions in terms of their actions by attributing a second-order belief such as "Person A thinks that Person B thinks that…". In the prediction tasks, people are expected to infer the actions of another person in terms of their intention. Their model predicted that while the operational demands of explanation tasks are higher than those of the prediction tasks for the second-order FBT, there is no difference for the first-order FBT. They also conducted experiments with children that confirmed these predictions. Even though their model cannot explain the developmental trend of false belief reasoning, their study showed how computational modeling could be used to explain the difference between first-order and second-order false belief tasks in terms of computational complexity.

Triona and colleagues (2002) constructed an ACT-R model in order to explain the developmental trend of false belief reasoning in children by using an unexpected contents task. Their model answers the control questions and the false belief question by identifying the content of the box using general knowledge, and updating the general knowledge by the specific input about the real content of the box. They hypothesized that the older children have had further experience with false belief-like questions while the younger children have not and they manipulated a parameter in order to explain the developmental trend. For this reason, their model could not suggest explanations about the development itself.

Another study on development of false belief reasoning is Bello and Cassimatis' (2006) Polyscheme model for the unexpected content and unexpected location tasks. Their model explains the difference between 3-year-olds' and 4 year-olds' performance by positing that they use two different rules. The rule that 3-year-olds apply is: "IF P wants OBJECT and P perceives object at LOCATION, THEN P will search for OBJECT at LOCATION". In contrast, the rule that 4-year-olds apply is: "IF P wants OBJECT and in his mind the object is at LOCATION, THEN P will search for OBJECT at LOCATION." Since Bello and Cassimatis' model is based on the rule-based approach, they concluded that attributing mental states to others or perspective taking is not necessary for passing false belief tasks. Again, instead of explaining the gradual development of false belief reasoning, this model explains an *'off-on'* like development.

More recently, Hiatt and Trafton (2010) constructed another ACT-R model regarding the development of first-order false belief reasoning. Their model uses Leslie and colleagues' (2004) dual-process approach, consisting of a theory of mind mechanism (ToMM) and a selection process (SP) mechanism. The model "watches" the first-order unexpected location FBT and explicitly notes what is happening during the story. While ToMM creates possible beliefs for the object's location, the SP mechanism of the model selects one of them. The model first selects its own belief without considering another agent's knowledge because of the higher utility value and after that, by using reward, the model learns to take into account the other's knowledge. They also manipulated parameters to show the gradual development of first-order false belief reasoning.

A Cognitive Model of Second-order False Belief Task

We used the hybrid cognitive architecture ACT-R (Anderson, 2007) to show the developmental transitions in false belief reasoning. Even though we used the second-order 'Chocolate Bar' story in our model, the model is not dependent on the particular features of this task. We are inspired by Taatgen and Anderson's (2002), and Van Rijn and colleagues' (2003) models and we use activation of declarative knowledge instead of utility learning to show the developmental transitions from zero-order to first-order and later to second-order reasoning. The reason is that we believe that reasoning about an answer to a question is a volitional process and that this type of reasoning cannot be explained in terms of utility learning, which corresponds more to implicit learning. More in particular, we assume that possible strategies are represented by declarative chunks, and that the model retrieves one of these chunks each time to solve a problem. Based on the success and/or feedback, the model will strengthen successful strategy chunks, or will add or strengthen an alternative strategy if the current one failed (see Meijering et al., 2013 for a similar approach).

First, we explain how the model goes through the developmental transitions and then we will explain the details of the model in the following subsection. Initially, in addition to the story facts, there is only a zero-order reasoning chunk in declarative memory. The representation of the chunk in declarative memory is as follows:

(reasoning-zero isa reasoning level 1)

This chunk is just a symbol in declarative memory and only represents the strategy to answer any object-location question with the location where they know the object is in reality. Other than representing a strategy, it has no role in the reasoning processes that are carried out by applying the production rules.

More specifically, after the second-order false belief question 'Where does Ayla think Murat will look for the chocolate?', the model automatically reasons about the real location of the object. Then, it requests the retrieval of a strategy chunk from its declarative memory in order to give an answer. Given that the zero-order reasoning is the only initial strategy, the model retrieves this chunk and gives an answer (toybox) to the second-order false belief question based on zero-order reasoning, which corresponds to the correct answer for the reality control question 'Where is the chocolate now?'. However, since it is not a correct answer to the second-order false belief question, the experimenter gives the feedback 'Wrong'. Given this feedback, the model increments the reasoning one level up and enters a new strategy chunk in declarative memory: a chunk that represents first-order reasoning. More in particular, this chunk represents the strategy in which the former (zero-order) strategy is now attributed to (in this example) Ayla.

