COMPETENCE & PERFORMANCE IN BELIEF-DESIRE REASONING:

PROPERTIES OF SPONTANEOUS THEORY OF MIND

By

LU WANG

A Dissertation submitted to the

Graduate School – New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Psychology

written under the direction of

Alan M. Leslie

and approved by

____

____

____

____

New Brunswick, New Jersey

October 2014

ABSTRACT OF THE DISSERTATION

Competence & performance in belief-desire reasoning:

Properties of spontaneous theory of mind

By LU WANG

Dissertation director:

Alan M. Leslie

Navigating the social environment requires us to understand and predict people's actions. This ability, known as theory of mind, develops early in children. Typically, children's theory of mind ability is assessed using the standard false-belief task in which children are asked to predict a character's action given his/her outdated belief. A new wave of research measuring participants spontaneous reactions, such as looking time and anticipatory eye gaze, have revealed an early understanding of false belief, even in infants. This dissertation reports a series of studies that investigate the nature of early theory of mind ability and how this relates to the later developed ability in preschoolers and adults. The first two studies in chapter 2 suggest that children and adults are able to reason about false belief spontaneously, as measured by their anticipatory looking; but their anticipation is subject to the same processing demands observed in children's performance in verbal false belief tasks, namely, a tendency to attribute to others a true belief derived from one's own knowledge whenever that information is available. Two studies in chapter 3 report evidence that three- and four-year-old children can represent two

distinct false beliefs and bind each belief to the correct agent when prompted explicitly in a verbal task, and two-year-olds can do the same spontaneously in a non-verbal task measuring anticipatory eye gaze. Early working memory therefore has the capacity to bind at least two distinct mental states to each of two agents. Finally, a study in chapter 4 reports that children around 39 months can predict a person's behavior based on an understanding of the person's second-order false belief in a task measuring children's preferential looking time. Together, these studies suggest that the early understanding of mind cannot be reduced to perceptual or behavioral primitives, but instead reflect the developmental basis of genuine theory of mind.

Acknowledgements

I would like to thank Alan M. Leslie for his guidance, support, encouragement, and great sense of humor. It has been an incredible pleasure working with him. I am deeply influenced by his enthusiasm in work, and views of cognitive science, which I greatly appreciate. I would also like to thank Randy Gallistel, Rochel Gelman, Renee Baillargeon, and Pernille Hemmer for their valuable advices and guidance, and Sydney Levine, Melissa Kibbe, Katya Saunders, Michelle Cheng, and Deena Weisberg for all the discussions that have inspired me. Special thanks to Sydney Levine for her patience in reviewing the early drafts, and detailed comments that tremendously improved this dissertation.

I would like to thank all the members of the Cognitive Development Lab, especially our lab manager Alex D'Esterre, the two fabulous honor students Felicia Sitrin and Joshua Curtiss, and all the smart undergraduates I have enjoyed working with, without whom data collection would have been impossible. Special thanks to all the parents and babies contributed to my research.

Last but not least, I want to thank my parents whose love and care is always around me, and Qibin Zhou for all the support, love, and laughter.

This research was supported by NSF grant BCS-0922184.

Table of Contents

List of Tables

List of Illustrations

I. Introduction

Making sense of people's actions is critical for us to navigate the social world. Adults do so by attributing a pair of mental states, desire and belief, to others. This ability, known as "theory-of-mind" (ToM), entails a conceptual understanding of mind and mental states. That is, what a person has in mind is a representation of the world (denoted as P), and the person's mental states (e.g. believes that P) are relationships between the person's propositional attitude (believe) and what he/she has in mind (P).

Psychologists have long been interested in the origin and development of ToM. Children's ToM ability has been extensively studied using standard false-belief tasks, such as the "Sally-and-Anne" task. In the story, children are told that a character, Sally, puts her toy in location A and leaves the room. Another character, Anne, moves the toy to location B. Children are asked to predict where Sally will search for her toy when she returns (Baron-Cohen, Leslie, & Frith, 1985; Wimmer & Perner, 1983). Starting at age four, children can correctly predict that Sally will search in location A because of her false belief. However, three-year-olds typically predict that Sally will search in location B, where the toy really is. This developmental pattern – that four year olds pass the task while three year olds fail – has been shown to hold across various experimental methods (e.g., change-of-location tasks as in Baron-Cohen et al., 1985; unexpected content tasks as in Gopnick & Astington, 1988) and replicated among cultures across the world (e.g. Callaghan, et al., 2005; Liu, Wellman, Tardif & Sabbagh, 2008).

Given these findings, a near consensus has emerged among most researchers in the field that children do not possess an adult-like concept of mind (especially belief) until age four. In other words, younger children do not understand that belief is a representation of the world, and that it can be true or false. On this view, the failure-to-success shift in false belief tasks reveals a conceptual revolution from a non-representational to a representational view of mind around the fourth birthday (for a review, see Wellman, Cross, & Watson, 2001). Some, however, continue to argue that such consensus was premature, and that developmental change does not preclude an early competence in ToM reasoning (Scholl & Leslie, 1999).

Proponents of the conceptual revolution view disagree on what factors bring about the shift, but the most commonly hypothesized factors are children's language ability (see a meta-analysis, Milligan, Astington, & Dack, 2007), the development of executive function (Carlson & Moses, 2001; Carlson, Moses, & Hix, 1998), and evidence-driven theory change (the theory-theory, Gopnik & Wellman, 1994; Perner, 1991). Championed by Leslie and his colleagues, proponents of the early-competence view argue that ToM is driven by an innate domain-specific module, the ToM mechanism (ToMM), that first becomes expressed in the second year of life (Leslie, 1987, 1994a). On this view, the child possesses a competent notion of belief (in which beliefs can have false as well as true contents) but encounters difficulties with applying this notion because of the processing demands of the experimental tasks.

Specifically, ToMM mandatorily interprets people's behavior in mentalistic terms, such as belief and desire. As a phylogenetically forged module (Fodor, 1983; Leslie, 1987), it calculates people's mental states from sparse observation of their behavior, and provides a representational view of mind as its output. In a false belief scenario, ToMM computes all plausible contents for belief attribution, including a 'true belief' (derived from one's own belief) and a 'false belief' (based on the other person's perceptual exposure). The ToMM comes with a default prior to attribute the 'true belief' to others. After all, people's knowledge of most every day matters is true. In other words, without information to the contrary, your belief that P is most likely the same as my belief that P. Therefore, an attribution of my own belief to others is a reasonable and efficient prior. However, in the false belief scenario, the default attribution of 'true belief' is guaranteed to be wrong, and it produces a bias that needs to be inhibited for the false-belief content to be selected.

Leslie and his colleagues have proposed a selection processor, a cognitive process that applies inhibition over the true-belief default, to account for the failure-to-success shift in false belief tasks (Leslie & Thaiss, 1992; Leslie, German, & Polizzi, 2005). According to the ToMM-SP model, three-year-old children's difficulties in false belief tasks are caused by poorly developed selection processor, and therefore, it is hard for them to inhibit the 'true belief' prior. However, if the content of a 'true-belief' is more easily inhibited by the selection processor in a given scenario, younger children should be more likely to succeed. That is, we should be able to observe systematic changes in performance in relation to the manipulations of the demands on selection processor.

Consistent with this hypothesis, when researchers developed tasks with lower demands on selection processor by balancing the salience of the true and false representations, three-year-old children were more likely to succeed in understanding false belief (Bartsch, 1996; Wellman & Bartsch, 1988; Zaitchik, 1991). For example, Zaitchik (1991) found that when told about a toy's location instead of seeing it, and therefore, the content of a 'true belief' was less salient and more easily inhibited, three-year-old children's performances in the false-belief task were dramatically improved.

Similar improvements were also achieved by highlighting the content of false belief and therefore decreasing the demands on selection processor (Mitchell & Lacohée, 1991; Bartsch & Wellman, 1989). These variations all suggested an early competence of ToM that is masked by younger children's immature selection processor.

Furthermore, even though it is uncontroversial that four-year-olds have the concept of belief, they may still fail in performance when the demands of selection processor increase (Cassidy, 1998; Friedman & Leslie, 2004a, 2004b; Leslie, German, & Polizzi, 2005; Leslie & Polizzi, 1998). For example, Leslie & Polizzi (1998) introduced an avoidance desire to the false belief task. They told four-and five-year-olds that a character wanted to avoid a certain target (e.g. a frog), yet they had a false belief about where the target was. Predicting an action out of an avoidance desire requires an inhibition of the target, independent from the inhibition of the true-belief prior. For example, for subjects to predict which box the character will choose if he wants any box but the one with a frog in it, subjects have to locate where the frog is and inhibit that location. Leslie & Polizzi (1998) found that if a character wanted to avoid a frog and he had a false belief about where the frog was, four-year-olds systematically erred in predicting the character's action, while they had no problem understanding the character's false belief. That is, four-year-old children's poor performances were not caused by the failure to understand false belief, but a failure of selection processor to marshal double inhibitions imposed by a false belief and an avoidance desire.

The role of inhibitory control in successful belief-desire reasoning is not exclusively a developmental phenomenon. Research with individuals undergoing cognitive aging and with lesion patients has shown that adults, whose belief concepts no one contests, still need sufficient inhibitory abilities to succeed in false belief reasoning. For example, inspired by developmental research, German & Hehman (2006) varied ToM tasks' demands on inhibition and found similar decline in adults' performances when the tasks' processing demands increased; such effects were greater in elder individuals who suffered from decline in inhibitory ability because of cognitive aging. Moreover, damage to prefrontal cortex, a well-known brain area involved in inhibitory control, results in compromised false-belief reasoning (Apperly, Samson, Chiavarino, & Humphreys, 2004). Furthermore, the performance of these patients could be facilitated by decreasing the demands on inhibitory control, similar to the systematic change in children's performances (Samson, Apperly, Kathirgamanathan, & Humphreys, 2005). These findings are consistent with the ToMM-SP model, and further suggest that inhibitory function plays an on-going role in belief-desire reasoning beyond ToM development in preschoolers.

Given these collected findings, the argument for the early competence view persists, despite fierce opposition. As stated above, several alternative theories have been proposed to explain the origin and development of ToM. They all agree that the conceptual understanding of mind, marked by an understanding of false belief, is learned over the first couple years of life; yet they disagree on what factors underlie the conceptual revolution. Three main arguments have been presented to explain the conceptual revolution: (1) ToM is built on top of natural language; (2) ToM is a byproduct of the development of executive function; and (3) ToM, especially the concept of belief, is constructed from experience. Below, I will briefly review the alternative theories as well as empirical evidence in support of these views. Then I will argue that the proposed causal roles of language, executive function, and accumulation of counter-evidence to non-representational hypotheses of mind are not adequate to explain the observed phenomena, with special attention to the recent findings that preverbal infants can understand false belief. Having established this theoretical framework, I will overview the projects in this dissertation.

Language & ToM

Several scholars have argued that language promotes the development of ToM (Astington & Jenkins, 1999; de Villiers & de Villiers, 2000; Harris, 2005; Slaughter & Peterson, 2012). Three different roles of language have been suggested:

(1). The pragmatic role – engagement in conversational exchanges helps children realize that other people may have different perspectives, and hence, facilitates their understanding of others' minds (Harris, 2005; Harris, de Rosnay, & Pons, 2005; Slaughter & Peterson, 2012)

(2). The semantic role – the learning of mental-state vocabularies makes it possible for children to map the mental verbs to their feelings, and therefore, helps them render abstract mental states (e.g. Olson, 1988)

(3). The syntactic role – the acquisition of complex syntactical knowledge, especially the knowledge of sentential complement and embedded propositions, serves as a conceptual framework for children to represent minds (Astington & Jenkins, 1999; de Villiers & de Villiers, 2000; de Villiers, 2005, 2007). Specifically, knowledge of sentential complements may give us the cognitive power to embed thoughts into thoughts

("John thinks that Mary wants him to pretend …"), and that is how children acquire ToM (e.g., de Villiers & de Villiers, 2000, 2003; de Villiers & Pyers, 2002).

These roles of language have received support from various studies, such as training study, longitudinal study, and research with deaf children and adults. For example Lohmann & Tomasello (2003) trained a group of three-year-olds who failed standard false belief tasks with four types of training – a non-language training (a control without mentioning different people's perspective nor any mental states), a discourse-only training (emphasizing different people's visual perspectives but no descriptions of mental states), a sentential-complement training (using mental-states verbs and sentential complements but no perspective-shifting), and a full training in which both perspective shifts and sentential complements were provided. Children's performances in subsequent false belief tasks were significantly improved in all three conditions compared to the non-language training, suggesting that discourse about various perspectives (the pragmatic role) and sentential complement (the syntactical role) both contributed to the improvement of children's false belief reasoning. In a longitudinal study, Ruffman and his colleagues (Ruffman, Slade & Crowe, 2002) found that mothers' use of mental-states terms (emotion, desire, know, think, etc.) uniquely predicted children's ToM performances, even when other influential factors were controlled for. Furthermore, when language development was delayed, as in deaf children and adults, their understanding of false belief was also delayed (Peterson & Siegal, 1999, 2000; Pyers & Senghas, 2009).

However, serious challenges have been leveled against each of the proposed roles of language in ToM development:

(1). ToM is not just an output of the pragmatic use of language, but also a prerequisite to it. For example, we need to infer others' communicative intentions and past knowledge to make sense of a conversation or to learn new words, and children as young as two can do that (O'Neill, 1996; for another example, Happé & Loth, 2002; Sabbagh & Baldwin, 2001). On this view, ToM is a precondition for language acquisition, rather than the other way around (see Sperber, 2000, for a extensive discussion on language and meta-representation).

(2). According to the semantic role of language, Olson (1988) proposed that learning about mental-state verbs opens a window for children to attend to and understand these concepts. However, this hypothesis presupposes the existence of these concepts. After all, what is there to attend to if one does not already know these concepts?

(3). With respect to the syntactic role of language, studies with aphasia patients reveal that the patients' understanding of false belief were relatively intact while their grammatical abilities, including syntactical knowledge, were severely impaired (Apperly, Samson, Carroll, Hussain, Humphreys, 2006; Varley & Siegal, 2000; Varley, Siegal, & Want, 2001). Similarly with typical adults, studies using neuroimaging technologies (fMRI or PET) have showed that belief reasoning and language processing recruit distinct neural structures (e.g. Dodell-Feder, Koster-Hale, Bedny, & Saxe, 2011; for a review, see Saxe, Carey, & Kanwisher, 2004). However, these studies were about adults' ToM reasoning; it is still possible that language acts exclusively as a developmental precursor of ToM. A study conducted by Hale & Tager-Flushberg (2003) suggests that this is not the case. The authors conducted a training study with a group of three- to four-year-olds who failed both false belief and sentential complement tasks. Results showed that both direct training of false belief and training on sentential complement improved children's performance in subsequent ToM tests. Interestingly, however, children who received training on false belief did not improve in their knowledge of sentential complement, suggesting that ToM development does not rely on a development in sentential complements.

