SECOND-ORDER THEORY OF MIND: THINKING ABOUT THINKING … ABOUT THINKING

By

ELIF NUR POYRAZ

A thesis submitted to the

Graduate School – New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Psychology

Written under the direction of

Alan M. Leslie

and approved by

New Brunswick, New Jersey

January, 2021

ABSTRACT OF THE THESIS

Second-order Theory of Mind: Thinking about thinking … about thinking

by ELIF NUR POYRAZ

Thesis Director:

Dr. Alan M. Leslie

Preschoolers' performance in the standard false belief task has been shown to be correlated with working memory and inhibitory control tasks. The Theory of Mind Mechanism (ToMM) account (Leslie, 1987; Leslie, 1992, 1994) posits that inhibition plays a role in the expression of theory of mind understanding (Leslie & Polizzi, 1998; Leslie, German, & Polizzi, 2005), and that the difference in performance between the ages of three and four can be explained by changes in children's inhibitory control. There is also evidence that children can track multiple minds each with distinct false beliefs (Cheng, 2018). Specifically, they found that 4-year-olds can track up to 4 minds, showing adequate working memory resources at the age of 4. In another vein, an unexplored part of theory of mind research is on how children understand second order false beliefs (FB): "Anne believes that Sally believes the marble is in the box.". Whereas initial studies suggested success around 6.5-7 years of age (Perner & Wimmer, 1985), simplified versions found success around the age 5.5 by reducing processing demands (Sullivan et al., 1994). It is still an open question what specifically those processing demands are. The initial aim of the current project is to replicate Cheng's (2018) findings, and as the main aim, we planned on testing children by manipulating two demand factors: working memory (by including more than a single agent in the story), and inhibitory control (by having low and high demand levels). However, the current project only contains data for 4- and 5-year-olds' performance in a double agent *low inhibitory* demand task. The results from Experiment 1 shows that five-year-olds were more successful compared to four-year-olds in both first and second-order false belief understanding. However, due to concerns about certain details of the stimuli and the delivery of the task in Experiment 1, Experiment 2 was conducted. Results from Experiment 2 were almost identical to Experiment 1, showing that the potential limitations thought to be affecting results in Experiment 1 were not influential in children's performance. Furthermore, both in Experiment 1 and 2 the results show a pattern that when children are successful in the first order questions, they also tend to be successful in the second order question. Further directions and need for testing a high inhibitory demand condition are discussed.

Acknowledgements

I would like to thank my advisor, Dr. Alan Leslie, who has provided immense support through the rough bits of this thesis; and also, to my committee members, Dr. Pernille Hemmer and Dr. Judith Hudson. Without the participating families and the curious minds of sweet preschoolers, this research wouldn't be possible! I'm thankful to the members of the Cognitive Development Lab for their support, and to Dr. Michelle Cheng, who has graciously shared her stimuli (and also guidance) for the Experiment 1. Lastly, my special thanks to Ashley Morales, who helped me with the native narration in Experiment 2.

Table of Contents

List of Figures

Chapter 1

Introduction

> *"What a nice vibe this place has!"*
>
> You are thinking to yourself, while you are having a nice cup of americano in a cozy little coffee shop with your friend Sammy. He's obviously ecstatic because he just *cannot* stop talking about that new post-modern art exhibit in MoMA he is going to that evening with a mutual friend, Anne. After you depart, you are by yourself walking home when you suddenly remember you actually ran into Anne earlier that day. You explicitly remember Anne's worried face. She told you her mom broke her arm:
>
> *"I'm on my way to go there now. I really hope she'll be fine until I get there, it's a 2-hour drive and she's all alone."*
>
> You realize that with all that fluster going on, Anne must have forgotten to cancel her plans with Sammy.
>
> *"Oh God, Sammy doesn't know that Anne won't make it. He's going to be crushed to find out. He was so eager to check this once-in-a-long-time exhibit with her."*

All of our social life, as well as enjoying fiction, depends on our ability to successfully reason about other people's minds. Premack and Woodruff (1978) coined the term "theory of mind" to refer to this ability. Theory of mind (ToM) allows us to make sense of others' observable behavior by inferring what might be the underlying mental state, such as believing, intending, knowing, desiring, or pretending, that could have produced such behavior. Despite the fact that Premack & Woodruff's study was focusing on chimpanzee's ToM, this complex ability quickly sparked interest in the cognitive developmental field, as researchers were interested in the roots of this ability and the underlying cognitive mechanism.

The standard false belief task (i.e., Sally-Anne task) became the most commonly used measure of ToM in the cognitive developmental literature and has been in use for the last 35 years (Baron-Cohen, Leslie, & Frith, 1985; also see the earlier version by

Wimmer & Perner, 1983). In this task, an agent, Sally has an object (e.g., marble), she hides it in Location A (e.g., basket), and then leaves the scene. While Sally is gone, another agent, Anne, removes the object from Location A and hides it in Location B (e.g., box). Children are then asked: "When Sally comes back, where will Sally look for her marble?". A persistent finding obtained using this paradigm is that, while the majority of the children older than 4-years-old are successful in predicting "Sally will look for the marble in the basket.", majority of the children younger than four respond with the current location of the object and report that "Sally will look for the marble in the box." (Wellman, Cross, & Watson, 2001).

Although the findings are regarded as well-established, the disputes between the theoretical accounts of the contributing factors to performance in the standard task are yet to be resolved. One theoretical account proposes that children acquire an adult-like concept of 'beliefs' around their fourth birthday (e.g., Perner, 1991). According to this theory-theory account, just as how scientists would revise their theories by trial and error, children build up the concept of beliefs through many instances of being involved in and surrounded by the social world (Gopnik & Astington, 1988; Gopnik & Meltzoff, 1997; Gopnik & Wellman, 1994; Perner, 1991; Perner, 1995; Perner & Ruffman, 2005; Wellman, 1990, 2002; Wellman et al., 2001). Reasoning about beliefs, especially false beliefs, requires the ability to represent others' mental contents and the fact that those contents can be different from one's own beliefs or the current reality. Before such concept acquisition, the representational power of ToM is limited and therefore, most of the three-year-olds failing to attribute a false belief (FB) in the standard task is dubbed as a conceptual deficit in their ToM.

In contrast, another account by Leslie and collegues proposes that a domain-specific modular mechanism, ***T**heory **o**f **M**ind **M**echanism (ToMM), underlies the children's ability to reason about others' minds (Leslie, 1994a, 1994b, 1992, 2000a; German & Leslie, 2000, 2001; Scholl & Leslie, 1999, 2001; Leslie & Thaiss, 1992). This theory postulates that ToMM comes online from infancy onwards, meaning that younger children in fact *do* have the competence to understand and reason about mental states. ToMM enables children to first attend to mental states by providing the organism with the 'initial push', i.e., with placeholders for mental statess like believing or desiring, in a way that enables children to later learn about their properties (Leslie, 2000b; German & Leslie, 2000, 2001). However, the standard FB task puts extra processing demands on the young brain, which exceeds their executive resource limits (especially inhibitory control) and masks the expression of an otherwise competent ToM (Leslie, Friedman, & German, 2004; Leslie, German, & Polizzi, 2005; Leslie & Polizzi, 1998).

1.1. Performance Factors in the Standard False Belief Task

Some researchers have suggested that younger children's limited working memory (WM) capacity might be playing a factor in their poor(er) performance in the standard FB task (Davis & Pratt, 1995; Gordon & Olson, 1998; Keenan, 1998; Keenan, Olson, & Marini, 1998; Olson, 1989). In the standard task children need to hold in mind both the representation of the current reality and Sally's false belief about the location of the object. Working memory accounts of false belief reasoning proposes that younger children's poor performance is related to their still-developing working memory capacity, which allows them to be able to only hold onto one representation – perceptual representation of the current reality that states 'The marble is in the box.'. Studies have found that WM tasks such as Counting & Labeling task (Gordon & Olson, 1998), the Counting Span task (Keenan et al., 1998), and Forward & Backward Digit Span task (Davis & Pratt, 1995) were significantly correlated with performance in the FB tasks. However, the studies arguing for a profound impact of WM on ToM performance is limited in their causal explanatory power because WM measures are only correlational to the ToM tasks, and they are not experimentally manipulated in the FB task itself. One researcher, Cheng (2018), have carried out such experimental manipulations in the standard FB task, and their findings will be discussed in Section 1.3.

Younger children's tendency to respond with the current reality in the FB tasks can also be explained by insufficient inhibitory control resources. Inhibition as an executive process has become a focus that researchers argue to be predominantly affecting children's performance on the standard FB task (e.g., Carlson, Moses, & colleagues; Leslie & colleagues). One proposed argument is that inhibitory control (IC) enables children to be able to *entertain* other's mental states, thereby filling in a crucial role in their *acquisition* of the concept of belief (Carlson & Moses, 2001; Carlson, Moses, & Breton, 2002; Carlson, Moses, & Hix, 1998). These researchers suggest that sufficient inhibitory control helps children in learning about false beliefs, thus facilitates both the acquisition (emergence) of the concept of belief, and also the expression of a matured concept of belief while performing in the standard FB task (Moses, 2001). Their findings suggest that performance in conflict inhibition tasks, which measure children's ability to inhibit a salient response, have found to correlate with, and further predict later ToM even when controlled for other factors such as general intelligence (Carlson et al., 2002) and verbal ability (Carlson & Moses, 2001). These findings are promising in further understanding relationships between IC and ToM; however, the exact interaction of them and how exactly IC works in the ToM mechanism can only be understood via manipulations of IC within a FB task.