However, since the model has more experience with zero-order reasoning, its activation at first is higher than the recently added first-order reasoning chunk. That is why the model retrieves the zero-order reasoning chunk instead of the first-order one for the next few trials and answers the second-order false belief question based on zero-order reasoning. Since the experimenter gives negative feedback and the model increments the zero-order reasoning one level up at each trial, the first-order reasoning chunk's activation value increases while the activation of the zero-order reasoning chunk decreases. When the activation value is high enough for its successful retrieval, the model gives an answer to the second-order false belief question based on first-order reasoning (toybox), which corresponds to the correct answer for the first-order false belief question 'Where will Murat look for the chocolate?'. Again, this is not a correct answer to the second-order false belief question. After the experimenter's feedback 'Wrong', again the model increments the reasoning one level up. This second-order reasoning now attributes the former first-order strategy to Ayla, and therefore now gives the correct answer (drawer). Given the positive feedback, the second-order reasoning strategy is further strengthened and finally becomes stable. In theory there is no limitation to the upper limit of the level of reasoning chunks. Nevertheless, in practice there is no need to use a very high level of reasoning, and even though one tries to apply more than third- or fourth-order reasoning, it will be very hard to apply that strategy in terms of memory. Note that in the 'Chocolate Bar' story, it is not possible to distinguish whether participants are applying zero-order or first-order reasoning, since both of them would yield the same answer (toybox). However, there are some modified versions of second-order FB stories in which the two reasoning levels can be distinguished (see Hollebrandse, in press).

Model Description

Before explaining the model in detail, it is important to explain the model's task-independent procedural knowledge to answer the second-order false belief question 'Where does Ayla think that Murat will look for the chocolate?' The task-independent procedural knowledge that the model implicitly knows comprises the following features: i) The difference between action verbs (e.g. 'to put') and perception verbs (e.g. 'to see'), ii) How to answer 'where' questions (e.g. retrieving from declarative memory a story fact that contains an action verb), iii) '*Seeing* leads to *knowing*', iv) People search for objects at the location where they have last seen them, v) The other people reason '*like me*'.

One of the assumptions of the model is that the experimenter told the story and only asks the second-order false belief question. Thus, the story facts are in the model's declarative memory and the model tries to answer only the second-order false belief question. Another assumption is that the model knows the temporal order of the story facts, meaning that it knows which events happened after or before a certain event. Since experimental studies exclude participants who give incorrect answers for the control questions from the data analysis, the retrieval threshold is set to a very low value, so that the model is always able to retrieve the story facts from declarative memory. Our model does not involve all of the sentences in its declarative memory but just the facts that are related to answering the control and false belief questions. These four story facts are as follows: i) "Murat put the chocolate into the drawer at time t1", ii) "Ayla put the chocolate into the toy box at time t2", iii) "Murat sees her at time t2", iv) "Ayla did not see him at time t2".

Let us describe the production rules in four steps to answer the false belief question[1]:

1. Retrieve a story fact that has an action verb in its slots.
2. Check the time slot of the retrieved story fact and if it is not the latest fact, request the latest one.
3. Request a retrieval of one of the reasoning levels from declarative memory.
4. If the k^{th}-order reasoning ($0<k\leq n$) is retrieved, determine whose knowledge the question is about and give the answer by reasoning as if that person employs $(k-1)^{th}$-order reasoning. Based on the success and/or feedback, the model will strengthen successful strategy chunks, or will add or strengthen an alternative strategy if the current one failed.

As can be seen from the above first two steps, the model always uses the same set of production rules before retrieving a reasoning level from its declarative memory. This feature of the model reflects the inevitable process of a person's reasoning from his/her own point of view (zero-order reasoning) even if the question concerns another person's belief, the so-called '*egocentrism*' or '*reality bias*' (Mitchell et al., 1996). After the model learns to inhibit its own answer (zero-order strategy) by retrieving one of the higher-order reasoning chunks, it gives an answer considering another person's belief. When we consider this feature of the model, which reflects simulation with the rules that were explained above in task-independent procedural knowledge, we can say that our model uses a *hybrid theory* approach to explain the development of false belief reasoning.

Another feature of the model is that when the model uses second-order reasoning, it also continues to use zero-order and first-order reasoning. More explicitly, when the model reasons about the second-order false belief question 'Where does Ayla think that Murat will look for the chocolate?' by using second-order reasoning, first it reasons about the *real* location of the chocolate (zero-order) by applying the first two steps of the explanation of the production rules that are stated above. After that, it reasons about Murat's knowledge about the location (first-order) by checking the slots of the story fact chunks whether Murat saw the latest action or not. Finally it reasons about Ayla's knowledge about Murat's knowledge about the reality (second-order) by checking the slots of the story fact chunks whether Ayla saw that Murat saw her or not.