In sum, language is not a precursor for ToM development. Instead, as claimed by Bloom & German (2000), the observed relationship between language and ToM is likely caused by the standard tasks' processing demands produced by linguistic and non-linguistic factors. On the one hand, most false-belief tasks require rather advanced linguistic competence (Bloom & German, 2000; Chandler, Friz, & Hala, 1989). When the tasks' linguistic processing demands decreased, several studies have reported three-year-olds' success (Lewis & Osborne, 1990; Siegal & Beattie, 1991; Surian & Leslie, 1999; Yazdi, German, Defeyter, & Siegal, 2006). On the other hand, standard tasks elicit direct responses, and this measurement might be problematic for pragmatic reasons (e.g. Helming, Strickland, Jacob, 2014; Rubio-Fernández, 2013; Rubio-Fernández & Geurts, 2013). For example, children might have misinterpreted the experimenter question as asking about the target object's real location (Helming, Strickland, & Jacob, 2014) Therefore, younger children's limited linguistic skills and the deliberate verbal questions might have masked their early competence in ToM reasoning.

Executive function & ToM

Another commonly suggested factor that is causally related to ToM development is executive function. Executive function is not a single ability but a construct of cognitive processes underlying problem solving and goal-directed actions, including working memory, planning, mental flexibility, and inhibitory control (Hughes, 1998; Russell, Mauthner, Sharpe, & Tidswell, 1991; Shallice, 1982, 1988; Zelazo, Carter, Reznick, & Frye, 1997). Given its various components, children's executive function is typically tested by using a battery of different tasks (Carlson & Moses, 2001; for an analysis of various tasks, see Carlson, 2005). In a commonly used card-sorting task, for example, children are required to sort cards following one of two dimensions, such as color or shape. For a given item, a Red Circle, children have to sort it with Red Triangle under the Color dimension, and sort it with Green Circle under the Shape dimension. To succeed, children have to keep both rules in mind, and inhibit one to respond based on the other.

It is not surprising that executive function is related to ToM. To begin with, development of ToM and executive function take place around the same age (e.g. Frye, Zelazo, & Palfai, 1995). Moreover, as stated earlier, to succeed in false belief tasks, children have to inhibit the obvious representation of the truth to correctly respond with another person's false belief. This process requires the core elements of executive function, such as inhibitory control and working memory. Consistent with the apparent relations, Carlson and her colleagues found that after rigorous control of children's age, verbal abilities, general intelligence, and other factors, children's inhibitory control in situations with conflict options significantly predicted their performances in a battery of false belief tasks (Carlson & Moses, 2001; Carlson, Moses, & Breton, 2002; Carlson, Moses, & Claxton, 2004). A recent meta-analysis also suggested that the association between executive function and false-belief reasoning is robust and largely consistent across various measures of executive function and different types of false belief tasks (Devine & Hughes, 2014).

Does the relationship imply that executive function underlies the development of the ToM concepts (e.g. Carlson, Moses, & Hix, 1998)? Studies conducted on the ToM competence of children with Autism Spectrum Disorder (ASD) suggest that this is not the case. Children with ASD are specifically impaired in understanding false belief, but they have no problem in non-mentalistic tasks requiring inhibitory control. For example, Leslie and Thaiss (1992) tested the performance of children with ASD in a standard false belief and a false photograph (Zaitchik, 1990) task. In the false photograph task, children were shown a cat at location A when a photo was taken, and the cat was moved to location B afterwards. Children were asked, "in the photograph, where is the cat sitting". This task was closely matched to the false belief tasks in structure and processing. In particular, inhibitory control was required to suppress the truth in order to succeed. However, there were no mental states involved in the false photograph task. Typically developing four-year-olds succeeded in both false-belief and false-photograph tasks; children with ASD with a similar verbal mental age only succeeded in the false photograph task, while they showed a specific problem with belief reasoning. These findings suggest that a mature executive function does not itself produce an understanding of belief (but see Sabbagh, Moses & Shiverick, 2006).

Therefore, executive function might be necessary for success in ToM tasks, but it is not sufficient. Otherwise, one cannot explain children with ASD's specific impairment in ToM with a relatively intact executive function.

Evidence driven: the theory-theory

By drawing a parallel between cognitive development and theory shift in science, Gopnik & Wellman (1992, 1994) proposed that children's ToM development should be explained as shift of theories in light of evidence. Particularly, infants are born with some understanding of the mind, including desire and perception. However, the early understanding of desire and perception is radically different from adults' understanding of mind, in the sense that, for the infant, neither the concept of desire nor the concept of perception is representational. In other words, an agent's desire directly interacts with the world, while an agent's perception is directly determined by the world, without being mediated by any representations of the world. This forms the beginning state of ToM development.

Around three years old, children form a similar non-representational theory of belief, and they treat belief as a direct copy of the world, rather than a representation divorced from the real states ("copy theory", Wellman, 1990). Therefore, the truth status of a belief is not considered when it is first constructed by the three-year-olds. Nevertheless, during the transitional state, three-year-olds begin to realize the existence of belief representations that may not be consistent with reality. Yet this alternative view of belief is an auxiliary hypothesis that is only deployed when children are forced to explain counter-evidence to the earlier non-representational theory. For example, when three-year-olds were forced to explain a committed action out of a person's false belief, they were able to do so, even though they failed to use the representational theory to predict a person's action (Bartsch & Wellman, 1989).

Eventually, around the fourth birthday, the accumulation of counter-evidence forces children to abandon the early theory completely, and endorse the representational view of mind as the only viable theory. Consequently, four- and five-year olds understand that an agent's action is entirely mediated by the agent's representation of the world, and various mental states are manifestations of different propositional attitudes (such as desire, perception, belief and pretense) towards the representation.

According to Gopnik & Wellman (1994), the theory-theory is particularly attractive because of its dynamic interaction with evidence, and its revisable character. However, there are some challenges, both theoretically and empirically, leveled against the theory-theory. First of all, how is the representational view of belief constructed? Specifically, how could a child shift from non-representational theory to representational theory without a capacity to entertain the representational nature of mind at the first place? Gopnik & Wellman (1994) suggested that children's understanding of representational mental states first appears in their understanding of pretense and imaginations. Between two and four, they also develop a representational view of desire (desire of a food item might change over time) and perception (e.g. level two perspective taking that children understand another person may see the same object differently from another angle). These concepts present in other mental states helps to construct a representational hypothesis for belief. Nevertheless, the question remains: how do children construct a conceptual understanding of pretense and imagination? Without acknowledging an innate mechanism, the theory-theory still presupposes the origin of the cognitive ability to view the mind as representations decoupled from the world.

Secondly, theory-theory regards children's performances as evidence of their theoretical competence. This makes it particularly difficult for theory-theory to explain four-year-olds' failures in false belief tasks coupled with avoidance desire (Cassidy, 1998; Friedman & Leslie, 2004a, 2004b; Leslie, German, & Polizzi, 2005; Leslie & Polizzi, 1998). After all, why should four-year-olds fail to predict a person's behavior when they have a competent understanding of the mind and action? If proponents of theory-theory agree that performance limitations may have masked children's competence, it will open the possibility that three-year-olds' failures are performance issues as well, in contradiction to the theory-theory.

In sum, (1) children's ToM competence is independent from their language development, and the correlations between language and children's performance may simply reveal the linguistic complexities of ToM tasks; (2) manipulation of tasks' inhibitory demands could lead to a systematic change of children's performance in false belief tasks, as predicted by the ToMM-SP model, suggesting that executive function is a performance factor in ToM development. (3) the theory-theory fails to provide any explanations of the origin of ToM concepts, and it is not sufficient to account for children's performances in tasks with various processing demands. Nevertheless, the most convincing evidence for an early competence in ToM comes from the new findings of infants' false belief reasoning, to which I now turn.

Infants' understanding of false belief

Recent studies have shown that infants in their second year of life can understand false belief (e.g. Buttelmann, Carpenter, & Tomasello, 2009; Kovács, Téglás, & Endress, 2010; Onishi & Baillargeon, 2005; Scott & Baillargeon, 2009; Scott, Baillargeon, Song, & Leslie, 2010; Scott, He, Baillargeon, & Cummins, 2012; Song & Baillargeon, 2008; Southgate, Chevallier, & Csibra, 2010; Southgate, Senju, & Csibra, 2007; Surian, Caldi, & Sperber, 2007; Träuble, Marinovic, & Pauen, 2010; for a review, see Baillargeon, Scott, & He, 2010). For example, Onishi and Baillargeon (2005) adapted the Sally-and-Anne story into a nonverbal looking-time task for 15-month-olds. In the story, an actor falsely believed that a toy was in location A while it was actually in location B. Infants saw that the actor either searched for the toy in location A, congruent with her false belief, or in location B. They looked reliably longer when the actor reached for the toy's real location, indicating a violation of the infants' expectation that the actor should behave in accordance with the actor's belief about the toy's location and not in accordance with the infant's belief.

Subsequent studies measuring infants' looking time confirmed and extended this finding, revealing that infants in their second year of life can understand a false belief about a toy's location (e.g. Träuble, Marinovic, & Pauen, 2010), an object's identity (Scott & Baillargeon, 2009), an object's appearance (Song & Baillargeon, 2008), and even an object's hidden property (Scott, et al., 2010). They can also attribute false beliefs to a non-human agent (Surian, Caldi, & Sperber, 2007; Surian & Geraci, 2012).

Studies measuring anticipatory looking provided evidence for toddlers' and preschoolers' ToM competence too (e.g. Clements & Perner, 1994; Southgate, Senju, &

Csibra, 2007). Clements and Perner (1994) were the first to use the anticipatory looking measurement with preschoolers. The authors modified the standard false belief task in the following way: Subjects were shown two containers, which acted as hiding places for a target object. A door behind each container allowed the main character (a mouse) to come through and reach that container. An anticipation prompt “I wonder where he is going to look” was used before the character reappeared from one of the two doors, and children’s anticipation by eye gaze was measured during the one to two seconds after the prompt. Children as young as 2 years and 11 months shifted their eye gaze to the door where the character would appear given his false belief, even though they failed when asked explicit questions about the behavior of the mouse immediately after the anticipation prompt.

Building on Clements & Perner’s (1994) work, Southgate et al. (2007) showed non-verbal videos to 25-month-olds and recorded their eye gaze via an automatic eye tracker. In the videos, a puppet hides the toy in one of two boxes when an actor is watching. Directly above each box is a window that the actor could reach through to open the box. After a chime signal, the actor reaches through the window above the box with the hidden toy and retrieves the toy. In the test trial, the actor is distracted and faces away after seeing the puppet hide the toy in one box; then, the puppet moves the toy into the other box, and eventually leaves with it. Afterwards, the actor turns back and the signal is played, indicating that she is about to search for the toy. Most 25-month-olds correctly shifted their first gaze to the window where the actor would reach given her false belief.

Another line of research uses an ecologically dynamic context to test infants' understanding of false belief, such as helping and word learning (Buttelmann, Carpenter, & Tomasello, 2009; Southgate, Chevallier, & Csibra, 2010). For example, in the study by Southgate et al. (2010), an experimenter puts two novel objects, one into each of two boxes, and leaves the scene. A second experimenter switches the locations of the two objects. When the first experimenter comes back, she points to one box and teaches the infants that the object inside is a "sefo". Infants were encouraged to retrieve the "sefo" for her. When experimenter 1 did not witness the switch-of-locations and therefore had a false belief about what she was pointing to, most 17-month-olds interpreted the non-pointed-at object as "sefo". Thus, the study clearly showed that 17-month-olds were able to reason about the speaker's wrong epistemic state to interpret her intended referent (Southgate, Chevallier, & Csibra, 2010).

Overview of dissertation research

Overall, studies measuring infants' and children's spontaneous responses have revealed an early understanding of mind (Baillargeon, Scott, & He, 2010). Yet these findings seem to be in direct conflict with children's failures in standard verbal tasks of ToM reasoning. Studies in this dissertation were aimed to address this puzzle. Specifically, the discrepancy leads to two broad questions: (1) What is the nature of the earliest competence? (2) How does this early competence relate to the later developed abilities?

One straightforward hypothesis is that the tasks that measure spontaneous responses reduce not only the demands of linguistic processing but also the demands of infants' immature executive processes, thereby allowing the early ToM competence to shine through. We tested this hypothesis in the two studies presented in chapter 2. Our results suggest that the discrepancy between early success and later failures cannot be entirely attributed to mode of response; the early developing ability is still subject to some processing demands. In chapter 3, we further explored the capacity and limit of the early developing system, by asking whether it permits binding a belief representation to a specific agent. Results suggested that the early developing ToM ability is flexible and allows children to keep track of at least two agents' mental states simultaneously. Finally, building on the results in chapter 3, in chapter 4 we took up the challenge of decisively determining the nature of the early competence, whether it is mentalistic (conceptual) or behavioral (non-conceptual), by testing children's spontaneous understanding of a second-order false belief.

Together, these projects will provide a finer picture of the nature and the limits of the early-developing system in ToM. Implications of these projects for the reconciliation of the early competence and later developed abilities will be discussed in chapter 5.

II. The true-belief default in intuitive theory of mind

As introduced in chapter 1, ToM development in typically developing children and its impairment in children with autistic spectrum disorders has been studied extensively using the “Sally-and-Anne false belief” task (Baron-Cohen, Leslie, & Frith, 1985). It is well-established that the majority of typically developing children over four years of age will pass this task by predicting a person’s behavior given her false belief.

Most researchers have reached a consensus that ToM does not develop until four years or later (Wellman, Cross, & Watson, 2001). Some, however, continued to argue that such consensus was premature (Scholl & Leslie, 2001) and that the failure of younger children was due to the processing demands made by standard false belief tasks (Roth & Leslie, 1998). For example, standard false belief task demands considerable executive resources if one is to inhibit the tendency to attribute one’s own belief to others. In the absence of contrary information, it is rational to attribute to others the belief one holds oneself because typically people’s everyday beliefs are true (Leslie & Thaiss, 1992). However, in false belief tasks this default attribution is guaranteed to fail and must first be inhibited if the correct belief content is to be selected.

The early competence theory has received support in recent years from studies that have adapted standard false belief tasks in such a way that the verbal content is reduced to a minimum or eliminated altogether and combined with non-verbal measures, such as violation-of-expectation looking time, anticipatory looking, and spontaneous helping (e.g. Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007; Buttelmann, Carpenter, & Tomasello, 2009; Scott, et al., 2010). For example, as described in chapter

1, in Southgate et al. (2007), infants watched a videoed scenario modeled on the Sally and Anne task in which an object is hidden in one container while being observed by an actor and then is moved to another container while the actor's back is turned. Subsequently the object is taken out of the new container and removed entirely from the scene, again, all the while the actor's back remains turned. Finally, the actor turns to face the scene once more and a chime signals that the actor is about to reach for the object. Eye tracking measures revealed that infants' eye gaze spontaneously anticipated that the actor would search the container in which she last saw the object being placed.

Early success on false-belief tasks brings later failure into sharp relief. However, the older findings do not simply disappear with the new; indeed, they highlight the need to understand the factors preventing successful performance at an explicit verbal level. An obvious and plausible possibility is that early-success tasks for one reason or another avoid making demands on immature executive processes thus allowing early competence to be revealed. It has been suggested that early-success tasks rely only on the representation of other people's belief, whereas later-failure tasks also demand response-selection and response-inhibition processes (Baillargeon, Scott, He, 2010; Scott, Baillargeon, Song, & Leslie, 2010). Specifically, when explicitly asked about other people's false belief, children have to select the correct representation for responses among all plausible candidates, one of which is one's own knowledge (true belief); and they have to inhibit the tendency to respond with the truth. Simultaneously engaging belief representation, response-selection and response-inhibition processes is too demanding for younger children, which results in their failures.