Importantly, a more specified account, ToMM theory (initially described in the previous section), is able to predict, and explain in more detail how inhibitory control demands are crucial in being successful in the standard FB task (Leslie et al., 2004, 2005; Leslie & Polizzi, 1998). The next section highlights the details of the ToMM theory and how children's performance is affected when IC is manipulated within FB tasks.

1.2. Theory of Mind Mechanism + Selection Process Model

Theory of Mind Mechanism + Selection Process (TOMM + SP) model posits that in order to be successful in the FB tasks, children, as well as adults, not only need to be able to reason about beliefs of others, but also they need to be able to successfully select the correct belief attribution among multiple belief candidates. In the standard FB scenario, ToMM spontaneously calculates the True Belief (TB) candidate (based on the object's current location) and also the FB candidate (based on Sally's lack of visual access to the scene). Then, these two candidates need to be evaluated. By default, ToMM attributes a TB because beliefs ought to be and generally are true (Dennett, 1989)[1]. In order to be successful in the FB task, this prepotent response needs to be *inhibited*, resulting in the attribution of the correct belief candidate, FB, as a response to the "Where will Sally look?" question. The inclusion of this extra mechanism, Selection Processing (SP), is implemented through inhibition (Leslie, 2000b; Friedman & Leslie, 2004a, 2004b, 2005).

---

[1] For further discussion of how knowledge about the true state of the world affects performance see Birch & Bloom, 2003, 2007.

Empirical evidence has demonstrated that when the overall demands of the task is sufficiently reduced, younger children's performance is helped. For instance, Setoh, Scott and Baillargeon (2016) showed that 2-and-a-half-year-olds can succeed in a traditional FB task. When the target object (e.g. marble) is moved to an unknown location at the end of the story, children's performance improved. Because children do not have a certain idea about the final location of the object – i.e. true state of the world/reality, this reduces the saliency of the true belief, making it easier to inhibit (Setoh et al., 2016; also see a direct replication by Grosso et al. 2019). Furthermore, other research have landed further support to improvements in younger children's performance with following manipulations; ***i)*** when children are only told about the location of the object but they themselves didn't witness it being moved to the final location (Carlson & Moses, 2001; Zaitchik, 1991); ***ii)*** when the object being moved is made out of 'Plasticine' and while being moved, it is shaped into an apple however it was initially looking like a hat (Kikuno, Mitchell, & Ziegler, 2017); and, ***iii)*** children see that both hiding locations are in fact empty while someone is announcing their belief about an object's location (Bartsch, 1996). The overarching theme in these experiments is that children's certainty/confidence about the current location of the object is reduced, and because the necessary IC that needs to be deployed to overcome the TB default is also reduced, children's performance is helped (Wang & Leslie, 2016). Another work shows that the minimal manipulation of asking children "Where will Sally look *first* for her marble?" improves younger children's performance (Siegal & Beattie, 1991; Surian & Leslie, 1999; Yazdi et al., 2006). It is argued that this question *increases* the salience of the initial, false belief, location as a probable candidate for Sally's belief, thereby reducing the need to deploy a higher level of inhibitory control (Surian & Leslie, 1999; Leslie, 2000b).

Similarly, older children who are otherwise successful in the standard task can fail at ToM tasks that require higher inhibitory control (Friedman & Leslie, 2004a, 2004b, 2005); and elderly individuals who have declined inhibitory control skills have shown to have difficulty in FB tasks (German & Hehman, 2006). In addition, research done with individuals with Autism Spectrum Disorder (ASD) show specific impairment in their ToM performance (Baron-Cohen et al., 1985; Leslie & Frith, 1988; Leslie & Thaiss, 1992; Roth & Leslie, 1998) and these individuals are not helped with manipulations mentioned above (Surian & Leslie, 1999).

These empirical findings land further support to existence of an innately specified, domain specific module, ToMM, which spontaneously computes mental states to explain and predict observable behavior (Leslie, 1994a, 1994b). This mechanism comes online sometime within the second year of life which is first observed by the emergence of pretend play behavior (Leslie, 1987, 1994a; Bosco, Friedman, & Leslie, 2006; Onishi, Baillargeon, & Leslie, 2007). For instance, when a child pretends to be drinking tea from a toy cup, similar to reasoning about false beliefs, they need to maintain two different representations about the world; the current state of the world "CUP IS EMPTY", and the pretend representation "CUP HAS TEA INSIDE" (Leslie, 1987, 1994a).

Furthermore, the child needs to have a higher-order understanding (meta-representation) that these two representations are separate/different from each other – one pertains to reality (Leslie, 1987). Of more importance, around the same time children produce such pretending behavior, they are also able to recognize pretense in others.

When the very same child sees their caregiver holding the same cup and pretending to drink from it, rather than being confused by this bizarre behavior of holding an empty cup and trying to drink from it, they are able to recognize this pretend behavior and join the pretend scenario (Leslie, 1987). The advent of pretend play is crucial in theory of ToMM because, in contrast to 'theory-theory' account that proposes a lack of conceptual and representational understanding of mental states, theory of ToMM proposes a representationally intact repertoire for understanding mental states from a very early age.

1.3. Multiple Agent Tracking

Although 'passing' the FB task could be interpreted as an indicator of a conceptual understanding of beliefs, it is not clear what 'failing' suggests. Dichotomous nature of the standard task is a major limitation for researchers that try to address the more intricate aspects of ToM performance. It creates an 'illusionary' 50% chance-level, which obscures further information that could be otherwise obtained from error analyses. Indeed, the manipulations that are empirically shown to be helping younger children only increase average performance to around 50% success, which is deemed to be insufficient of an increase for researchers that argue for a conceptual deficit account in younger children since the performance is still around 'chance' levels (Wellman et al., 2001; for discussions of this interpretation see Scholl & Leslie, 2001; and Moses, 2001)[2].

Coupled with this issue is also the fact that in real life belief-desire reasoning, there is rarely an instance one reasons only about a single person's mental states. In a regular social environment, children are at least surrounded by their small family groups, which

---

[2] For a discussion of whether it is still appropriate to use the standard FB task as a measure of ToM, see Bloom & German, 2000.

would require them to track distinct beliefs and desires of multiple people/agents (Cheng, 2018). Despite this fact, decades of research in ToM only looked into how children reason about a single agent's mental states, namely, Sally's. How about Anne, or a third, fourth, even a fifth person? Cheng (2018) in their studies aimed to address this gap along with the goal of testing the effects of working memory on the standard FB task by manipulating it within the actual task.

Cheng (2018) used variant versions of the standard FB task to test three- and four-year-olds. They manipulated both the working memory (WM) demand and the inhibitory control (IC) demand. For the former manipulation, instead of having only two agents/actors (i.e., Sally and Anne), the number of agents in the stories ranged from two up to five. This allowed for bypassing the correlational nature of WM-ToM studies by embedding the working memory demand into the false belief task. For the latter, the manipulation was to remove the object of desire to a vague, unknown location at the end of the story, which is shown to be helpful in reducing the IC demand in the FB task. As mentioned in the previous section, this manipulation increases performance in younger children by reducing the saliency of the True Belief (Wang & Leslie, 2016). Importantly, every agent that was added to the story had a distinct belief and all had false beliefs. For instance, in the triple-agent, low-IC demand task, there are three agents; Sammy, Aaron, and Jake. Sammy has a pet chick, he puts it in the *basket* and goes away. While Sammy is gone, Aaron takes the chick, plays with it and puts it in the *box*. Then, Aaron also goes away. While both Sammy and Aaron are gone, Jake takes the chick out of the box and puts it in the *chest*. Then, Jake also leaves the room. After all three agents are gone, the chick comes out of the chest and it is described as "… it jumped away, far far away, where nobody knows where it went". This last portion, where the chick jumps away and goes off-scene, makes the current location of the chick less certain – i.e., the true belief less salient, reducing the IC demand of the task. In the high IC demand version, the chick is described instead as jumping into another location, a jar located in the room, which makes the last location of the chick highly salient and certain.

These manipulations allowed children to produce error patterns in multiple ways unlike the standard FB task, although children could still be successful in a single way by attributing the correct false beliefs to the corresponding agents. Cheng (2018)'s findings showed that three-year-olds in fact can track up to 3 agents, and four-year-olds can track up to 4 agents, each with distinct false beliefs. Furthermore, what she observed in children's responses was that, although children could make errors in various possible ways, mainly two response types emerged; ***i)*** attributing the correct false beliefs to all agents, and ***ii)*** attributing the true belief to all agents. This pattern was observed across all levels of WM and IC demand manipulations, suggesting that WM plays a minimal role in the standard task, while the True Belief default has a heavy impact in false belief understanding and IC is a major contributor to performance in preschoolers' ToM performance (Cheng, 2018; Leslie et al., 2005; Leslie & Polizzi, 1998).

1.4. Second Order False Beliefs

When children are asked, "Where will Sally look for her marble?", they are representing only Sally's belief: "Sally believes that the marble is in the basket.". Since this requires reasoning about only a single person's mental state, a mental state (e.g., false belief) which does not mention another mental state, this is called reasoning about a *first order false belief.* However, in daily life children may also need to reason about what a person thinks about what another person knows, thinks, believes, or wants. As a reference to the standard FB task, understanding that "Anne believes that Sally believes the marble is in the box." is indeed reasoning about second order beliefs. If this feels too isolated as an example, think back to the scenario at the beginning of this chapter. As the reader, from the first-person perspective, you were thinking about your friends Sammy and Anne. You thought to yourself "Sammy doesn't know that Anne is thinking that she will go and see her mom.", or you could think "Sammy doesn't know that *I* know Anne won't make it.". These are all instances we are familiar with and can relate to in our daily lives and shows how mental state reasoning can get complicated very quickly.