Model Results

To assess the developmental transitions from zero-order to second-order reasoning, we ran the model 100 times, indicating that one child learns to apply second-order reasoning over time (years). To average the results across

[1] The code of the model can be downloaded from: http://www.ai.rug.nl/~barslan/styled/index.html

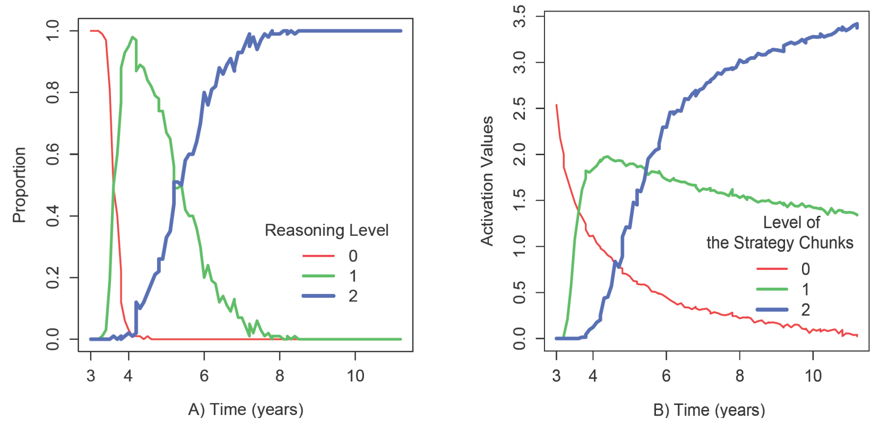

Figure 2: A) Proportions of the reasoning level the model applies, and B) The activation values of the strategy chunks

100 children, we made 100 simulations. Thus, we ran the model 10,000 times in total.The initial activation of the zero-order reasoning chunk is set to 6, indicating that children have a lot of experience with zero-order reasoning. Figure 2A shows the proportion of the levels of reasoning the model applies, and Figure 2B shows the activation values of the strategy chunks over time. As explained in the previous subsection, our model knows the rule 3, '*seeing*-leads-to-*knowing',* which is acquired by normally developing children around the age of 3 (Pratt & Bryant, 1990). So, the starting point of the x-axis of Figure 2 refers to 3-year-old children.

At the age of 4, the model learns to effectively use first-order reasoning (70%) instead of zero-order reasoning. Finally around the age of 6, the model starts to use second-order reasoning much more than (80%) first-order reasoning, and after the age of 8 second-order reasoning becomes stable (100%).

Discussion and Future Work

We constructed a computational cognitive model by using the cognitive architecture ACT-R in order to show the developmental transitions from zero-order to second-order false belief reasoning. Differently from previous ACT-R models, we did not introduce a new parameter in order to show the gradual developmental trend. Moreover, unlike Hiatt and Trafton's (2010) model, our model goes through the developmental transition not on the basis of utility learning but on the basis of the activation of the reasoning chunks in declarative memory. Utility learning does not have a method to introduce new strategies other than production compilation, which also requires that a strategy is represented in declarative memory first. Recently, Meijering et al. (2013) modeled reasoning about others in a sequential game without using utility learning but using the combination of declarative knowledge and problem solving skills. Their model's results nicely fit their previous empirical data.

Our model always starts by reasoning from its own perspective and if it learns to inhibit to give its own answer by retrieving a non-zero-order strategy from its declarative memory, then it reasons about another agent's belief by using its *a priori* task-independent knowledge. This feature of the model is important for two reasons. First, it explains the *'reality bias'* (Mitchell et al., 1996) of younger children by producing the model's own answer instead of taking into account others' knowledge. Secondly, unlike Bello and Cassimatis' (2006) pure rule-based model, it explains how the gradual development of false belief reasoning can be explained with the help of a *hybrid theory* by including the simulation and the rule-based approaches.

However, our model has production rules for reasoning about another agent's belief in advance. A more explanatory model for the development of second-order false belief reasoning would be a model that learns to construct new production rules for the higher levels of reasoning by itself. Another issue is that, unlike the behavioral data (Arslan et al., 2012), our model already becomes stable in two years (around the age of 8) when it learns to apply second-order reasoning correctly. The reason for that is that the model always gives a correct answer for the second-order false belief question when it retrieves the second-order reasoning chunk. However, different from first-order reasoning, we believe that successful second-order reasoning needs to overcome the constraints on serial processing involved in embedded structure of the beliefs (Verbrugge, 2009).

There are two important predictions of the model:

1. Children who are able to give correct answers for the first-order FB question but not for the second-order one do not give zero-order answers but first-order answers for the second-order FB question.

2. Children who are able to pass the first-order FB task learn to apply second-order reasoning with the help of feedback "Correct/Wrong".

A future step is to test these predictions by conducting a training study with 5-year-old children and to improve the current model so that it answers not only the second-order false belief question but also the ignorance question in order to explain the difference between them.