If spontaneous responding does not draw upon executive processes and various information processing demands, then we can explain both early success and later failure. However, up till now, two studies measuring subjects' anticipatory eye gaze have revealed that spontaneous belief reasoning is still subject to tasks' processing demands (Low & Watts, 2013; Schneider, Lam, Bayliss, & Dux, 2012). For example, in Schneider et al. (2012), adult subjects watched a non-verbal video modeled on the Sally-and-Anne story. Critically, subjects received an additional non-theory-of-mind task that may vary in cognitive demands. They found that adults' spontaneous anticipation of the actor's behavior was impaired when the dual-task increased in cognitive demands, implying that belief reasoning as measured by anticipatory eye gaze still draws on domain-general executive resources.

What produces the processing demands even in spontaneous belief reasoning? This question remains largely untested. It is important because we need to know whether these critical demands arise extraneously to ToM itself, as, for example, extrinsic working memory limitations, or whether, on the contrary, they reflect aspects of basic belief-desire competence. We designed two experiments to address the question.

Experiment 1

In experiment 1, we examined whether spontaneous responding is subject to the same executive demands as shown in verbal false belief tasks. Our manipulations build upon previous verbal studies suggesting that three-year-old children are more likely to calculate false beliefs when the target is entirely removed and the child does not know its whereabouts, compared to when the target is hidden in the scene (Bartsch, 1996;

Wellman & Barstch, 1988; Zaitchik, 1991). We produced videos that we hypothesized would differ in the executive demand, and used these as stimuli in a novel anticipatory looking task. Two- to three-year-old children and adults saw an actor put a dog puppet in one of two boxes; the actor then turned around to fetch some food for the dog. Different groups of subjects saw one of the following trials: when the actor's back was turned, (i) the dog crawled out of the original box and went into the other one (high-demand (HD) false-belief condition); (ii) after the same movement as in (i), the dog came out of the new box again, moved to the center of the stage, and entirely left the scene by jumping off the stage center front (low-demand (LD) false-belief condition); or (iii) as in (i), the dog moved from the original box to the other box, except this time the actor turned back to face the scene while the dog was climbing into the new box, and therefore, she saw the dog's final position (true-belief (TB) condition).

Previous studies measuring elicited responses have found that undermining the subject's confidence in their own knowledge of the target object's location makes it easier for them to attribute a false belief to the actor. For example, in Carpenter, Call, and Tomasello (2002), an actor points at a box in which she previously placed a novel object — but which, unbeknownst to the actor, had subsequently been removed — and says, "Let's get the dax!". Compared to another scenario in which the actor points at a box in which she previously placed a novel object, but, unbeknownst to her, had subsequently been replaced by another novel object, in the situation with an empty box, the child herself does not have a specific competing belief about the content of the box, making it easier for the child to attribute a different belief to the actor (also see Bartsch, 1996). That is, three-year-old children are more likely to succeed in tasks where the target is entirely removed from the scene (similar to (ii)); however, they fail in tasks in which the target remains in the scene (similar to (i)) until four years of age. We asked whether two-to three-year-olds' and adults' spontaneous anticipations are influenced by the same executive demands as present in children's verbal performances.

Methods

Subjects were 87 two- to three-year-olds (45 girls; mean age = 37.0 months, range = 24.0 – 47.8 months) and 113 undergraduate students from Rutgers University (54 females; mean age = 19.6 years, range = 18.0 – 28.3 years). An additional 19 children and 10 adults were excluded because no looking behavior during the anticipatory phase (6 children, 10 adults) was recorded or quality of looking was poor[1] (overall looking time towards both screens < 200 milliseconds; 13 children).

A Tobii (Stockholm, Sweden) T60 XL Eye Tracker was used to present the stimuli and record subjects' eye gaze. A computer with Tobii studio presentation software controlled the study and recorded gaze. Adult subjects were tested individually in the lab. Children either came with their parents to the lab or participated at their preschool. When tested at preschool, children were seated in a quiet separate room 60 to 80 cm from the eye tracker (individually calibrated to maximize tracking quality). When tested in the lab, children sat in a testing booth on their parent's lap, and were also individually calibrated at 60 to 80 cm from the screen. Adult subjects sat on a chair in the

[1] Specifically, low quality recording of looking may be caused by loss or degrading of signal to the Tobii eye tracker, typically due to rapid head movements or subjects looking away. Such recordings may not accurately reflect subjects' anticipation by first look. Degraded signal occurred more often in children than in adults; therefore, to be included children's overall looking time towards both split screens was required to be ≥ 200 ms; adults who made no anticipation to either screen within 600 ms were excluded.

same environment. Children's and adults' eye tracking was calibrated with appropriate 5-point Tobii calibration videos at the start of the experiment.

Each subject watched three videos of two familiarization trials followed by a single test trial. Each video shows an actor in a blue shirt behind a rectangular table, wearing a beige visor that covers her eyes. There are two wooden boxes, red and green, on opposite sides of the table. At the beginning of the first familiarization trial, the actor puts a toy bunny into the red box, closes the lids of both boxes, and says, "Let me get its carrot," and turns away. She turns back with a toy carrot, says, "Here it is," and puts the carrot down in the center of the table. The main picture then splits into two identical screens, each a copy of the main picture that has disappeared. Each screen is half the main picture's height and width presented against a black background. At 80cm, the visual angle of the main picture before splitting is 30.2°; the visual angle of each split-screen is 15.5°, and the visual angle for the gap between the two split screens is 1.5° (Fig. 1). The actor stands in the center of each screen, and says, "Let me get the bunny." After 600 milliseconds, in the screen on the left, she reaches to the left and touches the lid of the red box, and in the screen on the right, she reaches to the right and touches the lid of the green box. Her actions in both screens mirror each other and happen simultaneously. As soon as the actor touches the lid of a box, the action is frozen till the end of the trial, seven seconds after the initial split. The second familiarization trial was identical to the first, except that the actor placed the bunny in the green box.

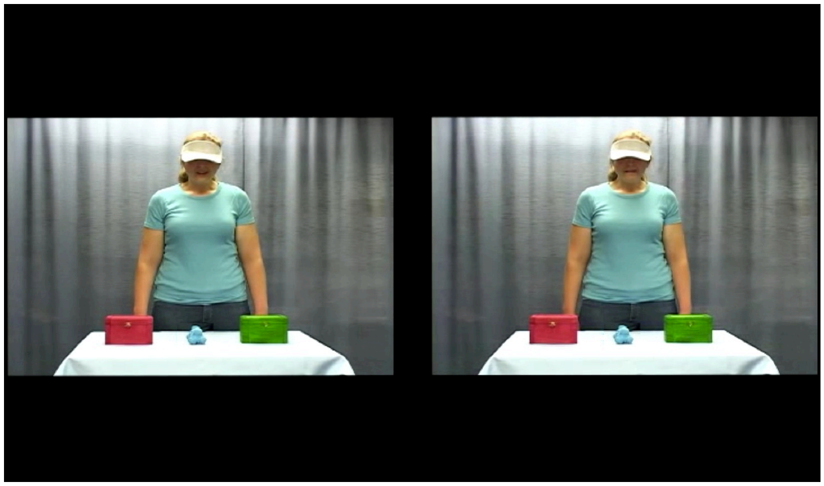


Figure 1 Split-screen effect in experiment 1. The main picture splits into two identical screens. During the first 600 ms, the actor stands in the center of each screen, and says, “Let me get the bunny” in the familiarization trials or “Let me get the dog” in the test trial.

In the test trial, the actor puts a dog puppet into the red box on the left, and says, “Let me get its bone,” and turns away. While her back is turned, the dog (i) crawls out of the red box and moves into the green box (HD false-belief condition); or (ii) after it moves into the green box, the dog comes out again, moves to the center of the table, and jumps off the table from the center front (LD false-belief condition). The actor then turns back holding a blue bone in her hand, and says, “Here it is.” She puts the bone down on the center of the table, and the split-screen phase identical to that in the familiarization trials starts. In both condition (i) and (ii), the actor has a mistaken belief that the dog is still in the red box. The difference between the conditions is whether the subject knows where the dog ends up (HD) or not (LD). The true-belief (TB) scenario starts the same way as in (i), except that the actor turns back to watch the dog trying to enter the green box. She again puts down the bone, saying, “Here it is.” Then she says, “Oh!” in a surprised tone, and turns her head to the dog who is still climbing into the green box, and she watches until the dog is completely hidden (video S1).

The initial hiding place of the toy was counterbalanced across subjects within each condition. We calculated the latency of subjects' first gaze shift to one or other screen following the first video frame showing the split-screens. Only gaze within the first 600 milliseconds following split, while the two screens were still identical before reaching outcomes began, was considered as anticipation. Immediately after the test trial, adults were asked, "Where will she look for her dog?" Their responses were recorded as Red box or Green box and further coded as 0 (inconsistent with the actor's belief) or 1 (consistent with the actor's belief).

Results & Discussion

The split-screen method shows both possible reaching outcomes in all conditions. For example, in the HD condition, one split-screen will show the actor reaching to the box consistent with her outdated belief (the false-belief (FB) screen). The other split-screen will show an outcome consistent with the subjects' own belief (the true-belief (TB) screen). Given that the two screens were identical during the first 600 milliseconds, and that the actor always reached to the left in the left-side screen and to the right in the right-side screen, any systematic preference for shifting first look to one side could only be explained by anticipation of what the actor will do in that screen. Therefore, the question of interest was whether subjects' anticipation by first look during the 600 milliseconds would be influenced by the processing demands of the test trial.

In the LD condition, both two- to three-year-old children and adults preferred to look first to the FB screen (21 out of 30 two- to three-year-olds, and 27 out of 36 adult subjects; two-tailed binomial tests, $p = .043$, and $p = .004$, respectively; see Fig. 2), replicating and extending Southgate et al (2007). The opposite preference was observed in the TB condition in both children and adults (21 of 30 two- to three-year-olds, and 28 of 42 adult subjects; two-tailed binomial tests, $p = .043$, and $p = .044$, respectively). However, subjects showed no preference for first look to the FB screen in the HD condition (12 out of 27 two- to three-year-olds, and 18 out of 35 adult subjects; Binomial tests, $p > .5$, n.s.; Bayes Factor analysis supports the null with odds of 3.63:1 for children and 4.75:1 for adults). Given that children and adults showed the same pattern of anticipation, we combined the results across age groups. Combined anticipations suggested that first looks between the LD and HD conditions, and between the HD and TB conditions were significantly different in the expected direction (LD vs. HD: Upton's $\chi^2 = 7.89$, p=.002, one-tailed; HD vs. TB: Upton's $\chi^2 = 3.74$, $p = .027$, one-tailed).

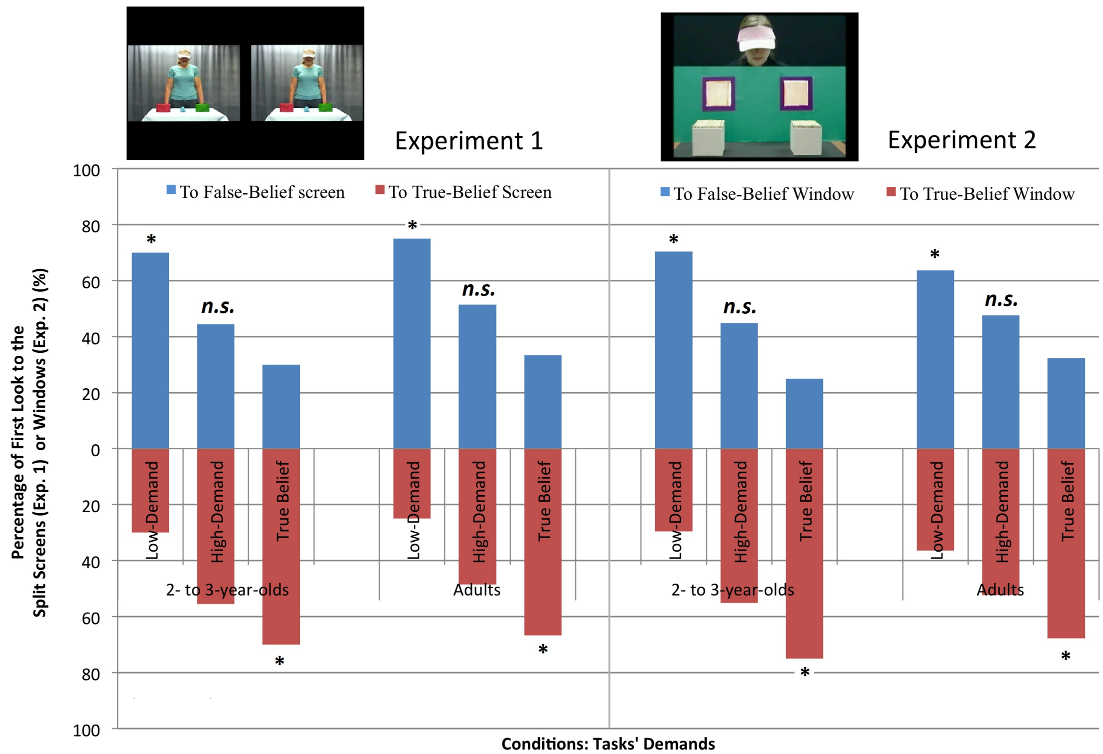


Figure 2 Results of anticipatory First Look, Experiments 1 (left panel) and 2 (right panel). Percentage of two-to three-year-old children's and adults' first look towards the False-belief screen/window or the True-belief screen/window.

* = Binomial test, $p < .05$, left panel two-tailed, right panel one tailed.

Immediately after the videos, adult subjects were asked, "Where will she look for the dog?" They correctly predicted the actor's behavior given her belief in all three conditions (30 of 36 in LD, 31 of 35 in HD, and 37 of 42 in TB; two-tailed binomial tests, $p < .001$). Verbal responses were more accurate than spontaneous anticipatory gaze in HD and TB conditions but not in LD (McNemar Binomial tests: in HD, $p = .001$; in TB, $p = .022$; in LD, $p > .5$).

Two- to three-year-olds and adults correctly attributed a belief to an actor in our true-belief scenarios, but also in our false-belief scenarios, at least when processing demands were low. However, knowing the hidden target's whereabouts reduced subjects' ability to anticipate the character's behavior in false-belief scenarios when processing demands were high. Strikingly, these demands were still evident in the reduced ability of our adult subjects. Together, these findings suggest that spontaneous anticipation does not escape the executive demands of false belief tasks when measured by anticipatory eye gaze. However, these findings come from a novel split-screen task; we therefore wished to check these results in our next study with a task that has been used more widely (Clements & Perner, 1994; Garnham & Ruffman, 2001; Low & Watts, 2013; Southgate, Senju, & Csibra, 2007; Schneider, Bayliss, Becker, & Dux, 2012).

At the same time, we need to address the question of what drives the processing demands in the high-demand false-belief task. There are two views in existing literature. One possibility is that subjects have a 'reality bias' (Mitchell, 1994; Russell, Mauthner, Sharpe, & Tidswell, 1991). Sometimes, when subjects should be calculating belief, they instead report reality. On this account, failures in false-belief reasoning reflect a failure to engage in ToM. Another possibility that has been argued for is that belief calculation is engaged spontaneously, but is subject to a 'true-belief bias' (Leslie & Thaiss, 1992). When subjects calculate belief, they have a tendency to attribute their own belief by default. Typically, people's beliefs about the everyday world are true. Therefore, without specific information to the contrary, it is probably the case that my beliefs are your beliefs too. That is to say, 'true-belief' attribution is a rational prior. But, in false belief scenarios, the default 'true-belief' attribution provides a bias that has to be inhibited for a false belief content to be selected. On this account, failures are caused not by subjects reporting reality, but by their efforts to engage in ToM.