The very first study which examined children's second order FB understanding was carried out by Perner and Wimmer (1985). They used 'the ice-cream truck story' to tell about a scenario in which two kids, Mary and John, wanting to get ice cream. In this story, initially Mary knows the ice cream truck to be in the park but goes back home to fetch her wallet. Later, John also goes to the park and sees the ice cream truck but unfortunately the truck is leaving the park. The truck passes by Mary's house, so Mary knows that the truck changed location. However, John doesn't know that Mary knows the ice cream truck has left the park to go to church. Children are then asked; "So John runs to look for Mary. Where does he think she has gone?" (Perner & Wimmer, 1985). Using this paradigm, Perner and Wimmer found that children who are almost all seven-year-olds and many six-year-olds were able to reason about second order false beliefs (1985). This indicated that preschool children (up until roughly 5,5 to 6-years-old) were not able to reason about second-order false-beliefs. Based on these findings Perner (1988) argued that children might indeed have another *conceptual milestone* in their mental state reasoning. Indeed, Perner (1988) stated that:

> "… second-order state attributions require understanding of the recursive nature of mental states (repeated embedding of propositions). It is plausible that this possibility for recursion is understood at a particular point in development…" (p.273)

It is unclear why there would need to be another milestone in which children acquire the concept of 'recursion', where in fact the very structure of the first order FB understanding relies upon effective embedding of propositions – "Sally thinks *that* the marble is in the box". Children by the age of 7 are competent in representing mental states of others, so why is there a developmental delay where children come to be passing the second order false belief tasks?

Sullivan, Zaitchik and Tager-Flusberg (1994) argued that the original 'ice cream truck story' could be too complicated as a task to measure the real second order FB competence since it included many different story episodes with many details to keep track of. They argued that this task was too high in 'processing demands' which could obscure an otherwise successful second order FB reasoning. Indeed, their arguments were supported with their findings that by the age of 5,5, 90% of preschoolers they tested were successful in their simplified version of the story.

There have been a few studies looking into children's second order FB understanding. Miller (2009, 2013a, 2013b) suggested that inclusion of deceptive intent in the second order FB stories is a facilitative factor in children's performance. They argued that deception highlights the existence of one character's lack of knowledge about the other character's information on the true state of the world; thus, this highlight facilitates performance. One study by Arslan, Hohenberger and Verbrugge (2017) looked into how syntactic recursion and working memory plays a role in second order FB. They tested children's performance on separate syntactic recursion, and simple and complex span tasks (as WM measures). Their findings suggested that although syntactic recursion was correlated with second order FB performance, working memory was the main predictor of performance (Arslan et al., 2017). Another study proposed that 5-year-olds' errors in second order FB tasks could be explained by whether they have selected the correct strategy in their first order FB reasoning (Arslan, Taatgen, & Verbrugge, 2017). They argued that with enough training or exposure, children could revise their strategies to get better at second order FB. Interestingly, another study using both a verbal and a low-verbal second order task found that 7-year-olds performed better in the verbal second order FB task compared to low-verbal task (Hollebrandse, van Hout, & Hendriks, 2012). Authors concluded that language might be supporting the *explicit* reasoning about second order false beliefs.

1.5. The Present Study

Taken together, studies to date do not provide a systematic manipulation of potentially relevant cognitive processes (e.g., WM or IC) within a second order FB task. All of the studies mentioned above indicate a direction in which second order FB reasoning could be more effectively understood via exploration of how other cognitive processes are interacting with it, rather than a proposal of another conceptual milestone (ala Perner, 1988).

The overarching theme of the current project is to parse through which performance factors are affecting preschoolers' second order false belief understanding.

Following from the arguments of the ToMM + SP model (Leslie, 1994a; Leslie et al., 2005), second order FB task should not impose a representational limitation to children older than four, which would be apparent in their success in the first order FB task. Furthermore, as Arslan et al.'s (2017) study has suggested, WM memory might be playing a more profound role in second order FB performance compared to obtaining success in the first order FB (standard FB task), which was shown in Cheng's (2018) studies. However, how reasoning about second order FBs interact with the True Belief default is yet to be explored. The experiments in this thesis provides the first steps of testing both WM and IC in second order FB tasks, by implementing adapted stories from Cheng (2018). We are going to be manipulating WM demand by including more than 1 agent. Furthermore, we planned on testing children also by manipulating IC demand by having a low and a high demand condition. However, due to certain concerns about the stimuli and the narration of the task in Experiment 1, the rest of the project focused on addressing those issues. As a result, only the low IC demand data is available in this thesis[3]. Further directions in research will be discussed in the end.

[3] Additionally, there were some persisting obstacles before the submission of this thesis. Firstly, our laboratory was closed due to building infrastructure problems for weeks on end, and then COVID19 has impeded our work for the last 9 months, many research-related activities has been suspended, and, unfortunately, the most negatively impacted was the data collection processes.

Chapter 2

Experiment 1

2.1. Method

2.1.1. Participants

A total of 60 children participated in the current experiment: 47 were 4-year-olds (M_{age} = 54.3 months, SD = 3.39; 19 females) and 13 were 5-year-olds (M_{age} = 63 months, SD = 2.23; 9 females). An additional 11 children were tested but excluded from the final sample due to following reasons: distraction/interruption (4), parent didn't consent for video recording (2), subject didn't respond to test questions (2), had participated in a similar experiment before (1), technical error (1), and experimenter error (1). The general exclusion criteria were predetermined before data collection, and the decision to exclude was taken by the author who collected the data. All the subjects were recruited from preschools in central New Jersey area. Subjects were either tested in the preschools or in the Cognitive Development Lab at Rutgers University, after obtaining parental consent. A verbal assent was taken from each subject prior to any testing session by asking "I have a really short and fun story to tell you today. Do you want to listen to it?". If the subject said "no", then no testing was done. Subjects received small thank you gifts (e.g. stickers and a kid-friendly certificate indicating their participation). If the subjects were tested in the lab, in addition to gifts given to subjects, families received a thank you gift (e.g. t-shirt or bib) as well as a small compensation for their travel.

2.1.2. Materials and Procedure

The current task was developed by Cheng (2018) and it is a version of the Standard False Belief Task with updates to accommodate demand manipulations (i.e. working memory and inhibition). The current task corresponds to their double-agent low

IC demand task. It is carried out as a story that is introduced to the subjects with a set of 9 pictures within a binder that follows a preset narrative. Each picture corresponds to an event that is taking place within the story. Sample layout of the stimuli that accompanies the stories can be seen in Figure 1. Before the story started, subjects were first introduced to the agents (characters) in the story with the use of paper cut outs of each agent. Then, they were asked to tell who is who, to ensure they understood and can remember the names of each agent (e.g. "Can you show me which one's Sammy?").

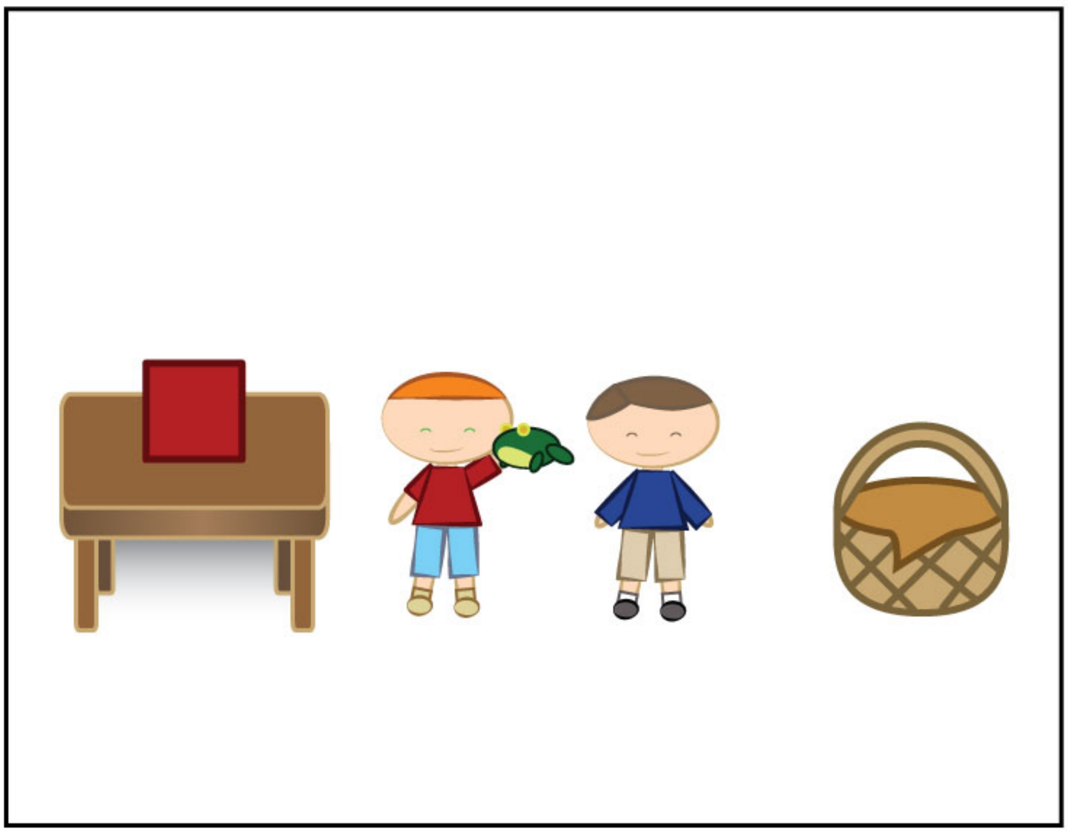

Figure 1. Sample story layout. Sammy (in red shirt, holding the frog), and Jack. A box on the table, basket on the right side.