Acknowledgments: We are grateful to the Netherlands Organization for Scientific Research for Vici grant NWO-277-80-01, awarded to Rineke Verbrugge.

References

Anderson, J. R. (2007). *How Can the Human Mind Occur in the Physical Universe?* New York: Oxford University Press.

Arslan, B., Hohenberger, A., & Verbrugge, R. (2012). The development of second-order social cognition and its relation with complex language understanding and working memory. *Proceedings of the 34th Annual Conference of the Cognitive Science Society,* Austin, TX: Cognitive Science Society, 1290-1295.

Bello, P. & Cassimatis, N. (2006). Developmental accounts of theory-of-mind acquisition: Achieving clarity via computational cognitive modeling. *Proceeding of the 28th Annual Meeting of the Cognitive Science Society*, 1014-1019. Vancouver: Cognitive Science Society.

Flobbe, L., Verbrugge, R., Hendriks, P. and Krämer I., (2008). Children's application of theory of mind in reasoning and language. *Journal of Logic, Language and Information*, 17(4), 417-442. Special issue on formal models for real people, edited by M. Counihan.

Gopnik, A. & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language*, 7(1-2),145-71.

Hiatt, L. M. & Trafton, J. G. (2010). A cognitive model of theory of mind. In D. D. Salvucci & G. Gunzelmann (Eds.), *Proceedings of the 10th International Conference on Cognitive Modeling*, 91-96. Philadelphia, PA: Drexel University.

Hogrefe, G., & Wimmer, H. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, 57, 567.

Hollebrandse, B., Van Hout, A. & Hendriks, P. (in press). Children's first and second-order false-belief reasoning in a verbal and a low-verbal task. *Synthese*.

Keysers, C., Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Progress in Brain Research*, 156, 379–401.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*, 12, 528-533.

Liddle, B., & Nettle, D. (2006). Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology,4(3-4)*, 231-246.

Meijering, B., Taatgen, N., Van Rijn, H. & Verbrugge, R. (2013). Reasoning about mental states in sequential games: As simple as possible, as complex as necessary. *Proceedings of the 12th International Conference on Cognitive Modelling* (this volume).

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin,* 135, 749-773.

Miller, S. A. (2012). Theory of mind: Beyond the preschool years. New York, NY: Psychology Press.

Mitchell, P., Currie, G. and Ziegler, F. (2009), Two routes to perspective: Simulation and rule-use as approaches to mentalizing. *British Journal of Developmental Psychology*, 27, 513–543.

Mitchell, P., Robinson, E. J., Isaac, J. E. & Nye, R. M. (1996). Contamination in reasoning about false belief: An instance of realist bias in adults but not children. *Cognition*, 59, 1-21.

Onishi, K.H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs?, *Science*, 308, 255-258.

Perner, J. & Wimmer, H. (1985). "John thinks that Mary thinks that...": Attribution of second-order beliefs by 5- to 10-year old children. *Journal of Experimental Child Psychology*, 5, 125-137.

Pratt, C., & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development,* 61, 973-983.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 4, 515-526.

Raijmakers, M.E.J., Mandell, D.J., Van Es, S.E. & Counihan, M. (in press). Children's strategy use when playing strategic games. *Synthese*, special issue on reasoning about other minds, ed. R. Verbrugge.

Van Rijn, H., Van Someren, M., & Van der Maas, H. (2003). Modeling developmental transitions on the balance scale task. *Cognitive Science*, 27(2), 227-257.

Siegal, M., & Beattie, K. (1991). Where to look first for children's understanding of false beliefs. *Cognition*, 38, 1-13.

Sullivan, K. & Winner, E. (1993). Three-year-olds' understanding of mental states: The influence of trickery. *Journal of Experimental Psychology*, 56, 135-148.

Surian, L., Caldi, S., & Sperber, D. (2007) Attribution of beliefs by 13-month-old infants. *Psychological Science*, 18(7), 580-586.

Tager-Flusberg, H., & Sullivan, K. (1994). A second look at second-order belief attribution in autism. *Journal of Autism and Developmental Disorders,* 24, 577-586.

Taatgen, N.A. & Anderson, J.R. (2002). Why do children learn to say "broke"? A model of learning the past tense without feedback. *Cognition,* 86(2), 123-155.

Triona, L. M., Masnick, A. M., & Morris, B. J. (2002). What does it take to pass the false belief task? An ACT-R model. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (p. 1045). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Verbrugge, R. (2009). Logic and social cognition: The facts matter, and so do computational models. *Journal of Philosophical Logic*, 38 (6), 649-680.

Wahl, S. & Spada, H. (2000). Children's reasoning about intentions, beliefs and behaviour. *Cognitive Science Quarterly*, 1, 5-34.

Wellman, H., Cross, D. & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false-belief. *Child Development,* 72 (3), 655-684.

Wimmer, H. M. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–128.