Most of the time, the ‘reality bias’ and the ‘true-belief bias’ will yield the same response. For example, in the standard Sally-Anne false-belief task, children younger than four usually predict Sally will look for the toy in its real position. Their failures could be either because they are reporting the reality (also see Schneider, Bayliss, Becker, & Dux, 2012; Schneider, Lam, Bayliss, & Dux, 2012), or because they are attributing their own belief (a ‘true belief’) to Sally.

However, if a ‘reality bias’ underlies our subjects’ anticipatory gaze in the HD condition of experiment 1, that is, instead of attributing belief, they were simply looking at the target’s real position, we would expect their first look to be towards the true-belief screen (just like their looking behavior in the TB condition). The fact that approximately half of them did not do so tends to speak against the ‘reality bias.’

Experiment 2

To better address the question of what produces the executive demands, a ‘reality bias’ or a ‘true-belief bias,’ we adapted the design of Southgate et al. (2007). In that study, an actor is shown standing behind a screen with her head poking above it. There are two windows in the screen, one on either side of the actor. In front of and under each window, there is a container. A puppet places and hides a toy in one of the two containers. Subsequently, the two windows illuminate, accompanied by a chime; together, these signal that the actor is about to reach for the toy. Infants were familiarized with this video in two trials, shown on an eye tracker. To test spontaneous belief attribution, infants were then shown a video in which the puppet entirely removed the toy from the scene while the actor had her back turned. The actor then turned around to face the scene, and the anticipatory signal was given while their eye gaze was recorded. This was a low-demand (LD) false belief condition because the target was removed from the scene.

Southgate generously shared her video with us, and we edited this to make two further videos that could be used to test true belief (TB) and high-demand false belief (HD). In the TB video, as the puppet is relocating the toy into the other container, the actor turns back to face the scene, and sees the toy in its final position. In the HD video, the actor does not turn back until after the toy has been relocated and the puppet leaves the scene empty handed. Therefore, in the HD condition, the actor should falsely believe that the toy is still in the original container.

Methods

Subjects were 87 two- to three-year-olds (47 girls; mean age = 36.3 months, range = 24.2 – 47.9 months) and 172 Rutgers University undergraduate students (85 females; mean age = 19.6 years, range = 17.0 – 28.1 years). An additional child failed to complete because of experimenter error. For all analyses of anticipatory gaze (with one exception[2]), out of the aforementioned totals, 6 children were excluded due to calibration error, 1 child and 16 adults had no recorded looking behavior during the four seconds of anticipation.

[2] The exception is that, to measure subjects' anticipation of the actor's behavior, we only analyzed subjects' looking to the windows. Data from those subjects who did not pay sufficient attention to either window were excluded.

Subjects watched three videos showing two main familiarization trials and a single test trial. The videos for the familiarization trials and the low-demand false belief trial were the same as used in Southgate et al (2007). The scenes of all the videos are the same: an actor wearing a visor is shown standing behind a screen with her head poking above it. There are two square-shaped windows in the screen, one on either side of the actor. In front of and under each window, there is a cube-shaped container; the lid of each container is below the bottom of the window on each side.

In the LD false-belief condition test trial, the puppet first puts the ball in the left-side container and leaves. A phone ring starts to play at this point, and the actor turns her head around as if she is distracted and attending to the ring. While the actor has her back turned, the puppet comes up again, opens the lid of the left-side container, takes the ball out, opens and places the ball into the right-side container. After he closes both lids, the puppet pauses for a second in the center, and then opens the lid of the right-side container again, takes out the ball, closes the lid, and finally takes the ball entirely off the scene. The phone ring stops immediately after the puppet disappears, and the actor turns back to face the scene. The windows light up and the chime sounds, indicating that the actor is about to search for the ball. The windows light lasts two seconds and the image is frozen for another two seconds. No reaching outcomes were ever shown to the subjects.

The HD false-belief condition was obtained by editing the above video, such that the puppet left the scene empty-handed after he relocated the ball from the left-side container into the right-side one. The video for the TB condition was also obtained by editing the LD video, such that the actor turned back to face the scene while the puppet was putting the ball into the right-side container, and therefore saw the ball's final position. The container in which the actor last saw the ball being placed was counterbalanced across subjects.

A Tobii T60 XL eye tracker automatically recorded the latency of subjects' first saccade and the duration of their gaze towards each of five areas of interest: both windows, both containers, and the head of the actor (see Fig. 3). After the four seconds anticipation, adults were asked, "Where will she search for the ball?" Their responses were recorded as Left-container or Right-container from their own perspective, and further coded as 0 (inconsistent with the actor's belief) or 1 (consistent with the actor's belief).

Results & Discussion

If subjects' eye gaze spontaneously anticipates the actor's search, responses should be opposite in the LD and TB conditions. In the HD condition, there are three possible gaze patterns: the same as LD, or the same as TB, or evenly split as in experiment 1. Since the windows and containers in these videos are separated, the stimuli allow us to determine more precisely what the focus of subjects' interest is in these scenarios. If the interest is in the real position of the target ('reality bias'), then subjects should look at the containers; however, if the interest is in the actor's belief and the action comes from it ('true-belief bias'), they should look at the window where the action will first appear.

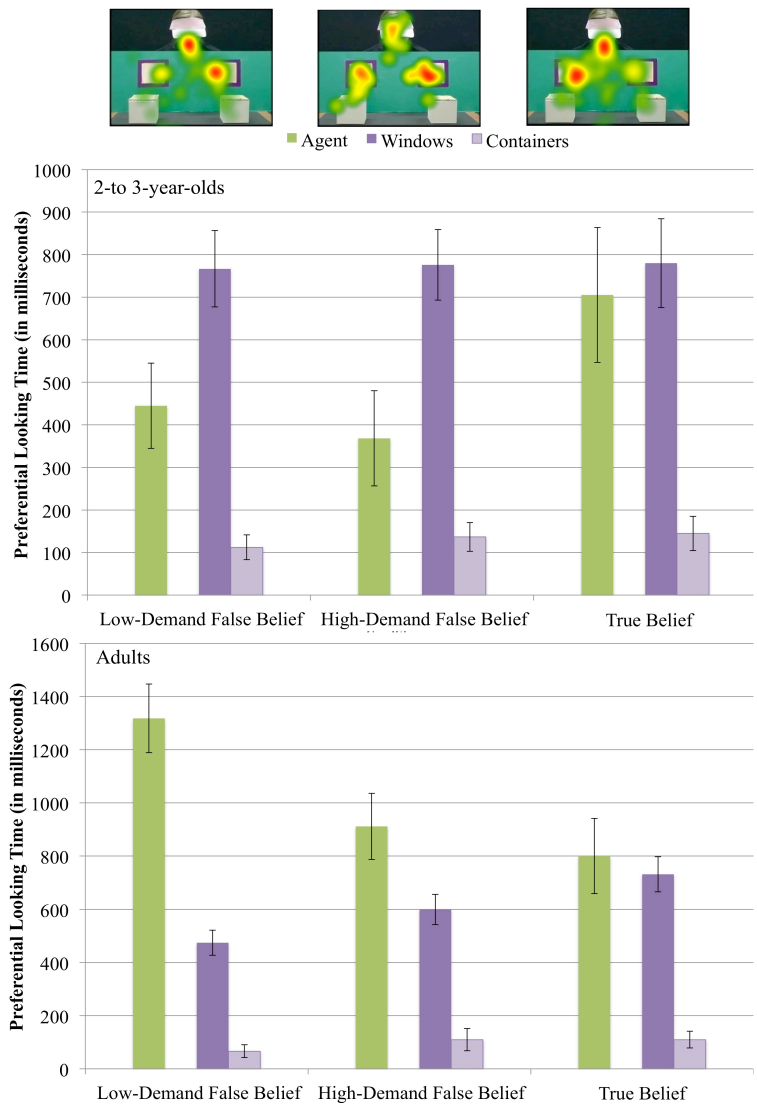


Figure 3: Two- to three-year-old children's and adult's looking time towards the windows, containers, and the agent's head over the 4 seconds of anticipation in experiment 2. Subjects' preferential looking was averaged across the two windows and two containers to facilitate comparison with looking to the agent. On the monitor of Tobii T60 XL Eye Tracker (24 inches) with a 800*600 screen resolution, the physical areas of each AOI are 94.1cm^2 for the agent's head, 56.8cm^2 for each window, and 54cm^2 for each container.

In general, subjects were interested in the agent and the windows, and they paid relatively little attention to the containers (as in Fig. 3; subjects looked significantly longer at the windows than at the containers, t (86) = 11.82, $p < .001$ for children, and t (171) = 12.01, $p < .001$ for adults). The preference for windows held no matter whether the target object was hidden in the scene (as in TB or HD conditions, t (53) = 9.29, $p < .001$ for children; t (105) = 9.62, $p < .001$, for adults) or completely removed from the scene (as in LD, t (32) = 7.20, $p < .001$ for children; t (65) = 7.41, $p < .001$ for adults).

Although subjects paid relatively little attention to the containers overall, was their looking at each container nevertheless biased by the hidden object ('reality bias')? To examine this, we first combined subjects' looking times to either of the containers in the HD and TB conditions (because the target object was on scene but hidden in these videos), and compared that to subjects' looking times to the containers in the LD condition (because both containers were empty in this video). Subjects did not look longer to the containers when the target was hidden on scene (t (85) = 0.70, $p = .487$ for children, Bayes Factor = 4.73:1, odds favoring the null; t (170) = 1.09, $p = .278$ for adults, Bayes Factor = 4.64:1, odds favoring the null). Secondly, when they looked at the containers in the HD and TB conditions, subjects spent equal amounts of time looking at the empty one and the one with the hidden object (combined subjects' looking times to the TB container in the HD and TB condition, and compared that to the combined looking times to the empty container in the HD and TB condition, for children, t (53) = 0.97, $p = .336$, Bayes Factor = 5.92:1, odds favoring the null; for adults, t (105) = 0.96, $p = .341$, Bayes Factor = 8.29:1, odds favoring the null).

Therefore, taken together, subjects' looking patterns in experiment 2 provide substantial evidence that there was no bias to look at the hidden object, contradicting the 'reality bias' hypothesis[3].

Looking at the actor's face cannot reveal which of her behaviors subjects anticipate. At the same time, the data presented above shows that looking at the containers is also not informative. Therefore, following Southgate et al (2007), we analyzed subjects' looking times to the windows as a measure of subjects' anticipation of the actor's behavior. Data from those subjects who did not pay sufficient attention to either window (for children, n = 3 in LD, n = 1 in HD, and n = 1 in TB; for adults, n = 17 in LD, n = 16 in HD, and n = 13 in TB) were excluded.[4]

In the LD condition, two- to three-year-olds and adults were more likely to shift first look to the window through which the actor should reach given her false belief (FB window) (19/27 children and 28/44 adults; one-tailed binomial tests, $p = .026$, and $p = .048$, respectively; see Fig. 2). The opposite looking pattern was observed in the TB condition: for their first look, subjects preferred the window through which the actor

[3] With regard to subjects' first look, the majority of the subjects shifted first look to the windows (73 out of 80 children and 144 out of 156 adults). When they looked first at the containers in the HD and TB conditions, subjects were not biased to look first at the container with the hidden toy (we combined the number of children and adults firstly looked at the TB container in the HD and TB condition, n = 10, and compared that to the combined number of children and adults firstly looked at the empty container in the HD and TB conditions, n = 13, Binomial test, $p = .34$, one-tailed; Bayes Factor = 3.27:1, odds favoring the null).

[4] As explained in experiment 1, low quality recording of looking may not accurately reflect subjects' anticipation. Therefore, data from the bottom 5th percentile (for children) and bottom 25th percentile (for adults) of total looking time to the windows in each condition were rejected. For children, the bottom 5th percentile of total looking time to windows corresponded to 33 ms in LD, 78 ms in HD, and 233 ms in TB; For adults the bottom 25th percentile of total looking time to windows corresponded to LD=479 ms, HD=654 ms, and TB=783 ms.

should reach given a true belief (TB window) (15/20 children and 21/31 adults; one-tailed binomial tests, p = .021, and p = .036, respectively). In the HD condition, however, subjects had no preference for first look (16/29 children and 20/42 adults; one-tailed binomial tests, p = .356, and p = .439, respectively; Bayes Factor for children 3.79:1 favoring the null and for adults 5.02:1 favoring the null). Given that the same pattern of anticipations was observed in both children and adults, we combined their first looks in each condition. Combined anticipations suggested that first looks between the LD and HD conditions, and between the HD and TB conditions were significantly different in the expected direction (LD vs. HD: Upton's χ^2 = 3.48, p =.031, one-tailed; HD vs. TB: Upton's χ^2 = 5.49, p = .010, one-tailed).

Given that no reaching outcomes were ever shown to subjects in test trials, their preferential looking time over the four seconds also indicated their spontaneous anticipations. We calculated a differential looking score (DLS) that reflected subjects' relative preference for looking longer at the FB window in all conditions. DLS is calculated by subtracting subjects' looking time towards the TB window from that to the FB window, divided by total looking time to both windows. Therefore, DLS will vary from -1 to +1, with a positive number indicating a preference for the FB window. Preliminary analysis suggested that children's age (two-or three-year-old), gender, initial position of the toy had no main effect on their DLS. Similarly, neither adults' gender nor toy's initial position played a role in adults' DLS. There were no interactions between the above factors and conditions (LD, HD, TB). These factors were dropped in further analyses.

Two by three ANOVA with age (two-to three-year-olds or adults) and conditions suggested that only the main effect of condition was significant ($F(2,187) = 8.29$, $p < .000$, $\eta^2 = .081$). To further explore the main effect of condition, two planned analyses were conducted using separate 2 (age) by 2 (two of the three conditions, LD & HD or HD & TB) ANOVAs on DLS. No age differences nor the interactions between age and conditions were observed; subjects' DLS were significantly different between LD and HD, and between HD and TB conditions, respectively (LD vs. HD: $F(2,138) = 4.41$, $p = .038$, $\eta^2 = .031$; HD vs. TB: $F(2,118) = 4.55$, $p = .035$, $\eta^2 = .037$). Specifically, in the LD condition, subjects showed a strong preference for the FB window (mean DLS = 0.22 for children (t (26) = 1.89, $p = .035$, one-tailed), mean DLS = 0.30 for adults (t (43) = 2.93, $p = .003$, one-tailed)), and in the TB condition, subjects showed a strong preference for the TB window (mean DLS = -0.28 for children (t (19) = 2.06, $p = .027$, one-tailed), mean DLS = -0.20 for adults (t (30) = 1.70, $p < .05$, one-tailed); see Fig. 4). In the HD condition, subjects spent equal amounts of time looking at the TB and FB windows (mean DLS = 0.01 for children (t (28) = 0.06, $p = .475$, one-tailed; Bayes Factor = 6.96:1, odds favoring the null), mean DLS = 0.03 for adults (t (41) = 0.34, $p = .367$, one-tailed); Bayes Factor = 7.85:1, odds favoring the null).