Afterwards, the story started. The story script is as follows:

> "This is Sammy, and this is Jack [Pointing to agents]. Oh look! Look what Sammy has in his hands! Yes, it's a frog! Sammy was playing with his frog, but now Sammy has to go away. Sammy puts his frog into this box over here, then Sammy

> goes away. Now, Jack wants to play with the frog. Jack goes over to the box and takes the frog out. But now, Jack needs to go away, and then Jack puts the frog into the basket! Did Sammy see that? No, that's right! Then, Jack goes away. But, look what happens while the two kids are gone! The frog comes out of the basket, and then it jumps *away*, *far far away*, where no one knows where it went! Did Sammy see that? No, that's right! How about Jack, did Jack see that? No, that's right!".

If subjects said 'yes' to any of the See Questions (e.g. "Did Sammy see that?"), they were corrected by saying "Sammy wasn't in the room, right? He can't see that!". After the story ended, children were asked several control questions to ensure that they followed the storyline correctly, namely the *Memory Questions* ["Where did Sammy (or Jack, both asked) put the frog?"] and the *Reality Question* ["Where is the frog right now?"]. After the control questions, test questions were asked. First Order False Belief (FB) questions were asked first: "When Sammy comes back, where will Sammy look for the frog? / Where will Jack look for the frog?". Afterwards, the Second Order Question was asked "Where does Jack think Sammy will look for the frog?".

With each Memory Control and First Order FB question, the appropriate cut-out of the agent was brought back and was placed right in the middle of the story visuals (equal distance to box and the basket). These cut-outs were used to aid children in remembering the characters in the story. Since performance in recall of characters was not of interest for the purposes of the current study, making sure that children were able to correctly identify who was who while answering the study questions were important for making the task less taxing. If the subject failed any memory questions, the story was repeated only once starting from the very beginning. If the subject failed any memory questions after the story was repeated, no further test question would be asked, and the testing session would be terminated. However, none of the children gave incorrect answers to the memory questions after the second narration. If the subject was hesitant in giving an answer to the reality question and/or pointed to an irrelevant location (either on the story visuals or within the actual physical room the testing was taking place), they were asked "Do you remember what happened to the frog?". Since the manipulation that the frog jumps away to an indefinite location at the end of the story is vital for the experiment, this further question enabled us to make sure the subject understood/followed that portion of the story correctly in the case they were indecisive or confused. Since the last location of the frog is vague, when asked the reality question, children generated various locations themselves that are not mentioned in the story at all (e.g., 'it's in the pond', 'it's in the forest') or sometimes just responded saying 'I don't know'. Since indeed frog is described as going somewhere that nobody knows, this response as well as the self-generated responses are all counted as true belief responses.

Where the object was hidden first (box or basket) was counterbalanced across all subjects, as well as the order of the Memory Control and the First Order FB questions. Each session was videotaped. Two coders watched videos for all subjects, and recorded answers to the questions.

2.2. Results

Figure 2 shows 4-year-olds' responses. Each square (or hexagons for the 5-year-olds) indicates a single response to one of the three belief questions. Each column represents the response sequence of an individual subject. The top two rows are responses to the first order FB questions, whereas the bottom row (third row, below the straight line) indicates responses to the second order FB question. For the first order FB questions 17 out of 47 four-year-olds correctly attributed both agents their respective ("fully correct") false belief (binomial 1/9, p <.001; BF_{10}= 4346.039) while 22 out of 47 four-year-olds incorrectly attributed both agents a true belief (binomial 1/9, p <.001; BF_{10}= 2.633e +7). For the second order FB questions, which can be observed in the bottom row in Figure 2, 15 out of 47 four-year-olds correctly responded to the second order FB question (binomial 1/3, p = 1; BF_{10} = .172) while 21 out of 47 four-year-olds incorrectly attributed a true belief (binomial 1/3, p = .121; BF_{10} = .663).

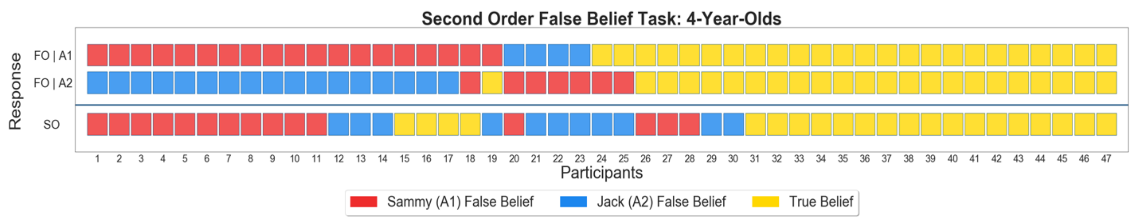

Figure 2. 4-year-olds' individual responses for Experiment 1. Each column represents the response sequence of an individual subject. FO rows marks the responses to the first order FB questions; FO | A1 row indicates the belief attributed to Agent 1; FO | A2 indicates the belief attributed to Agent 2. SO row indicates the responses to the Second Order question. Correct response sequence is FO | A1 = red; FO | A2 = blue; SO = red.

Figure 3 shows responses for 5-year-olds. For the first order FB questions, 9 out of 13 five-year-olds correctly attributed both agents their respective ("fully correct") false belief (binomial 1/9, p <.001; BF_{10}= 62047.997) and 4 out of 13 five-year-olds attributed a true belief to both agents (binomial 1/9, p = .048; BF_{10}= 1.892). For the second order FB questions, which can again be observed in the bottom row in Figure 3, 8 out of 13 five-year-olds correctly responded to the second order FB question (binomial 1/3, p = .4; BF_{10} = 2.727) while 3 out of 13 five-year-olds incorrectly attributed a true belief (binomial 1/3, p = .564; BF_{10} = .388).

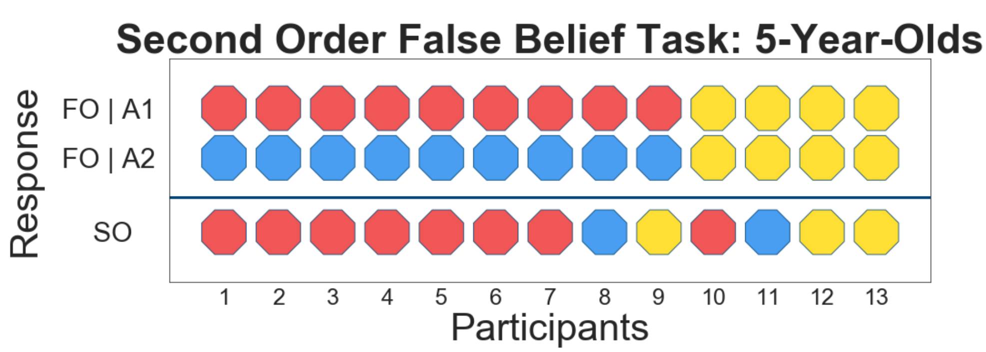

Figure 3. 5-year-olds' individual responses for Experiment 1. Each column represents the response sequence of an individual subject. FO rows marks the responses to the first order FB questions; FO | A1 row indicates the belief attributed to Agent 1; FO | A2 indicates the belief attributed to Agent 2. SO row indicates the responses to the Second Order question. Correct response sequence is FO | A1 = red; FO | A2 = blue; SO = red.

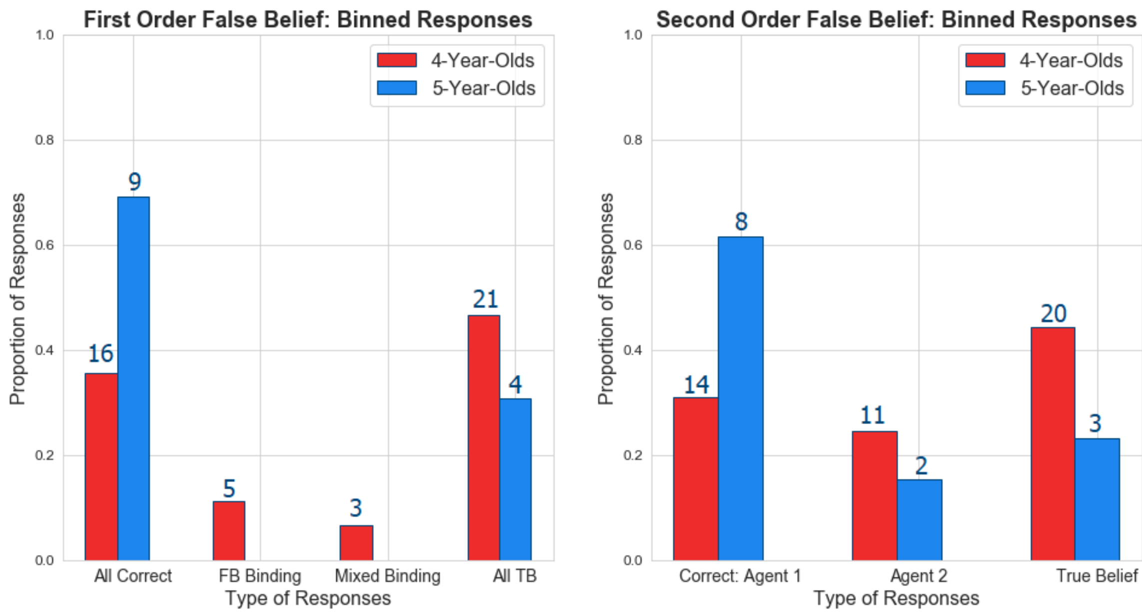

Figure 4. Proportion of binned responses for Experiment 1. Responses were categorized into one of the following four categories: All Correct False Belief Attribution (All Correct), False Belief Binding Error (FB Binding), Mixed False Belief- True Belief Binding Error (Mixed Binding), and All True Belief Error (All TB). The frequency of each response is displayed above the respective bar.