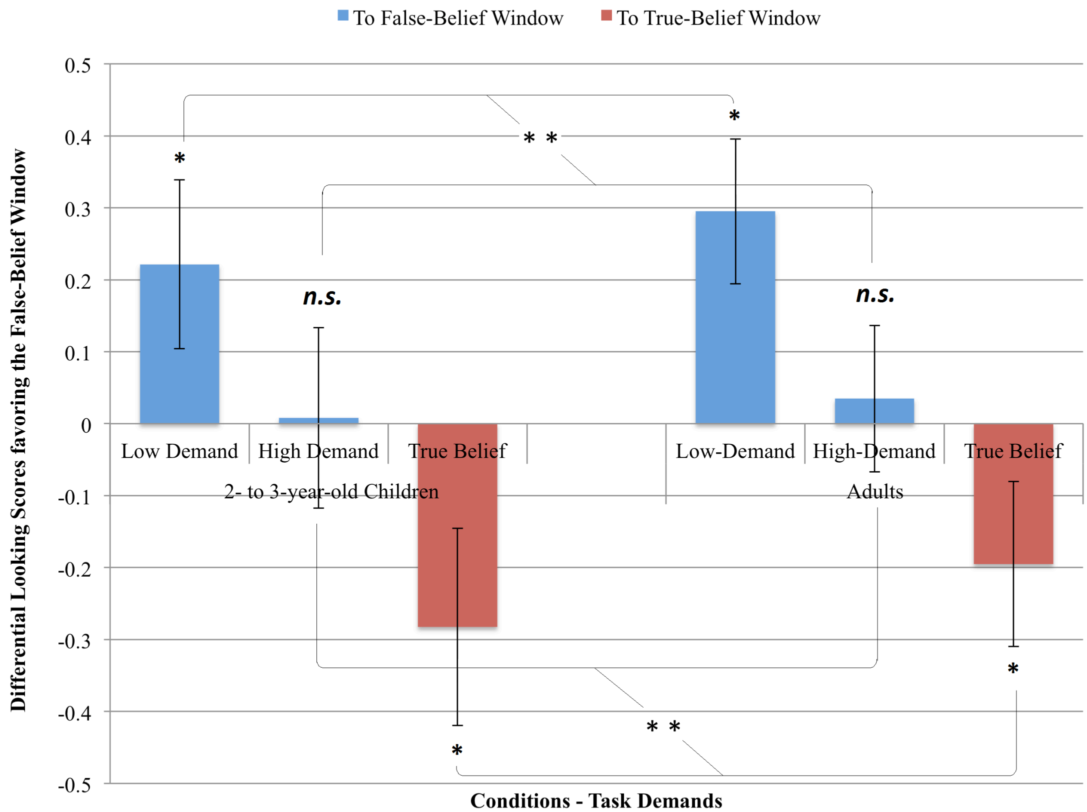


Figure 4 Results of Differential Looking Scores (DLS), Experiment 2. A positive DLS represents a preference for (looking longer at) the False-Belief Window over the 4 seconds of anticipation. A negative DLS represents a preference for (looking longer at) the True-Belief Window over the 4 seconds of anticipation.

* = one-sample t-test, $p < .05$, one-tailed; ** = 2 by 2 ANOVA, $p < .05$, two-tailed.

As in experiment 1, adult subjects were asked, “Where will she look for the ball?” immediately after the videos, and they correctly predicted the actor’s behavior based on her belief in all three conditions. Results were reported for all adult subjects (172), and for those who were included for the anticipatory looking analysis (117), separately: for all adult subjects, 60/66 in LD, 57/62 in HD, and 38/44 in TB correctly predicted the actor's behavior given her belief (two tailed binomial tests, all *P's* < .001). Their verbal responses were more accurate than their spontaneous anticipation by first look in all conditions (McNemar Binomial tests two-tailed: in LD, $p < .001$; in HD, $p < .001$; in TB, $p = .013$). For those who were included in the analysis given their looking behavior towards the windows, 42/44 in LD, 38/42 in HD, and 26/31 in TB correctly predicted the actor's behavior based on her belief (two tailed binomial tests, all *P's* < .001). Again, their verbal responses were more accurate than their spontaneous anticipation by first look in both LD and HD conditions (McNemar Binomial tests two-tailed: in LD, $p = .001$; in HD, $p < .001$; in TB, $p = .227$).

To sum up, we replicated the findings of experiment 1 that spontaneous anticipation measured by anticipatory eye gaze is still subject to the executive demands of false belief tasks. Moreover, neither children's nor adults' anticipatory gaze patterns reflected a bias to attend to the physical presence of the hidden object, inconsistent with a 'reality bias'; but instead the data was more consistent with a bias to attribute a true-belief to the actor whenever possible, namely, whenever the content of such a belief is known to the subject. The source of the executive demands in false belief tasks appears to be the employment of ToM ability itself.

Discussion

For many years, it was widely assumed that false-belief reasoning first developed after four years of age (Perner, 1991; Wellman, Cross, Watson, 2001). Contrary to this assumption, the findings of experiments 1 and 2 support more recent findings that subjects spontaneously attribute true- and false-beliefs in the second year of life when measured by various spontaneous behaviors (Onishi, Baillargeon, 2005; Southgate, et al. 2007; Buttelmann, et al., 2009; Scott, et al., 2010; Kovács, et al., 2010). Successful spontaneous belief reasoning has also been reported in toddlers and preschoolers (Clements, Perner, 1994; Garnham, Ruffman, 2001; He, Bolz, & Baillargeon, 2011; Scott, He, Baillargeon, & Cummins, 2012).

The findings from early success tasks support an early competence view of ToM (Scholl, Leslie, 2001; Roth, Leslie, 1998). However, the wealth of findings from the later failure tasks does not simply disappear. Early competence theories explained later failure in terms of processing demands that the child is not yet able to meet (Scholl, Leslie, 2001; Roth, Leslie, 1998; Leslie, Thaiss, 1992; Bloom, German, 2000). Given this explanation, it is natural to assume that the early success tasks do not make these demands (Baillargeon, Scott, & He, 2010). The present studies address this hypothesis directly and examine the long-term continuity of spontaneous early success measures.

We varied the processing demands of belief reasoning tasks in experiments 1 and 2 and measured two-to three-year-old children's spontaneous anticipations. As predicted by early competence theories, we observed opposite gaze patterns in the low-demand false-belief and true-belief scenarios. Two- to three-year-old children expected an actor to behave in accord with her false belief when the actor believed that a now-absent target object was still hidden in one of the containers. Children anticipated the actor would behave consistently with a true belief when the actor saw the target object being placed into its last position. However, children showed a third pattern of eye gaze in the high-demand false-belief scenario by anticipating that the actor would search equally in either container when, unbeknownst to the actor, the target object had been relocated from one container to the other. This pattern of anticipation cannot be explained by a preschooler's lack of competence, since the same pattern was also found in the adult subjects, whose competence is not in question (also see Low & Watts, 2013; Schneider, Bayliss, Becker, & Dux, 2012; Schneider, Lam, Bayliss, & Dux, 2012). Indeed, our adult subjects correctly predicted the actor's behavior when asked verbally, subsequent to viewing the videos, thus, demonstrating a dissociation between eye-gaze and verbal measures. The assumption that early-success false-belief tasks are free of processing demands is wrong.

What explains the above patterns of anticipation? One proposal is that overcoming a bias to respond to reality creates high processing demands. Nevertheless, the gaze patterns from experiment 2 shows that neither preschoolers nor adults spend much time looking at the hidden object's real position; even when they looked at the containers, they did not distinguish between the empty container and the one with the hidden object. Instead, the evidence favors a bias towards attributing a true belief to the actor *whenever the subjects know the current location of the object*. In many cases, a 'reality bias' and a 'true-belief bias' are easily confused because, for the belief attributer, 'reality' is the same as what he/she believes is true. However, responding only to reality does not actually require ToM; by contrast, attributing a true belief to someone does.

Indeed, there are theoretical reasons for supposing that any successful belief-desire reasoning system will have a 'true-belief bias'. Most of our simple beliefs about everyday matters usually are similar. Therefore, in the absence of specific information to the contrary, my own belief that P will always be my best guess about your belief that P. This means that, for a belief-desire reasoning system, attributing a 'true belief' *by default* is an efficient and rational prior. Accordingly, the ‘true-belief bias’ should be a fundamental part of ToM competence itself.

In the low-demand false-belief condition, subjects did not know the current location of the object. Thus, they could not attribute any specific location for the ‘true belief’, making its default attribution moot or at least easier to overcome. In the high-demand false-belief condition, by contrast, the ‘true belief’ was obvious, making its default attribution harder to inhibit. In this condition, however, subjects’ anticipatory gaze was not solely driven by the ‘true belief’, but was equally balanced between action predictions from the ‘false belief’ and the ‘true belief’. This behavioral pattern supports the hypothesis proposed by Leslie and colleagues that the theory-of-mind mechanism calculates both contents for belief attribution – own belief for the ‘true’ content and the agent’s perceptual exposure for the ‘false’ content (Leslie, German, & Polizzi, 2005). Previous studies measuring children’s verbal responses revealed only the outcome of selection between all plausible candidates. The spontaneous eye gaze to the windows in our studies provides direct evidence for the proposal that the theory-of-mind mechanism actually computes two candidates, a default ‘true belief’ and a context-contingent ‘false belief’. Interestingly, the eye gaze pattern we observed shows that both contents are equally attractive.

Why do infants succeed in high-demand false-belief tasks when measured by VOE looking time? In the study by Onishi & Baillargeon (2005), the target object was hidden in the scene, and infants looked significantly longer when the actor with a mistaken belief searches the object’s real position. Thus, infants’ looking times apparently overcame the ‘true-belief’ default. One possibility is that VOE tasks are less sensitive to tasks' processing demands than anticipatory looking tasks. Interestingly, when appropriate comparisons could be made, Scott et al. (2010, p.383) found evidence that infants looked longer when viewing behavioral violations of false beliefs than when viewing behavioral violations of true beliefs. This finding suggests that even VOE looking time can measure the increased processing demands of tracking an actor's false belief. Future research is needed to explore the possibilities.

Why do children succeed in some high-demand false-belief tasks measuring their anticipatory looking? Clements & Perner (1994) developed tasks measuring subjects' anticipatory looking embedded in a verbal Sally-and-Anne story, by using a narration saying "I wonder where she is going to look" in a self-addressed manner and measuring children's eye gaze after the prompt (Clements & Perner, 1994; Garnham & Perner, 2001; Garnham & Ruffman, 2001). This line of research was carried on recently by Scott, Baillargeon and their colleagues (Scott, et al., 2012; He, et al., 2012). For example, in Scott et al. (2012), they showed children a special storybook in which each page of the book presented two scenarios, one consistent with the story narration while the other was not (e.g. the narration said "Emily has an apple" when a picture of an apple and a picture of a banana were presented). Children heard a story similar to the Sally-and-Anne task in which a girl (Emily) put her apple in location A, and it was relocated to location B when Emily was away. In the end, Emily came back, and the narration said "Emily is looking for her apple." Again, children saw two pictures of Emily reaching towards different boxes, one consistent with her false belief while the other showing her reaching where the apple really was. Scott et al. (2012) found that 2.5-year-old children preferred looking at the picture consistent with the girl's belief states, even though it was a high-demand task.

Are our findings in the high-demand condition contradicting the results of Garnham (née Clements) - Scott line of work? We believe not. Particularly, there are important task differences between our studies and the Garnham - Scott line of work. Firstly, they used verbally presented tasks and measures eye gaze manually (by observers analyzing video recordings of the subjects' gaze), whereas our tasks are non-verbal and measure eye gaze automatically using a Tobii eye tracker. Additionally, we measure anticipatory eye gaze occurring within 600 ms of a prompt (experiment 1) and four seconds of a signal (experiment 2), whereas Scott et al. (2012) discards eye gaze within the first two seconds and analyzes gaze as it evolves over the subsequent 6 seconds. There are, therefore, substantial differences in time scale. It is possible that the selection between true-belief and false-belief candidates can be resolved in a longer period, such that the balanced looking pattern we observed in the high-demand condition might be replaced by a positive looking towards false-belief option later on (some behavioral and neuroimaging evidence suggests that the larger the perceived differences between self and other, the longer it takes for the subject to judge other's preference (Tamir & Mitchell, 2013), and the stronger the neural responses at MPFC, a brain region largely involved in inhibition, is (Tamir & Mitchell, 2010)). Moreover, the contrast that has emerged in recent years between early non-verbal and later verbal ToM reasoning rests on how different neurocognitive mechanisms are engaged by different tasks at different points in development. The differences in time scale together with verbal factors are likely to affect which processes are being measured. Thus, we believe that there is no contradiction but instead an enrichment of the date set that any integrated theory must account for.[5]

In sum, our studies find evidence that two-to three-year-old children's spontaneous belief reasoning is subject to the same or similar executive demands as shown in standard verbal tasks. This demand effect, when measured in the same way in the same tasks, persists unresolved in adulthood. We argue that default theory-of-mind reasoning, whenever possible, spontaneously attributes a 'true belief'. The 'true-belief' prior remains a cornerstone of theory-of-mind competence throughout the lifespan, and, for the brain systems controlling spontaneous eye gaze, is never overcome.

[5] Interestingly, Low & Watts (2013) designed a non-verbal study similar to the high-demand false-belief task in experiment 2, and they found that both children and adults successfully anticipated the actor's behavior. However, instead of measuring the precise focus of subjects' gaze using an automatic eye tracker, they manually coded and categorized subjects' looking at the scene as to the left or to the right. It is curious how the coding based on a force-choice between left-and-right may categorize looking that would otherwise be categorized as towards the actor's head using an automatic eye tracker (e.g., we found that subjects spent significant amount of time looking at the actor's head, as in Fig.3). Thus, without comparable observations, we leave it an open question whether the seeming discrepancy is real.

III. Can toddlers track two false beliefs

In chapter 1, I presented the argument that the seeming regression in development from infants' successes to three-year-olds' failures in false belief reasoning is readily explained by an early competence theory proposed by Leslie and his colleagues (Leslie & Thaiss, 1992; Leslie, German, & Polizzi, 2005).

Two studies in chapter 2 provide evidence that the early competence is subject to tasks' processing demands. When children knew the target object's whereabouts, and therefore, the true-belief content was salient, their anticipation was balanced between the true- and false-belief option; however, when children were uncertain about the truth, and therefore, the true-belief was moot or at least easier to inhibit, children were able to anticipate an actor's behavior given her false belief spontaneously. Surprisingly, adults' spontaneous anticipation showed the same pattern as children's, even though they had no problem predicting an actor's behavior in a verbal high-demand false belief task. These findings suggest a dissociation between spontaneous and verbalized ToM reasoning. Specifically, the spontaneous system has limited access to one's inhibitory resources, while the verbalized ToM reasoning benefits from its interaction with one's executive ability.

In the current chapter, I will present two studies showing that children's spontaneous reasoning of false belief is flexible in the sense that belief representations are bound to specific agents.

Introduction

Studies on theory of mind (ToM) have mostly focused on origins, on documenting the first signs that a child can understand another person's mental sates. However, our social environment usually involves multiple people, for example, meals shared with the family or conversation with friends at a party or colleagues in a work group. Navigation in our social world typically requires us to track multiple agents and their social relations (Meyer & Lieberman, 2012; Meyer, Spunt, Berkman, Taylor, & Lieberman, 2012). Yet, it is largely unknown what our capacity to process simultaneously the minds of multiple agents is, or how this ability develops in children.

In contrast, the simultaneous tracking of multiple non-social objects has been extensively studied since Miller's (1956) classic demonstration that WM is limited. The brain may use two distinct systems to represent objects, namely, the "where/for-action" system that processes objects' spatiotemporal position, and the "what/for-knowledge" system that processes objects' identifying features, such as color and shape (e.g., Goodale, Milner, 1992; Ungerleider & Mishkin, 1982). Adults can simultaneously keep track of the locations of about four randomly-moving featureless objects (Pylyshyn & Storm, 1988; Scholl & Pylyshyn, 1999). When identifying featural information is bound to object representations, a trade-off appears between the number of objects adults can track and the amount of information about each that adults can remember (Alvarez & Cavanagh, 2004; Awh, Barton, & Vogel, 2007; Eng, Chen, & Jiang, 2006; Zhang & Luck, 2008; for a recent review, Fukuda, Awh, & Vogel, 2010). Thus, any information bound to an object representation tends to reduce the maximum number of objects that can be actively retained in WM. The distinction between limits on objects in locations and on identifying featural information also seems relevant to the development of object WM in infants (for a review, see Leslie, Xu, Tremoulet, & Scholl, 1998; Feigenson, 2007; Feigenson & Carey, 2003; Káldy & Leslie, 2003, 2005; Kibbe & Leslie, 2011, 2013; Leslie & Káldy, 2007; Mareschal & Johnson, 2003; Ross-Sheehy, Oakes, & Luck, 2003).

An analogous distinction may provide a framework for studying WM in theory of mind. Agents are a kind of object, they move around, and each has a distinct identity that needs to be tracked when we navigate small social groups. Beliefs, on the other hand, need to be 'attached' or bound to specific agents because two agents may have either the same or different beliefs. Therefore, as far as WM is concerned, analogies might hold between object- and agent-indexes, on the one hand, and, on the other, between an object's features and an agent's belief. These analogies lead us to ask whether maintaining the bindings between multiple agents and their various beliefs about a situation would create a WM load.

Although there have been suggestions that working memory capacity crucially limits preschool children's performance in single-agent false-belief tasks (Carlson, Moses, & Breton, 2002; Davis & Pratt, 1995; Gordon & Olson, 1998), the question has never been directly investigated.[6] In experiment 1, using a verbal task, we ask whether three- and four-year-old children can track the beliefs of two agents each of whom has different false beliefs. In experiment 2, we present a non-verbal double-agent task modeled on the verbal story, and measure two-year-old children's anticipatory eye gaze to determine whether spontaneous ToM can track multiple false-beliefs.

---

[6] Tracking multiple people's beliefs is implied in second-order false belief tasks in which children have to reason about a person's understanding of another party's belief that P (e.g. "John falsely believes that Mary thinks that P", Perner & Wimmer, 1985; Sullivan, Zaitchik, Tager-Flusberg, 1994). Second-order false belief, however, requires more than the binding between belief representations and agents, but understanding the recursive nature of mind.

Experiment 1: Verbal double-false-belief task

We designed a single-false-belief and a novel double-false-belief scenario, both modeled on the standard false belief task. As introduced in chapter 1, a number of studies over the years have found that undermining the subject's confidence in their own knowledge of the target object's location makes it easier for them to attribute a false belief to the actor. For example, in Carpenter, Call, and Tomasello (2002), an actor points at a box in which she previously placed a novel object — but which, unbeknownst to the actor, had subsequently been removed — and says, "Let's get the dax!". Compared to another scenario in which the actor points at a box in which she previously placed a novel object, but, unbeknownst to her, had subsequently been replaced by another novel object, in the situation with an empty box, the child herself does not have a specific competing belief about the content of the box, making it easier for the child to attribute a different belief to the actor (also see Bartsch, 1996). Following chapter II, We label this a 'low-demand' false belief task.

We used the low-demand task in our studies. In the single-false-belief story, Sally hides her cookie in one of two boxes, then, unbeknownst to her, a puppet takes her cookie and runs away, departing the scene entirely. In the double-false-belief task, after Sally has hidden her cookie and departed, Anne takes it and hides it in another box, and departs too. Then, a puppet takes the cookie and runs off with it. Therefore, in both single- and double-agent stories, the character(s) have false belief(s) about the target object's final location, while its real location is unknown both to the story character(s) and to our subjects.

Methods

Forty-six three-year-olds and 24 four-year-olds from local preschools participated in this study. Out of these, 20 of the three-year-olds were assigned to the single-false-belief task (8 girls; mean age = 42.8 months, ranging from 36.3 − 47.9 months), and 26 to the double-false-belief task (15 girls; mean age = 44.5 months, ranging from 38.0 − 47.7 months). All 24 four-year-old children participated in the double-false-belief task (16 girls; mean age = 55.1 months, ranging from 48.4 − 59.1 months). Another 27 three-year-olds (11 in the single-false-belief task and 16 in the double-false-belief task) and 4 four-year-olds participated, but were excluded from the analysis because they failed control questions.

Participants were tested in a quiet room at their schools. The single-false-belief story was presented through a series of images in a binder. Children were told that a character, named Sally, has a cookie. She puts it into a red box on the table (a red and a yellow box were in the scene) and goes out to play. When she is away, a cat climbs onto the table and opens up both boxes. Then the cat takes the cookies out of the red box, and runs away with it. No one knows where the cat goes. At this point, children were asked three control questions: "Did Sally see that [the cat took the cookie and ran away]?", "Where did Sally put her cookie at the beginning of the story", and "Where is the cookie now?" Children were corrected if they got any of the control questions wrong, and the story was told again (at most once). After that, they were asked a single belief question, "Sally is coming back. Where will she look for the cookie?" Children's responses were recorded and coded as "pass" if they answered "the red box", or "fail" if they said "the yellow box”, “don’t know” or “with the cat”. The original container where Sally put her cookie was counterbalanced between subjects.

The double-false-belief story was presented and tested with children in a similar way, except that there were two characters, Sally and Anne in the story. Sally originally puts her cookies in a red box and leaves the room. While she is away, Anne wants the cookie, so she takes the cookie out of the red box and puts that into the yellow one. After that, Anne leaves the room to talk to her mom. Then a cat climbs onto the table, takes the cookie out of the yellow box and runs away with it. No one knows where the cat goes. Similar to the procedure in the single-false-belief story, at this point, children were asked several control questions: “Did Sally see that?”, “Did Anne see that?”, “Where did Sally put her cookie at the beginning of the story?”, “Where did Anne put it?” and “Where is the cookie now?” Children were corrected and the story was told again (at most once) if they got any of the control questions wrong. After that, they were asked two belief questions, “What if Sally comes back? Where will Sally look for the cookie?” and “What if Anne comes back? Where will Anne look for the cookie?” Children’s responses were recorded and coded as “pass” if they answered “the red box” to Sally’s belief and “the yellow box” to Anne’s belief. Any other answers were coded as “fail”, such as “don’t know” or “with the cat”. The original container where Sally put her cookie was counterbalanced between subjects. Half of the children were asked to predict Sally’s action first and Anne’s second. The other half received the test questions in the reverse order.

Results

In the single-false-belief task, 15 out of 20 three-year-olds correctly predicted Sally's action given her false belief (Binomial test, $p = .041$, two-tailed). The other five children failed by attributing their own true belief to the agent (e.g. saying "no where" or "with the cat").

In the double-false-belief task, children can answer each prediction question with one of the three options: correctly predict the agent's behavior given her false belief; making a binding error by attributing one agent's false belief to the other; or making a true-belief error by attributing their own belief about the cookie to one or both agents. Together, there are nine possible combinations to answer the prediction questions; therefore the probability of making any one response pattern is 0.11.

The majority of both three- and four-year-old children were able to correctly attribute false beliefs to both agents (15 out of 26 three-year-olds (57%) and 18 out of 24 four-year-olds (75%); Binomial tests, ps < .001, two-tailed; Bayes Factors (BFs) favoring H_1 > 100:1, both age groups). Children rarely made mistakes by binding the belief to the wrong agent (one three-year-old (4%) and two four-year-olds (8%); Binomial tests, *p*s > .05; BF = 4.10:1 for three-year-olds and BF = 6.38:1 for four-year-olds, favoring the null of no systematic binding errors). The only systematic error was observed in three-year-olds who failed by attributing their own true belief to both agents (Binomial test, $p = .011$; BF = 9:1 favoring H_1 that the error of attributing one's own belief to both agents was different from chance). Four-year-olds no longer made this error (Binomial test, p > .05, BF = 3.84:1 favoring the null of no systematic true-belief error). Table 1 summarizes these results.

Table 1. Error analysis in the double-false-belief task. All Binomial probabilities have chance level = 1/9. Also shown are Bayes Factor testing either the null (H_0) or H_1 that a given number is different from chance (Gallistel, 2009).

Age	**No errors**	**Binding errors**			**One binding error & one TB error** Text		**True-belief errors**			**N**
		one agent wrong		both agents wrong			one TB error		both TB errors	
	A1A2	A1A1	A2A2	A2A1	A2-T	T-A1	A1-T	T-A2	T-T	
Threes	15 ($p < .001$; BF >> 100:1) (H_1)	0	0	1 ($p = 0.398$; BF = 4.1:1) (H_0)	0	0	2 ($p = 0.872$; BF = 6.4:1) (H_0)	0	8 ($p = 0.011$; BF = 9.0:1) (H_1)	26
Fours	18 ($p < .001$; BF >> 100:1) (H_1)	0	0	2 ($p = 0.984$; BF = 6.4:1) (H_0)	0	0	0	0	4 ($p = 0.548$; BF = 3.8:1) (H_0)	24

Discussions

Consistent with previous findings, three-year-old children succeeded in the single false belief when they were not certain of a target object's precise whereabouts (Bartsch, 1996; Carpenter, Call, and Tomasello, 2002; Wellman & Bartsch, 1988; Zaitchik, 1991). In a meta-analysis on children's false-belief reasoning, Wellman, Cross & Watson (2001) found that reducing subjects' certainty of a target object's whereabouts only improved three-year-olds' performance to a chance level, not to success (better than chance). It is possible that the three-year-olds in our study were older than those tested previously (mean age was 3.5 years), or the design of our study made the disappearance of the target object more salient (the cat runs away with the cookie, and no one knows where the cat is gone). Thus, we found a better-than-chance performance in three-year-olds in the single false-belief task.

More importantly, both three- and four-year-old children were able to bind two respectively correct false beliefs to two agents simultaneously. When they erred in action prediction, three-year-olds failed by attributing their own knowledge to the agents, while they rarely made any binding mistakes by attributing one agent's belief to another. Four-year-olds rarely made either type of mistake. These findings suggest two things. First, working memory capacity to bind false belief representations to agents can accommodate at least two agents with distinct false beliefs. Second, even in the low-demand task, three-year-old children's problems in false belief tasks are due to intrusion of their own knowledge (Leslie, Friedman, & German, 2004; Leslie, German, & Polizzi, 2005; Roth & Leslie, 1998) rather than working memory limitations (Davis & Pratt, 1995; Gordon & Olson, 1998).

Is sufficient working memory capacity to track two agents' false beliefs available only in explicit belief-desire reasoning? To explore whether children can represent more than one belief and bind the representation to more than one agent in an implicit task with a spontaneous behavioral measure, in experiment 2, we gave two-year-olds a nonverbal anticipatory looking task with double agents and double false beliefs.

Experiment 2: Spontaneous double-false-belief task

Methods

Subjects were 40 two-year-olds (19 girls; mean age = 30.7 months, s.d. = 3.72 months). They were randomly assigned to condition "Sally" (n=21) and condition "Anne" (n=19). An additional 17 two-year-olds were tested but excluded because their anticipatory looking was biased to one side in all trials (3 children) or tracking quality was below the 25th percentile in the corresponding test trial[7] (14 children, 6 in condition Sally).

A Tobii T60 XL Eye Tracker was used to present the stimuli and record subjects' eye gaze via Tobii studio software. Children either came with their parents to the lab or participated at their preschool. When tested at preschool, children sit on a chair by themselves in a quiet room, 60 to 80 centimeters in front of the eye tracker (adjusted for tracking quality in calibration). When tested in the lab, children sat on their parent's lap, 60 to 80 centimeters in front of the monitor in a separate booth surrounded on three sides with dark curtains. A 5-point moving toy calibration was applied before the experiment.

[7] The bottom 25th percentile of total looking time to windows corresponded to 194 ms in Sally's condition and 0 ms in Anne's condition.

Each subject watched five videos with two pre-familiarization trials, two familiarization trials and a single test trial. Fig.5 shows the general scene: There are two wooden boxes on stage, one green and one red. In the back wall of the stage, there are two square windows, one above each box. Both windows are covered by a blue curtains. Two purple curtains are hanging behind the stage, with a gap as wide as a person in the middle. In the two pre-familiarization trials, an actor wearing a visor and an orange shirt is shown standing in the center behind the back wall of the stage, with her head and shoulders showing above it. In the first pre-familiarization trial, there is a toy on top of the green box. Immediately after the trial starts, a chime sounds and the two windows illuminate for two seconds; then, the actor reaches through the left window and retrieves the toy. The second pre-familiarization trial is identical to the first one, except that the toy is on top of the red box and the actor reaches through the right window to retrieve it.

In the two familiarization trials, two actors are shown standing behind the back wall of the stage. The actor on the left is wearing an orange shirt (Sally), and the actor on the right is in a light-green shirt (Anne). At the beginning of the first familiarization trial, a puppet is present at the center of the stage, and it climbs into the green box on the left (all the while both actors are following the puppet with obvious head movements). Sally says, “Let me get its carrot”, and disappears behind the purple curtain on the left. Immediately, Anne says, “Let me get its ball”, and disappears behind the purple curtain on the right. Then, Sally appears from behind the curtain on the left. She lifts her hand to show that she has got a carrot, and says, “Here it is!” Then she steps forward to the center behind the back wall, and the anticipatory signal of chime and windows illumination occurs. After two seconds, Sally reaches through the left window, puts down the carrot down by the green box, opens up the box and retrieves the puppet. The second familiarization trial is similar to the first one, except that the puppet climbs into the red box on the right, and Anne comes back in the end with a ball and retrieves the puppet from the red box.

In the test trial, a puppet dog climbs into the green box, and Sally says, “Let me get its bone”. Immediately Sally disappears behind the left-side curtain, the dog comes out of the green box, and Anne sees it. The dog moves to the red box on the right while Anne is watching. After the dog enters the red box, Anne says, “Let me get its ball”, and goes behind the curtain on the right. The dog climbs out of the red box, moves center stage, and eventually jumps off-stage, center front. Then, either Sally (condition Sally) or Anne (condition Anne) comes back, and the anticipatory signal occurs, indicating that the actor is about to search for the dog. The final image of the anticipatory signal is visible for another two seconds while latency of subjects’ first saccade and duration of their gaze to each area of interest (AOI) - windows and the actor - was automatically recorded. No reaching outcomes were shown in the test trial. The box in which the dog entered first was counterbalanced between subjects.

Results

With respect to children’s first look towards the windows, their anticipations in Sally and Anne conditions were significantly different in the expected direction (Fisher’s exact test, p = .001, one-tailed). Specifically, in condition Sally, 15 out of 21 children shifted first look to the window consistent with Sally’s belief (Sally’s window; one-tailed Binomial test, p = .039). By contrast, in condition Anne, only 4 out of 19 children’s first look was at Sally's window, while the other 15 looked first at the window consistent with Anne's belief (Anne's window; one-tailed Binomial test, $p = .010$) (Fig.5).