Although there were nine possible response patterns, not all patterns were observed in the answers to the first order FB questions. Following the criterion in Cheng (2018), the responses are grouped into four main categories: ***i)*** All Correct FB attribution, this is when subjects attribute both agents the correctly corresponding FBs; ***ii)*** False Belief Binding Error, this is when subjects attribute false beliefs to both agents but not to the correct corresponding agent (i.e., attributed Sammy's FB to Jack, Jack's FB to Sammy); ***iii)*** Mixed FB-TB Binding Error, this is when subjects attribute a TB to one agent, and a FB to the other agent; and finally ***iv)*** All True Belief Error, this is when subjects attribute the TB to both agents – i.e. where the frog really is. The All Correct FB attribution is the only response sequence that is correct, and the remaining eight possible response sequences are all incorrect. Figure 4 left panel shows the proportion of each category both for 4- and 5-year-olds. Figure 4 right panel shows the proportion of different responses to the second order FB question. Subjects could respond with any possible three belief candidates; Agent 1's FB, Agent 2's FB, and the TB. Although the number of subjects in the two age groups are different, general trends in proportions show that five-year-olds had a higher All Correct FB attribution rate and were providing the correct response for the second order question more frequently.

Further analyses excluding children who were not successful in the first order FB questions yielded that 11 out of 17 four-year-olds were successful in second order FB (binomial 1/3, p = .009; BF_{10} = 9) (see Figure 5). With the same exclusion criteria, 7 out of 9 five-year-olds were successful in second order FB (binomial 1/3, p = .008; BF_{10} = 13.67) (see Figure 6).

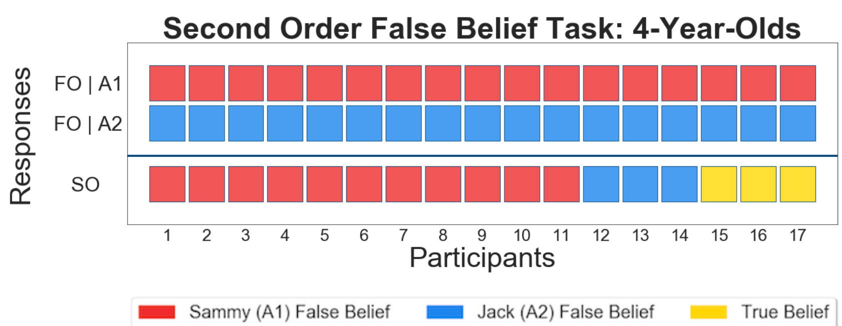

Figure 5. 4-year-olds' individual responses for Experiment 1; showing only the subjects who were successful in the first order FB questions. Each column represents the response sequence of an individual subject. FO rows marks the responses to the first order FB questions; FO | A1 row indicates the belief attributed to Agent 1; FO | A2 indicates the belief attributed to Agent 2. SO row indicates the responses to the Second Order question. Correct response sequence is FO | A1 = red; FO | A2 = blue; SO = red.

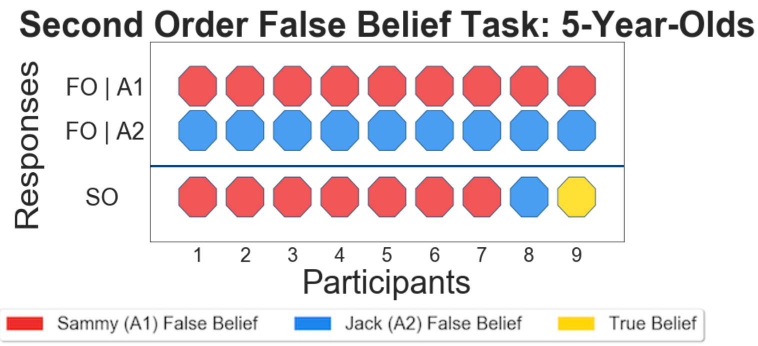

Figure 6. 5-year-olds' individual responses for Experiment 1; showing only the subjects who were successful in the first order FB questions. Each column represents the response sequence of an individual subject. FO rows marks the responses to the first order FB questions; FO | A1 row indicates the belief attributed to Agent 1; FO | A2 indicates the belief attributed to Agent 2. SO row indicates the responses to the Second Order question. Correct response sequence is FO | A1 = red; FO | A2 = blue; SO = red.

2.3. Discussion: Exp 1

Compared to Cheng (2018)'s findings in the double agent low IC demand task, the four-year-olds' performance in the first order FB questions were on the lower end. Cheng (2018) found that 32 out of 50 (64%) four-year-olds were successful in the first order FB questions; whereas in the current experiment we have observed that 17 out of 47 (36%) were successful. Since first order FB performance is essential in interpreting the second order FB in a sound manner, this raised a need to conduct some follow-up studies. Firstly, the set of visuals being used might have confusing details for children. The object being tracked is very important for this task. It might be the case that the picture of the frog used in the story visuals was too ambiguous and this created an extra confusion in children's minds. Indeed, during the warm-up questions when children were asked "Look what Sammy has, do you know what this is?", 74% of children provided a response other a frog. Although a chi-square test didn't seem to be significant (χ^2; 4-YOs: p=.75, BF_{10}=.3; & 5-YOs: p= .56, BF_{10}=.66), this might still be important to address. Moreover, the way agents left the room could have been perceived to be a bit confusing because in one scene they are there, whereas in the next scene they sort of 'vanish'. Secondly, the experimenter (i.e. author of this thesis) was a non-native English speaker. Although there was no perceived difficulty in communicating with the children (as the experimenter is fluent in English), and children were able to understand the stories (apparent in their ceiling performance in all control questions), this could still be posing some challenges. There are mixed findings about children's accent understanding. Although there is research showing that when listening to accented speakers, children are able to understand less number of words (Nathan, Wells, & Donlan, 1998; Barker & Turner, 2015), there is also research showing that children's story comprehension was not affected when delivered by an accented speaker (Barker & Turner, 2015). Based on these mixed results, the possibility that accented speech could have affected the performance in the task cannot be ruled out.

In the following Pilot study, we aimed to address the first follow-up issue by fixing the story visuals.

2.4. Pilot

2.4.1. Method

Participants

A total of 15 children participated in this pilot study: 8 were 4-year-olds (M_{age} = 58.9 months, SD = 4.17; 3 females) and 7 were 5-year-olds (M_{age} = 61 months, SD = 4.23; 5 females). Parental consent for and verbal assent from each subject was obtained. Recruitment and compensation procedures were exactly the same as Experiment 1.

Materials and Procedure

The story in this pilot study was developed by the author. It is modeled after the task in Experiment 1 to include minor changes both to the stimuli and the narrated story. It is carried out as a story that is introduced to the subjects with a set of 20 pictures that were animated using Microsoft Power Point, accompanied with a preset narrative. Each picture corresponds to an event that's taking place within the story. Sample layout of the stimuli that accompanies the stories can be seen in Figure 7. Before the story started, subjects were first introduced to the agents (characters) in the story with the use of paper cut outs of each agent. Then, they were asked to tell who is who to ensure they understood and can remember the names of each agent (e.g. "Can you show me which one's Sammy?"). Afterwards, the story started. The story script is as follows:

> "This is Sammy, and this is Anne [Pointing to agents]. Oh look! Look what Sammy has in his hands! Yes, it's a <u>bunny</u>, that's right! Sammy was playing with his bunny, but now Sammy needs <u>to get some water</u>. Sammy puts his bunny into the box over here, then Sammy <u>goes to the kitchen</u>. <u>Bye Sammy!</u> Now, Anne wants to play with the bunny. Anne goes over to the box and takes the bunny out. But now, Anne needs to <u>do her homework.</u> So, look what Anne does! Anne puts the bunny into the basket! Did Sammy see that? No, that's right! And then, Anne goes away <u>to her room.</u> <u>Bye Anne!</u> But, look what happens while the two kids are gone! The bunny comes out of the basket, and then it jumps *away*, *far far away*, where no one knows where it went! Did Sammy see that? No, that's right! How about Anne, did Anne see that? No, that's right!" [Underlined parts indicate the differences between the narrative in Exp 1 and current pilot script].

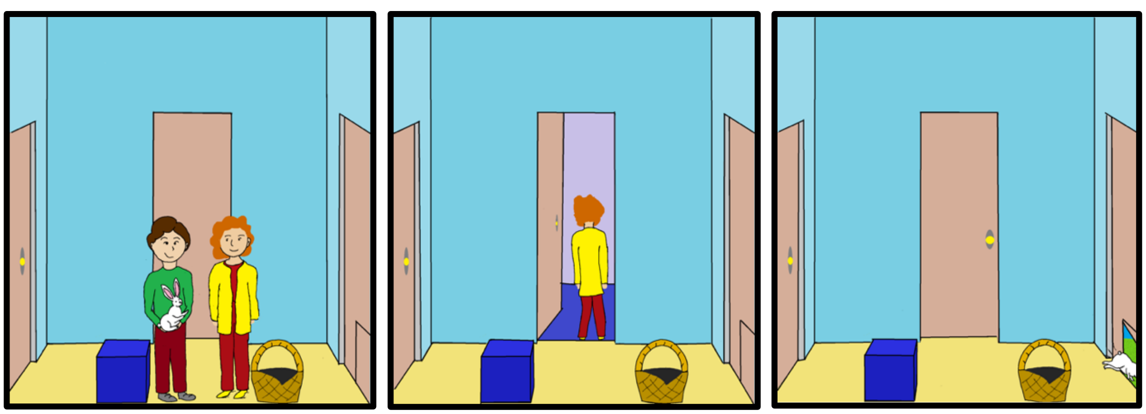

Figure 7. Sample study scenes. ***Left:*** Sample story visuals. Sammy (in green shirt, holding the bunny), and Anne. A box and basket on the floor. ***Middle:*** Anne leaving the room. ***Right:*** bunny jumping through the flap.

The same procedure in Experiment 1 was followed. If subjects said 'yes' to any of the See Questions (e.g. "Did Sammy see that?"), they were corrected by saying "Sammy wasn't in the room, right? He can't see that!". After the story ended, children were asked several control questions, the *Memory Questions* ["Where did Sammy (or Anne, both asked) put the Bunny?"] and the *Reality Question* ["Where is the bunny right now?"]. After the control questions, First Order False Belief (FB) questions were asked first: "When Sammy comes back, where will Sammy look for the bunny? / Where will Anne look for the bunny?". Afterwards, the Second Order Question was asked "Where does Anne think Sammy will look for the bunny?". The story was only repeated once if subjects failed any of the control questions. Where the object was hidden first (box or basket) was counterbalanced across all subjects, as well as the order of the Memory Control and the First Order FB questions. Each session was videotaped. Two coders watched videos for all subjects, and recorded answers to the questions.

2.4.2. Results & Brief Discussion

Figure 8 shows the individual responses obtained for the Pilot study with the new study visuals. A similar trend to Experiment 1 was observed, as only 3 out of 8 four-year-olds, and 4 out of 7 five-year-olds were successful in the first order FB questions. For this reason, we set out the address the second issue addressed in the previous section which pertains to eliminating the potential effects of accented speech on children's performance.

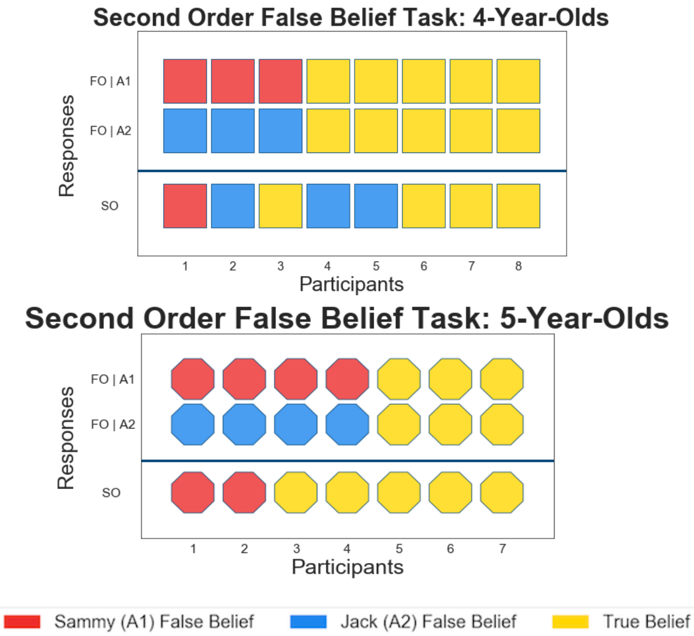

Figure 8. Individual responses in Pilot Study. ***Top:*** 4-year-olds individual response sequences. ***Bottom:*** 5-year-olds individual response sequences.

Chapter 3

Experiment 2

3.1. Method

Participants

A total of 56 children participated in this experiment: 43 were 4-year-olds (M_{age} = 53.73 months, SD = 3.32; 26 females) and 13 were 5-year-olds (M_{age} = 63 months, SD = 3.1; 6 females). One additional subject was tested but excluded from the study because they didn't want to listen to the test questions. Parental consent for and verbal assent from each subject was obtained. Recruitment and compensation procedures were exactly the same as Experiment 1 and the Pilot Study.

Materials and Procedure

The story visuals and the narrated story were exactly the same as the previous pilot study, with a difference in the medium they were introduced. Each sentence in the narrative was carefully recorded by a native English speaker (a lab member). In several occasions, in order to make the visuals match the story script more naturally, sentences were divided into sections that would correspond exactly to what was happening at the visuals for that scene.

The visuals in the pilot study and voice records obtained from the native speaker were combined by creating a program in MATLAB (MATLAB, 2018). The story was introduced to each subject by simply running the program. The script was written in a way that the experimenter had full control over proceeding to the next scene/story line, enabling her to monitor children's attentiveness to the story at any given point. *All* the questions were asked via the recordings done by the native speaker. If children did not respond at first, experimenter had the availability to repeat it via the program. The experimenter remained as silent as possible throughout, acting like she was listening to the narration together with the subject, sharing eye contact and smiles when necessary (e.g., when the subjects looked at her rather than looking at the screen). However, sometimes the experimenter would encourage subjects to pay attention, or to answer the questions in the story, in some other occasions experimenter interacted with the subjects simply when the subjects wanted to talk to, or interact with the experimenter (e.g., they wanted to tell a story about when they saw a bunny themselves). These instances were unavoidable as to ensure that the rapport between the experimenter and the subject is maintained, and that subjects feel safe and comfortable. The rest of the procedure was exactly the same as Experiment 1 and the Pilot study.

3.2. Results

Figure 9 shows 4-year-olds' responses. Each square (or hexagons for the 5-year-olds) indicates a single response to one of the three belief questions. Each column represents the response sequence of an individual subject. The top two rows are responses to the first order FB questions, whereas the bottom row (third row, below the straight line) indicates responses to the second order FB question. For the first order FB questions 14 out of 43 four-year-olds correctly attributed both agents their respective ("fully correct") false belief (binomial 1/9, p <.001; BF_{10}= 202.123) while 16 out of 43 four-year-olds incorrectly attributed both agents a true belief (binomial 1/9, p <.001; BF_{10}= 3824.28). For the second order FB questions, which can be observed in the bottom row in Figure 9, 12 out of 43 four-year-olds correctly responded to the second order FB question (binomial 1/3, p = .52; BF_{10} = .226) while 24 out of 43 four-year-olds incorrectly attributed a true belief (binomial 1/3, p = .003; BF_{10} = 17.8).

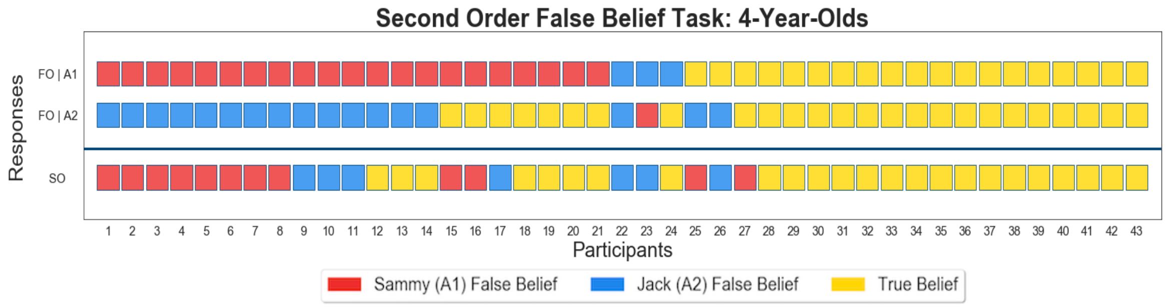

Figure 9. 4-year-olds' individual responses for Experiment 2. Each column represents the response sequence of an individual subject. FO rows marks the responses to the first order FB questions; FO | A1 row indicates the belief attributed to Agent 1; FO | A2 indicates the belief attributed to Agent 2. SO row indicates the responses to the Second Order question. Correct response sequence is FO | A1 = red; FO | A2 = blue; SO = red.

Figure 10 shows responses for 5-year-olds. For the first order FB questions, 10 out of 13 five-year-olds correctly attributed both agents their respective ("fully correct") false belief (binomial 1/9, p <.001; BF_{10}= 1.241e+6) and 2 out of 13 five-year-olds attributed a true belief to both agents (binomial 1/9, p = .649; BF_{10}= .271). For the second order FB questions, which can be observed in the bottom row in Figure 10, 10 out of 13 five-year-olds correctly responded to the second order FB question (binomial 1/3, p = .002; BF_{10} = 49.81) while 3 out of 13 five-year-olds incorrectly attributed a true belief (binomial 1/3, p = .564; BF_{10} = .389).