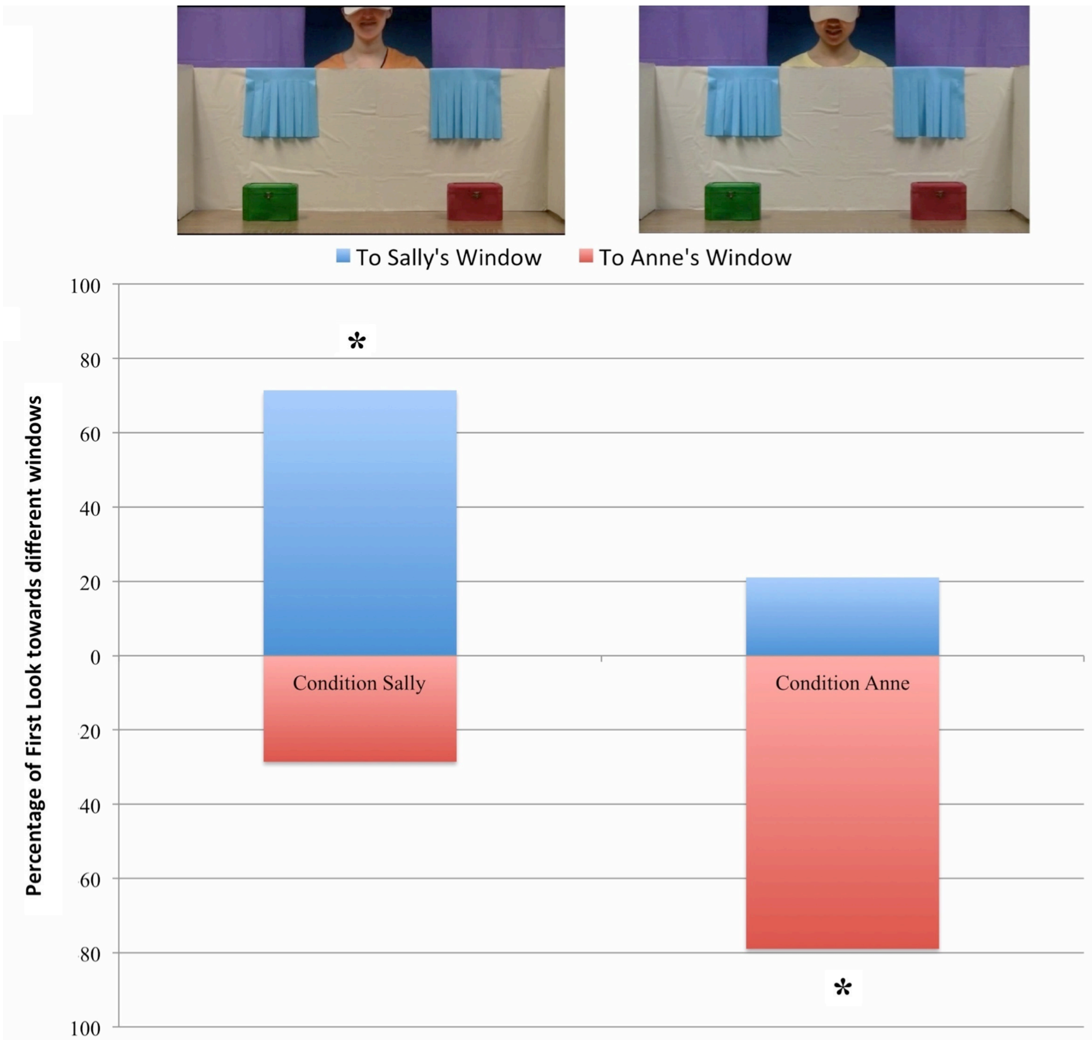


Figure 5. Results of anticipatory first look in experiment 2. Percentage of two-year-old children's first look towards the window consistent with Sally's false belief or the window consistent Anne's false belief.

* = Binomial test, $p < .05$, one-tailed.

Given that no reaching outcomes were ever shown to subjects in the test trials, their preferential looking time over the four seconds from the anticipatory signal also indicated their spontaneous anticipations. A differential looking score (DLS, Senju & Csibra, 2008; Senju, et al., 2009) was calculated by subtracting subjects' looking time towards Sally's window from that to Anne's window, and divided by total looking time to both windows. Therefore, DLS varied from -1 to +1, with a positive number indicating a preference for Sally's window and a negative number indicating a preference for Anne's window.

One sample t-tests on subjects' DLS in each condition confirmed the results from the first look measure (Fig.6). Specifically, in condition Sally, children had a preference for looking at Sally's window during the four seconds (mean DLS = 0.37, t (20) = 2.41, p = .026, two-tailed); in condition Anne, they showed the opposite preference and looked longer at Anne's window (mean DLS = -0.35, t (18) = -2.13, p = .047, two-tailed).

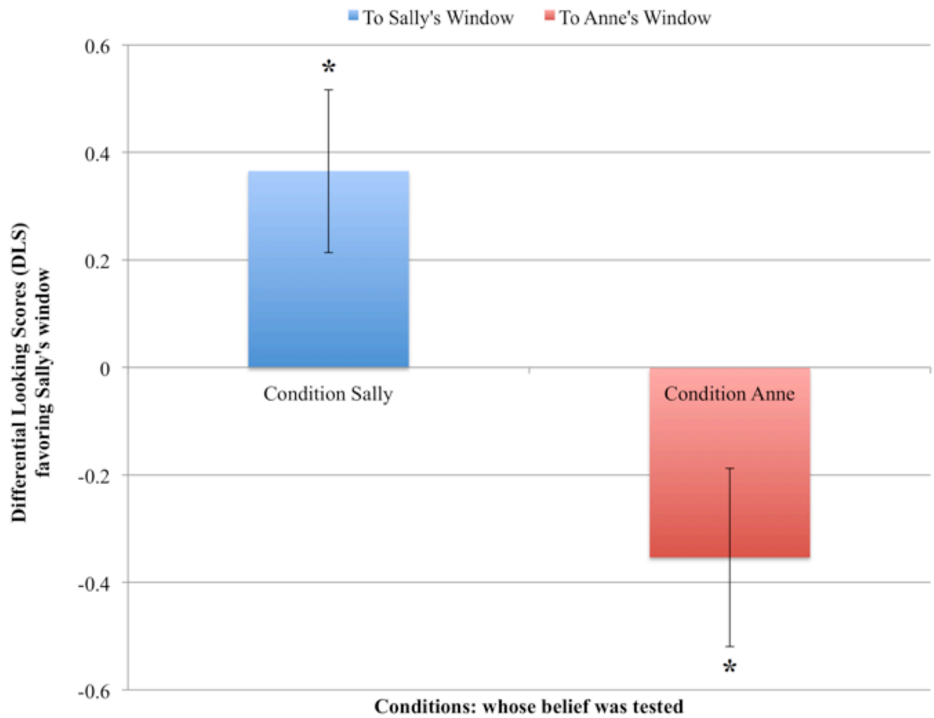


Figure 6. Differential Looking Scores (DLS) in experiment 2. A positive DLS represents a preference for Sally's window over the four seconds of anticipation. A negative DLS represents a preference for Anne's window over the four seconds of anticipation. * = one-sample t-test, $p < .05$, two-tailed.

Discussions

We tested only one character's false belief per subject, randomly assigned. We reasoned that because subjects could not know which character would be relevant, they had to track both to succeed. However, in the literature of tracking multiple objects, Káldy & Leslie (2003) pointed out that "the question of whether infants attend to all of the objects or just to one of them arises." Applying the question to the current study, can two-year-olds succeed in experiment 2 by only representing one agent's belief? We consider two possible models of representing one agent's belief — children were representing the same agent's belief (model 1), or children were tracking one agent's belief, selected at random (model 2).

Following Káldy & Leslie (2003) that attending only to the last-hidden agent would require a minimal memory load, while attending to the first-hidden agent would be more demanding, model 1 suggests that if the child only tracked one agent's belief, it would be Anne's. Therefore, in condition Sally, children would have performed at chance. Our findings that two-year-olds correctly predicted Sally's action in condition Sally spoke against this model. According to model 2, children may have tracked either agent's belief, selected at random. In other words, only half of the subjects in condition Sally (condition Anne) were tracking Sally's (Anne's) belief, while the other half were tracking Anne's (Sally's) belief, and therefore, half of the subjects would have performed correctly and the other half would have performed at chance. If model 2 were correct, the effect size of children's performance in each condition would be at most half the size as the effect of children's performance in a single-agent false belief task (as in chapter II). Our findings in chapter II showed that when subjects were tracking a single agent's low-demand false belief, the mean DLS favoring the correct window was 0.22, with an effect size of Cohen's d = 0.74. Two-year-olds' performances in experiment 2 were comparable to children's performances in the single-agent false-belief task, if not better, with an effect size of Cohen's d = 1.08 in condition Sally, and d = 1.00 in condition Anne. Therefore, model 2 seems an unlikely explanation of two-year-olds' success, and children were indeed tracking two agents' false beliefs in an implicit non-verbal task.

Discussion

Navigating the social world requires us to track multiple agents and their mental states simultaneously. A critical and unique property of agents that varies over time is their mental states. The non-stationarity of agents' mental states may favor a frequently updated WM, rather than a more slowly changing passive short-term store, but it demands the employment of executive resources (e.g., Engle, 2002). It is widely agreed that the development of executive functions influences performance on Sally and Anne false belief tasks from three years onwards (e.g., Carlson, Moses, & Breton, 2002; Roth & Leslie, 1998; Russell, Mauthner, Sharpe &Tidswell, 1991), though there are a range of views about what that role is exactly (Leslie, German, & Polizzi, 2005). Our results suggest that as soon as children are able to attribute a single false belief to a single agent, they are able to do so for two agents. This in turn suggests that performance may not be constrained by WM capacity itself and that changes in that capacity do not explain the famous three to four shift.

In experiment 1, three- and four-year-old children succeeded in representing both one and two characters' false beliefs explicitly. Children did not know the target object's precise whereabouts, making their true belief less salient and easier to be inhibited than usually posed in standard false belief task (i.e. Baron-Cohen, Leslie, & Frith, 1985). Despite being less salient, most of the errors three-year-olds' made were due to attributing their own indefinite 'true' belief. Four-year-olds made no systematic errors. Neither three- nor four-year-olds failed by making binding errors (misattributing one agent's belief to another).

In experiment 2, two-year-old toddlers tracked two people's false beliefs spontaneously as measured by anticipatory looking. These findings suggest that implicit WM capacity for agents and their beliefs also extends beyond a single agent.

Although we have not yet found the WM limit for theory of mind, we continue to assume there must be one. Whereas it is important to find where that limit lies, it is even more important to determine what in theory of mind reasoning produces the WM load, what exactly is limited and how. Leslie (1987, 1994; see also Kovács, Téglás & Endress, 2010) theorized that others' mental states are represented in memory as metarepresentions (representation of representation) which comprise four parts, "Agent - Attitude - Anchor - 'expression'" (Leslie, 1994, p217). Each part stores information about each of the different aspects of a mental state: an agent is identified who owns the mental state; the type of the mental sate (attitude) is identified (e.g., believes, desires, pretends, … etc.); an object (anchor) in the world is indexed which forms the topic of the mental state; a decoupled representation ("expression") is formed which is the content of the mental state. Each of these four parts may, theoretically, contribute to WM load. In addition, each part needs to be bound to the other parts correctly, again theoretically, the bindings could contribute to WM load. Taken together, this framework of ToM reasoning can be used as a guide for future research.

Our studies provide the first evidence that three- to four-year-old children can represent and track at least two agents and their false beliefs simultaneously in a verbal task, and that two-year-old toddlers can solve the same problem spontaneously in a non-verbal task. With an agenda to explore the WM limit in theory of mind (Cheng, Wang, & Leslie, 2013), and a theoretical framework to guide the research on what in theory of mind reasoning produces the WM loads, we believe that the binding problem represents a new frontier in theory of mind research.

IV. Spontaneous reasoning of second-order false belief

Studies in chapter 2 and 3 have shown that early understanding of ToM is both limited and flexible: it is limited by its lack of access to other cognitive resources, particularly inhibitory control, even in adults; it is flexible in the sense that it can track at least two agents and their distinct false beliefs simultaneously.

However, it is hotly debated whether the early competence in ToM is conceptual or behavioral in nature. Specifically, Perner & Ruffman (2005) summarized three possible non-theory-of-mind mechanisms to explain infants' success in spontaneous false belief reasoning. In this chapter, I wish to investigate the nature of the early competence, whether it is mentalistic (conceptual) or behavioral (non-conceptual), by testing children's spontaneous understanding of a second-order false belief.

Second order belief

Adults can easily entertain a person's understanding of another individual's mind. For example, readers of Romeo and Juliet will not forget the tragic climax of the drama: Romeo committed suicide because he did not know that Juliet had faked her death with the help of Friar Lawrence, and thus he falsely believed that Juliet had died. Besides that, we can also understand Friar Lawrence's worry when he heard that the messenger failed to notify Romeo about Juliet's fake death. In other words, we can appreciate a person's (Friar Lawrence's) understanding of another person's (Romeo's) false belief about some affairs in the world (Juliet's death). This ability requires an understanding of the recursive nature of mind (“Mary thinks that John believes that…”), beyond the representation of a first order false belief as tested in the Sally-and-Anne task.

Do children understand second-order belief? To study this, researchers developed second-order false belief tasks (Perner, Wimmer, 1985) in which children are asked to predict a character’s behavior given the character’s understanding (which is false) of another party’s belief (Perner & Wimmer, 1985; Sullivan, Zaitchik, Tager-Flusberg, 1994). For example, Perner & Wimmer (1985) told children that a girl Mary and her friend both saw an ice cream van at location A; later, Mary’s friend learned that the van was moving to location B, while he did not know that Mary also had learned about the change. Children were asked where Mary’s friend thought she would go to find the van, location A or B. To correctly answer the question (location A), children need to understand that Mary’s friend thought Mary still believed that the ice cream van was at location A, while Mary knew it was at location B. Therefore, each individual had a true belief about the van’s location, while one individual’s second-order belief about the other’s knowledge state was false. Studies using this story showed children did not understand second-order false belief until 7 (Perner & Wimmer, 1985; Hogrefe, Wimmer, Perner, 1986).

What causes the developmental lag between children’s performances in standard first-order and second-order false belief tasks? Perner (1988) claimed that understanding second-order false belief requires another conceptual shift on top of the development of belief concept between three-and four-year-olds, namely, the content of one’s belief could be another person’s mind. However, some researchers have suggested that the delay is caused by the second-order false belief task’s processing demands (Sullivan, Zaitchik, Tager-Flusberg, 1994; Miller, 2009; Tager-Flusberg, Sullivan, 1994). Indeed, the second-order task involves more characters, more scenes and much longer stories. When the story was simplified, children's performance was significantly facilitated, and children as young as five were able to reason about second-order false belief (Sullivan, Zaitchik, & Tager-Flusberg, 1994; Coull, Leekam, Bennett, 2006).