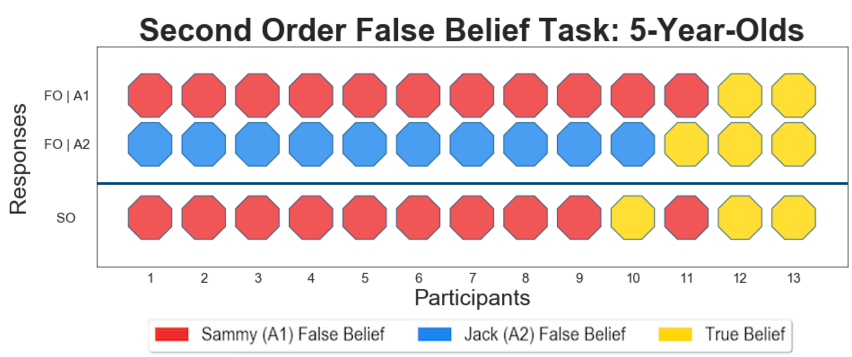

Figure 10. 5-year-olds' individual responses for Experiment 2. Each column represents the response sequence of an individual subject. FO rows marks the responses to the first order FB questions; FO | A1 row indicates the belief attributed to Agent 1; FO | A2 indicates the belief attributed to Agent 2. SO row indicates the responses to the Second Order question. Correct response sequence is FO | A1 = red; FO | A2 = blue; SO = red.

Figure 11. Proportion of binned responses for Experiment 2. Responses were categorized into one of the following four categories: All Correct False Belief Attribution (All Correct), False Belief Binding Error (FB Binding), Mixed False Belief- True Belief Binding Error (Mixed Binding), and All True Belief Error (All TB). The frequency of each response is displayed above the respective bar.

Responses were grouped into four main categories, following the same criteria as Experiment 1. Figure 11 Left Panel shows the proportion of each category both for 4- and 5-year-olds. Figure 11 Right Panel shows the proportion of different responses to the second order FB question. Although the number of subjects in the two age groups are different, general trends in proportions show a pattern that replicates Experiment 1: five-year-olds had a higher All Correct FB attribution rate and were providing the correct response for the second order question far more frequently.

Similar to Experiment 1, further analyses were carried out excluding children who were not successful in the first order FB questions. These analyses showed that 8 out of 14 four-year-olds were successful in second order FB (binomial 1/3, p = .085; BF_{10} = 1.66) (see Figure 12, top panel). With the same exclusion criteria, 9 out of 10 five-year-olds were successful in second order FB (binomial 1/3, p < .001; BF_{10} = 268.63) (see Figure 12, bottom panel).

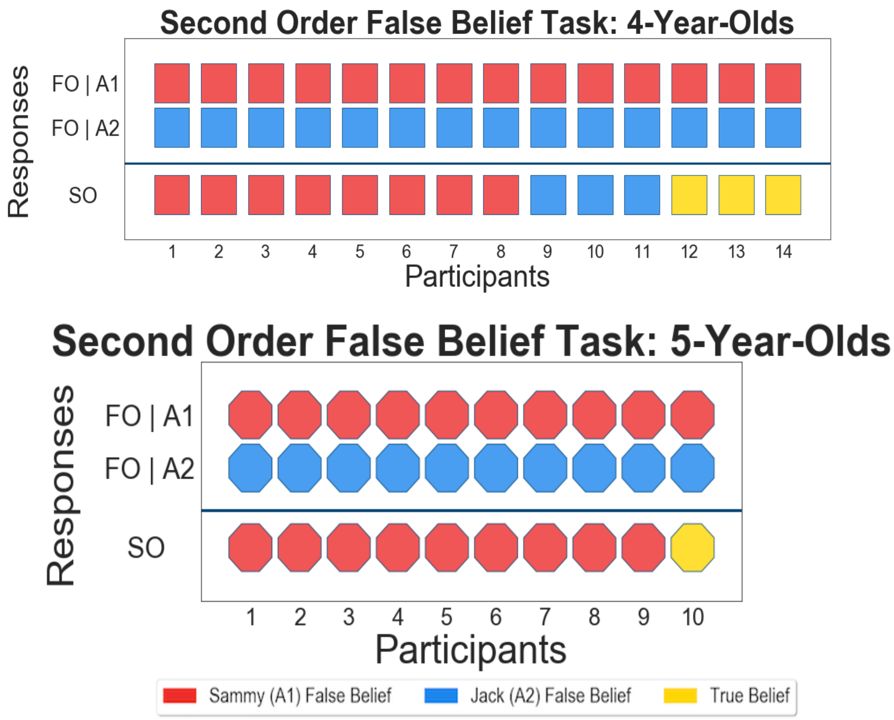

Figure 12. 4- and 5-year-olds' individual responses for Experiment 2; showing only the subjects who were successful in the first order FB questions. Each column represents the response sequence of an individual subject. FO rows marks the responses to the first order FB questions; FO | A1 row indicates the belief attributed to Agent 1; FO | A2 indicates the belief attributed to Agent 2. SO row indicates the responses to the Second Order question. Correct response sequence is FO | A1 = red; FO | A2 = blue; SO = red.

Chapter 4

General Discussion of Exp 1 & 2

The present studies set out to explore how working memory (WM) and inhibitory control (IC) demands would affect performance in preschoolers' second order false belief (FB) reasoning. Although there have been a few studies looking into how other cognitive processes or other factors might be interacting with second order FB understanding (Arslan et al., 2016, 2017; Hollebrandse et al., 2014; Miller, 2009, 2013a, 2013b), no study to date manipulated both WM and IC within the false belief task itself. Initial studies suggested that children around the age of 6.5-7 started to successfully reason about second order FBs (Perner & Wimmer, 1985), although subsequent research showed that with an 'easier' task – i.e. requiring less processing demands, 90% of the children were successful by the age of 5,5 (Sullivan et al., 1994). The current studies aimed to initiate a much larger project in which WM and IC demands would be systematically manipulated to have a deeper understanding of which 'processing demands' are playing a role in second order theory of mind.

The current studies employed a double agent low IC demand task. In this task there were two agents (characters) playing with a frog (Exp1) or a bunny (Exp2). Agent 1 put the bunny in the box and left the scene. Then, Agent 2 moved the bunny to the basket, and left the scene. While both agents were gone, bunny came out of the basket, and jumped away to an unknown location. The task required children to keep track of both Agent 1 and Agent 2's false beliefs about the whereabouts of the bunny. Each agent had a distinct false belief about where the bunny currently was, and only the child themselves knew about what really happened to the bunny – although their confidence was low as to where the bunny was. At the end of the story, besides being asked about the first order FBs of both agents, children are asked to reason about what another agent would think about another agent's mental states. Namely, children were asked: "Where does Anne think that Sammy will look for the bunny?". This question required children to recursively represent two agents' beliefs.

In line with Sullivan et al.'s (1994) findings showing success by the age of 5,5, the five-year-olds in both of our experiments were successful in the second order FB question. After obtaining somewhat lower success rates in 4-year-olds' first order FB performance (compared to that of Cheng, 2018) in Experiment 1, we set out to address any potential issues relating to our initial experiment. However, the results of the Experiment 2 were almost identical to Experiment 1 (with somewhat higher success in 5-year-old performance), suggesting that details that were potentially concerning about the study visuals and the experimenter's accented speech were not influential in children's FB performance. Furthermore, in both experiments, further analyses -where the children failing in the first order FB was excluded, showed that both for 4- and 5-year-olds, the children that were successful in the first order FB questions also tended to be successful in the second order FB question. These findings land an initial and a tentative evidence against the proposal that children are undergoing a conceptual shift where they acquire the concept of recursion or "understand the possibility of recursion" in representing mental states in a later point in development (Perner, 1988). However, further exploration of 4-year-olds' first order FB performance in multiple agent tasks is needed to make any further theoretical claims.

The current studies were limited in several ways. First of all, the difference in results between Cheng's (2018) findings in 4-year-olds' double agent, low IC task, and that of current studies should be further examined. A replication of Experiment 2's methodology with a new and also a bigger sample could help elucidate this issue. Secondly, within the scope of this thesis, only the low inhibitory control demand condition was considered. This leaves the question of how inhibitory control might interact with second order FB understanding to be still open. Could the high performance of five-year-olds in this study be due to the low demand nature of this task? How would the performance be affected when children need to keep track of three agents, instead of two? Thirdly, the second order belief studies to date have always included a 'belief revision' component. For instance, a belief revision would be needed where in the story when Anne is moving the bunny to the basket, Sammy peeks through the window and witnesses the new location of the bunny. Therefore, the child would need to update Sammy's belief about bunny's location. They would also need to update Anne's belief about Sammy's belief, because Anne would be unaware that Sammy was peeking through the window. So, "Anne doesn't know that Sammy knows the bunny is now in the basket.". Reasoning about this scenario means that children would need to account for Anne's false belief about Sammy's false belief. They both hold FBs in such scenario because in reality the bunny is gone; whereas in our current studies, children were reasoning about Anne's *true* belief about Sammy's false belief – since there was no revision/update of Sammy's belief state.

This thesis initiated the very first steps into understanding children's representational abilities for second order theory of mind. There are definitely many more interesting questions remained to be explored. Future projects will include testing of these possibilities.

References

Arslan, B., Hohenberger, A., & Verbrugge, R. (2017). Syntactic Recursion Facilitates and Working Memory Predicts Recursive Theory of Mind. PLoS ONE, 12(1), Article e0169510.

Arslan, B., Taatgen, N. A., & Verbrugge, R. (2016). Five-year-olds' systematic errors in second-order false belief tasks are due to first-order theory of mind strategy selection: A computational modeling study. *Frontiers in Psychology, 8,* Article 275.

Barker, B.A., & Turner, L.M. (2015). Influences of foreign accent on preschoolers' word recognition and story comprehension. *Applied Psycholinguistics, 36*, 1111-1132.

Baron-Cohen, S., Leslie, A.M. & Frith, U. 1985: Does the autistic child have a 'theory of mind'? *Cognition, 21*, 37–46.

Bartsch, K. (1996). Between desires and beliefs: Young children's action predictions. *Child Development, 67*, 1671-1685.

Birch, S.A.J., & Bloom, P. (2003). Children Are Cursed: An Asymmetric Bias in Mental-State Attribution. *Psychological Science*, *14*(3), 283–286.

Birch, S.A.J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*(5), 382-386.

Bloom, P. & German, T.P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*, B25-B31.

Bosco, F., Friedman, O., & Leslie., A.M. (2006). Recognition of pretend and real actions in play by one- and two-year-olds: Early success and why they fail. Cognitive Development, 21, 3-10.

Carlson, S.M., & Moses, L.J. (2001). Individual Differences in Inhibitory Control and Children's Theory of Mind. *Child Development, 72*, 1032–1053.

Carlson, S.M., Moses, L.J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contribution of inhibitory control and working memory. *Infant & Child Development, 11*, 73-92.

Carlson, S.M., Moses, L.J., & Hix, H.R. (1998). The role of inhibitory control in young children's difficulties with deception and false belief. *Child Development, 69*, 672-691.

Cheng, M. (2018). Small group theory of mind: Beyond the single agent false belief task (Unpublished doctoral dissertation). Rutgers University, New Brunswick, USA.

Davis, H.L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology, 47*, 25-31.

Dennett, D. (1989). The intentional stance. The MIT Press.

Friedman, O., & Leslie, A.M. (2004a). A developmental shift in processes underlying successful belief-desire reasoning. *Cognitive Science, 28*, 963-977.

Friedman, O., & Leslie, A.M. (2004b). Mechanisms of belief-desire reasoning: Inhibition and bias. *Psychological Science, 15*(8), 547-552.

Friedman, O., & Leslie, A.M. (2005). Processing demand in belief-desire reasoning: Inhibition or general difficulty? *Developmental Science, 8,* 218-225.

German, T.P., & Leslie, A.M. (2000). Attending to and learning about mental states. In P. Mitchell, & K. J. Riggs (Eds.), Reasoning and the mind (pp. 229–252). Psychology Press.

German, T.P., & Leslie, A.M. (2001). Children's inferences from 'knowing' to 'pretending' and 'believing'. *British Journal of Developmental Psychology, 19*, 59–83.

German, T.P., & Hehman, J.A. (2006). Representational and executive selection resources in 'theory of mind': Evidence from compromised belief-desire reasoning in old age. *Cognition, 101,* 129-152.

Gopnik, A., & Astington, J. W. (1988). Children's Understanding of Representational Change and Its Relation to the Understanding of False Belief and the Appearance-Reality Distinction. *Child Development, 59*, 26-37.

Gopnik, A. & Meltzoff, A.N. (1997). Words, Thoughts and Theories. The MIT Press.

Gopnik, A. & Wellman, H. (1994). The "theory theory". In Hirschfiled, L. & Gelman, S. (Eds.) *Domain specificity in culture and cognition*. Cambridge University Press.

Gordon, A., & Olson, D.R. (1998). The relation between acquisition of theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68*, 70-83.

Grosso, S., Schuwerk, T., Kaltefleiter, L., & Sodian, B. (2019). 33-month-old children succeed in a false belief task with reduced processing demands: A replication of Setoh et al. (2016). Infant Behavior and Development, 54, 151-155.

Hollebrandse, B., van Hout, A., & Hendriks, P. (2014). Children's first and second-order false-belief reasoning in a verbal and a low-verbal task. *Synthese, 191*(3), 321-333.

Keenan, T. (1998). Memory span as a predictor of false belief understanding. *New Zealand Journal of Psychology, 27*(2), 36-43.

Keenan, T., Olson, D.R., & Marini, Z. (1998). Working memory and children's developing understanding of mind. *Australian Journal of Psychology, 50*(2), 76-82.

Kikuno, H., Mitchell, P., & Ziegler, F. (2007). How do young children process beliefs about beliefs? Evidence from response latency. *Mind & Language, 22*(3), 297-316.

Leslie, A.M. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review, 94*(4), 412-426.

Leslie, A.M. (1994a). Pretending and believing: Issues in the theory of ToMM. Cognition, 50, 211–238. (Reprinted in J. Mehler & S. Franck (Eds.), *Cognition on Cognition*, 193–220, MIT Press, 1995).

Leslie, A.M. (1994b). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*, 119–148, Cambridge University Press.

Leslie, A.M. (1992). Pretense, autism, and the theory-of-mind module. *Current Directions in Psychological Science, 1*(1), 18–21.

Leslie, A.M. (2000a). How to acquire a 'representation theory of mind'. In D. Sperber (Ed.), *Metarepresentations: A Multidisciplinary Perspective,* 197-223, Oxford University Press.

Leslie, A.M. (2000b). "Theory of Mind" as a mechanism of selective attention. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed.), 1235-1247, The MIT Press.

Leslie, A.M., Friedman, O., & German, T.P. (2004). Core Mechanisms in "Theory of Mind". *TRENDS in Cognitive Sciences, 12*, 528-533.

Leslie, A.M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology, 50*(1), 45-85.

Leslie, A.M., & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology, 6*, 315-324.

Leslie, A.M., & Polizzi, P. (1998). Inhibitory processing in the false belief task: Two conjectures. *Developmental Science, 1*(2), 247-253.

Leslie, A.M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition, 43*(3), 225-51.

MATLAB. (2018). *9.7.0.1190202 (R2019b)*. Natick, Massachusetts: The MathWorks Inc.

Miller, S. A. (2009). Children's Understanding of Second-Order Mental States. *Psychological Bulletin, 135*(5), 749-773.

Miller, S. A. (2013a). Effects of Deception in Children's Understanding of Second-order False Belief. *Infant and Child Development, 22*, 422-429.

Miller, S. A. (2013b). Understanding of Second-Order False Belief: Comparisons of Content and Method of Assessment. *Infant and Child Development, 22*, 649-658.

Moses, L.J. (2001). Executive Accounts of Theory-of-Mind Development. *Child Development, 72*(3), 688-690.

Nathan, L., Wells, B., & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: a preliminary investigation. *Journal of child language*, *25*(2), 343–365.

Olson, D.R. (1989). Making up your mind. *Canadian Psychology, 30*, 617-627.

Onishi, K.H., Baillargeon, R., & Leslie, A.M. (2007). 15-month-old infants detect violations in pretend sequences. Acta Psychologica, 124, 106-128.

Perner, J. (1988). *Higher-order beliefs and intentions in children's understanding of social interaction.* In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (p. 271–294). Cambridge University Press.

Perner, J. (1991). Understanding the Representational Mind. Cambridge, MA: MIT Press.

Perner, J. (1995). The many faces of belief: reflections on Fodor's and the child's theory of mind. *Cognition, 57*, 241–69.

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science, 308*, 214-216.

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, 39*, 437-471.

Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 512-526.

Roth, D., & Leslie, A.M. (1998). Solving belief problems: Toward a task analysis. *Cognition, 66*,

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development and 'theory of mind.' *Mind & Language, 14*(1), 131–153.

Scholl, B. J., & Leslie, A. M. (2001). Minds, modules, and meta-analysis. Commentary on "Meta-analysis of theory-of-mind development: The truth about false belief." *Child Development, 72*(3), 696–701.

Setoh, P., Scott, R.M., & Baillargeon, R. (2016). Two-and-a-half-year-olds succeed at a traditional false belief task with reduced processing demands. *PNAS, 113*, 1336-13365.

Siegal, M., & Beattie, K. (1991). Where to look first for children's knowledge of false beliefs. *Cognition, 38*(1), 1–12.

Surian, L., & Leslie, A. M. (1999). Competence and performance in false belief understanding: A comparison of autistic and normal 3-year-old children. *British Journal of Developmental Psychology, 17*, 141-155.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers Can Attribute Second-Order Beliefs. *Developmental Psychology, 30*(3), 395-402.

Wang, L., & Leslie, A. (2016). Is Implicit Theory of Mind the 'Real Deal'? The Own-Belief/True-Belief Default in Adults and Young Preschoolers. *Mind & Language, 31*, 147–176.

Wellman, H. M. (1990). The child's theory of mind. Cambridge, MA: MIT Press.

Wellman, H. M. (2002). Understanding the psychological world: Developing a theory of mind. In U. Goswami (Ed.), Blackwell handbook of childhood cognitive development (pp. 167–187). Oxford, UK: Blackwell, 167–187.

Wellman, H.M., Cross, D. and Watson, J. 2001: Meta-analysis of theory of mind development: the truth about false-belief. *Child Development, 72*, 655–84.

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

Yazdi, A.A., German, T.P., Defeyter, M.A., & Siegal, M. (2006). Competence and performance in belief-desire reasoning across two cultures: The truth, the whole truth and nothing but the truth about false belief? *Cognition, 100*, 343-368.

Zaitchik, D. (1991). Is only seeing really believing? Sources of true belief in the false belief task. *Cognitive Development, 6*, 91-103.