Might there be a way to demonstrate even earlier second-order belief competence in children? In light of the new wave of research showing infants' competence in false belief reasoning, tasks measuring children's spontaneous reactions might be more informative to test their understanding of second-order false belief. Cumulative studies have suggested that early success in ToM tasks only presents when children's spontaneous reactions are measured, no matter whether the stimuli are verbal or non-verbal (e.g. Scott, et al., 2012; He, Bolz, & Baillargeon, 2011; Clement & Perner, 1994; Ruffman et al., 2001).[8] For example, Scott et al. (2012, experiment 2) measured children's spontaneous understanding of first-order false belief in a verbal task, and it presents some prima facie evidence of younger children's second-order belief reasoning. In their study, 2.5-year-old children watched a live show of the standard Sally-and-Anne story, with an extra character in the show who would be asked to predict Sally's action. In other words, instead of being asked, "Where will Sally look for her marble?" children

---

[8] Two studies were conducted to investigate second-order belief with non-verbal stimuli (with an aphasia patient, in Apperly, Samson, Carroll, & Hussain, 2006; with children, in Hollebrandse, van Hout, Hendriks, 2012), and they measured subjects' explicit, rather than spontaneous, reactions. For example, Hollebrandse et al. (2012) showed 6- to 9-year-old-children a nonverbal video in which two people, A and B, both saw a toy hidden in a box at the beginning, and they individually witnessed that the toy was replaced by another object, while A did not know B also saw a replacement and vice versa. Children were asked explicitly what A thought B believed to be in the box. Children's performances were worse than their performances in a standard verbal second-order task.

watched another character in the video responding to the prediction question. Therefore, a second-order belief of “she thinks that Sally believes that P” was encouraged, and the second order belief was true. The authors took cautions in their interpretation, and explained 2.5-year-olds’ success a strong evidence of first-order false belief reasoning. This cautious interpretation is endorsed because children may have succeeded on this task given their own understanding of Sally’s first-order false belief and their own predictions of what Sally would do, without attributing a second order belief to the character per se. Similar arguments were leveled against a rich interpretation of understanding of true belief as evidence of a possession of belief concept (Dennett, 1978). Therefore, even though this is some tentative evidence in support of the claim that children as young as 2.5 can reason about second-order belief, it remains unclear.

We designed a study measuring children spontaneous looking behavior in a verbal task in which children’s success cannot be explained simply by their competence with first-order belief reasoning.

Current study

Our task adopted a communication scenario independently developed by Happé & Loth (2002) and Carpenter, Call & Tomasello (2002), and further extended by Southgate, Chevallier, & Csibra (2010). Studies suggest that toddlers can understand a speaker’s communicative intention and uncover the reference of the speaker’s labeling (e.g. Tomasello, 2001; Bloom, 2000, 2001; for a general account of the importance of intentionality in commutation, see Grice, 1957; Sperber & Wilson, 1986); that is, they understand what a speaker intends to refer to when he/she uses a novel label. For example, Tomasello & Akhtar (1995) showed that two-year-olds were sensitive to a speaker's intention of either introducing a novel object or producing a novel action, and mapped a novel word either to the object or the action accordingly. Importantly, uncovering the specific content of a speaker's communication is dependent on the learner's understanding of the speaker's epistemic states, that is, what the learner thinks the speaker is referring to. For example, if mom points to a box and labels the toy inside a "blicket," the child's learning of the word "blicket" is contingent on his/her knowledge of what mom thinks is inside the box. It is possible that mom thinks toy A is inside, whilst the child knows that dad has replaced A with another novel toy when mom is away; and therefore, a correct inference that "blicket" refers to toy A, not the toy currently inside the box, is critically dependent on the child's understanding of mom's false belief. Studies suggest that 17-month-old infants and preschoolers indeed consider a speaker's belief state, true or false, in such context (Carpenter, Call, & Tomasello, 2002; Happé & Loth, 2002; Southgate, Chevallier, & Csibra, 2010).

Adding another twist to the studies measuring children's learning of reference of a novel label based on false belief reasoning, we ask if children can predict another individual (a learner) to consider a speaker's belief states in uncovering the intended reference, especially when the learner falsely thinks that the speaker has a false belief, while both the learner and the speaker have a true belief. For example, a girl puts one novel toy into box A while a boy is watching; when the girl is away, the boy relocates the toy into box B, puts another novel toy into box A, and then he goes outside. Immediately, the girl comes back and discovers the new toy in box A. Therefore, both characters (and the child) have a true belief about the toy's identity in box A, while the boy (wrongly) believes that the girl thinks the old toy is still in box A. In the end, the boy comes back, and the girl names the toy in box A her “blicket.” The boy wants to play with the girl’s blicket, and he searches for it. Given the girl’s true belief, children could infer that the ‘blicket’ is indeed the toy in box A. However, the boy will believe that the girl is referring to the old toy now hidden in box B. Therefore, based on an understanding of the boy’s belief about the girl’s belief (second-order), children should predict that the boy will search in box B. However, any first-order belief attribution (true belief) will lead to a different prediction, namely that the boy will search in box A. We showed children both outcomes – the boy reaches box A or the boy reaches box B – simultaneously, and measured their preferential looking while listening to the story narration.

Can younger children understand second-order false belief when spontaneous looking behavior is measured? Answers to this question will enrich our understanding of the nature of early ToM competence, and help us address the debate of whether early success in false belief tasks can be explained by non-ToM behavioral regularities.

Methods

Subjects were 29 children around 39 months old (16 girls; mean age = 39.2 months, s.d. = 5.27 months). Another ten children were recruited but excluded: two because they preferred the mismatching picture in the critical setup trial (setup 7, see below) in which the key difference between true-belief and false-belief conditions were revealed; one because the child failed to look at the matching picture in 75% of the setup trials (6 out of 8); two because they preferred the same side of the storybook in setup trials (looked at the same side in 7 out of 8 setup trials, or spent over 70% of overall looking time at the same side); two because they were overly active, and three because they were distracted during the experiment. Children were randomly assigned to the true-belief (n = 14, mean age = 39.0 months) and false-belief (n=15, mean age = 39.3 months) conditions.

Materials

The verbal preferential looking task followed the design in Scott, et al (2012, experiment 1) in which children were presented with a large picture book with two pictures on each page while listening to a story, and one of the two pictures on each page matched the story narrations. Children's preferential looking time at either picture was measured.

Children either participated in the lab or in a separate room at local schools. When tested in the lab, the child was seated at a table by himself/herself, with his/her caregiver seated right behind. In front of the child was a picture book raised at a 30º angle, about 30 - 50 cm distance from the child. An experimenter sat behind the storybook, facing the child. A camera in the middle behind the storybook (or to the left/right of the storybook for some of the subjects) captured the child's eyes, and a camera in the back above the head of the child captured the stimuli and the experimenter's face. The setting was the same in school except that no caregiver was present.

The picture book includes 11 pairs of printed color photos (20.6 cm by 27 cm). Each pair of color photos were put in two transparent plastic sheets (30.5 cm by 24.6 cm) attached in the center, forming one double-page. Each of the two transparent plastic sheets of a double-page was attached to the top of a black binder (30.5 cm by 25.4 cm), with three binder rings. The two black binders were connected at the center, and they were attached to a laptop stand (42 cm by 31 cm) that provided support to the double-pages picture book. The gap between the two color photos on each double-page was 6 cm (at the distance of 30 cm, visual angle is 12º; see Fig.7).

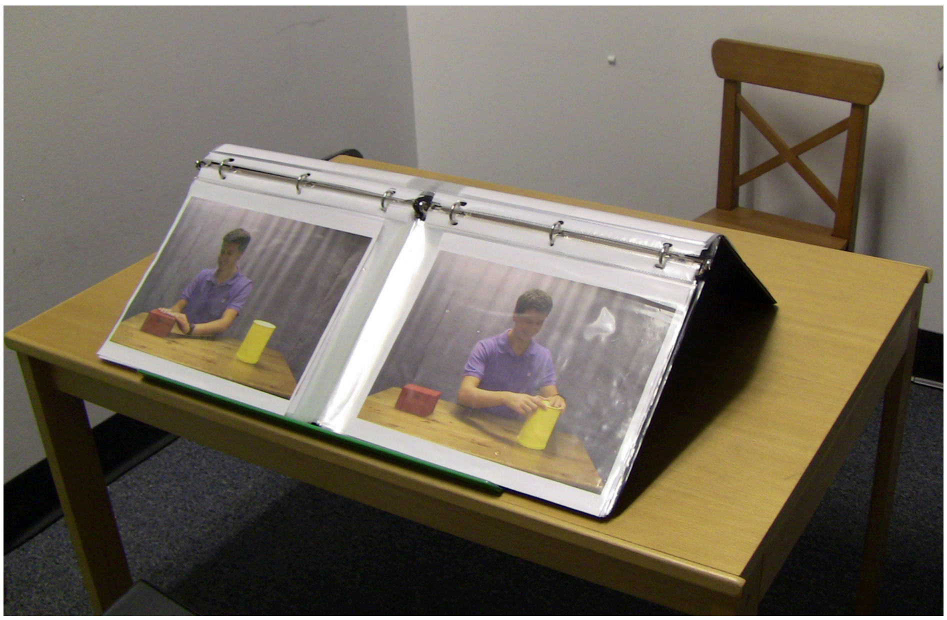


Figure 7. Picture book used in the experiment.

Design & Procedure

In the first two of the eleven double-pages, one side was a color picture and the other side displayed white paper of the same size as the color photo. In the subsequent nine double-pages, one side demonstrated a picture consistent with the story line, while the other side displayed an image of one of the characters doing something irrelevant (see Fig.8).

The procedure followed that from Scott et al. (2012) with the following exceptions: (1) the 11 double-pages were organized into two introduction trials, eight setup trials, and one test trial. In the introduction trials, the images presented the main character, Emily (introduction 1) and her friend Tim (introduction 2). In the setup trials, children were told that Emily found a novel toy (setup 1), and she put the toy into a red box (setup 2; or a yellow bucket, counterbalanced across subjects in each condition) while Tim was watching. Emily went to take a nap (setup 3), and while she was sleeping, Tim removed the toy and put it into the other container (setup 4). Then Tim found another novel toy and he put it into the now-empty container (setup 5). At this point, the false-belief and true-belief conditions differed.

In the false-belief condition, Tim went outside to play (setup 6), and while he was outside playing, Emily woke up and found the new toy in the original container where she put the first toy (setup 7). This generates a second-order false-belief because Tim still thinks that Emily thinks the original toy is in the box. Then Tim came back, and Emily told Tim the toy in the box was her “dax/blicket” (setup 8).

In the true-belief condition, after Tim puts the new toy in the newly empty container, he went to read his book (setup 6). Then Emily woke up, and while Tim was watching, Emily found the new toy in the box (setup 7). Emily told Tim the toy in the box was her “dax/blicket” (setup 8).

In both introduction and setup trials, the side of the matching picture was counterbalanced across trials, across children in each belief condition, and across other controlled variables (including which toy was used at the beginning, and in which container the toy was initially put). Overall, in each trial, half of the children saw the matching picture on the left, and the other half saw it on the right.

The story ended with a single test trial, and children were told, “Tim wants to play with Emily’s dax/blicket. He searches for it.” Then the double-page was flipped and children were told, “See? He is looking for it.” The left image of the double-page showed Tim searching in the box and the right image showed him searching in the bucket (see Fig.8). Half of the time the left image was matching the narration (given the characters’ belief) and half of the time the right image was matching the narration.

Common pictures for both true-and false-belief conditions: introduction 1-2, and setup 1-5

‘This is a story about a girl named Emily.’

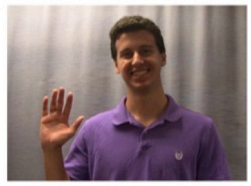
Introduction 1

‘There is Emily! Do you see Emily?’

(Pause 3 s)

‘Emily has a friend named Tim.’

Introduction 2



‘There is Tim! Do you see Tim?’

(Pause 3 s)

‘Emily finds a toy.’

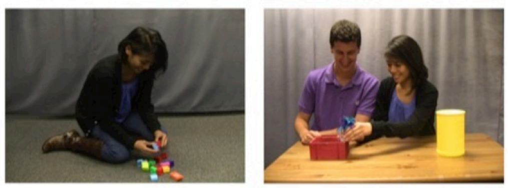
Setup 1

‘See? Emily finds a toy.’

(Pause 3 s)

‘Emily puts the toy into a box.’

Setup 2

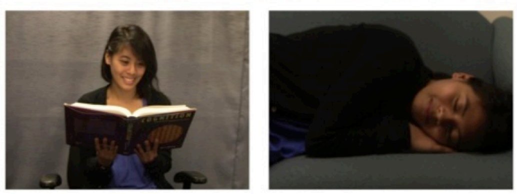


‘See? Emily is putting the toy into a box.’

(Pause 3 s)

‘Then, Emily is tired. She goes upstairs to take a nap.’

Setup 3

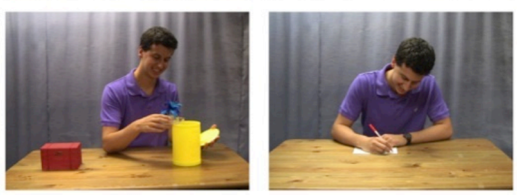


‘See? There she is taking a nap.’

(Pause 3 s)

‘While Emily is sleeping, Tim puts the toy into the bucket.’

Setup 4

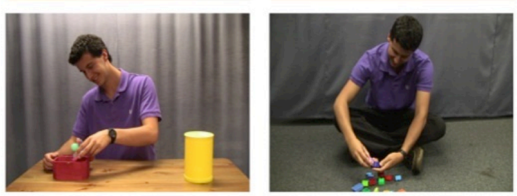


‘See? Tim is putting the toy into the bucket.’

(Pause 3 s)

‘Then, Tim finds a new toy. He puts the new toy into the box’

Setup 5



‘See? He is putting a new toy into the box.’

(Pause 3 s)

Figure 8. Example of the picture book, showing a false-belief condition in which the boy falsely believed that the girl thought the feather object was still in the red box, while the girl learned that the ball object was actually hidden in the red box; and the different pages (Setup 6 – 7) in the true-belief condition in which the boy saw that the girl discovered the ball object in the red box. False-belief condition and true-belief condition only differ in Setup 6 and Setup 7.

Coding

For each setup trial and the test trial, observers coded where the child looked (left, right, or neutral) from silent videos recorded by the camera capturing the child's eyes, frame-by-frame. Specifically, each video was exported as frame-by-frame static images, with 28 frames per second; and for each frame, an observer coded whether the child was looking at the left-side picture, the right-side picture, or neither. For the eight setup trials, the timing started immediately when the double-page was visible to the child, prolonged to a 7s-window (196 frames of image for each trial). It ended during the 3s-pause after the narration was repeated, prior to the narration of the next page. For the test trial, the timing was prolonged to a 8s-window (224 frames of image). All children were coded by one observer. Reliability was assessed by having another observer blind to the experimental conditions recode 95% of the data. The average agreement on the direction of children's gaze was over 97% between the two observers, that is, over 97% of the frames were coded the same by the two observers. Frames that the two observers coded differently were resolved by discussion.

Children's preferential looking time at the left (or the right) picture in each trial was then calculated by dividing the number of frames in which the child was looking at the left picture (or the right) by 28. For example, a child who was looking at the left picture in 56 out of the 196 frames in one setup trial means the child spent 2 s looking at the left picture in that trial.

Results

Across the eight setup trials, children looked reliably longer at the matching picture than at the mismatching picture (mean time looking at the matching picture = 3.92 s, at the mismatching picture = 1.93 s; t (28) = 13.79, $p < .001$), and this pattern held for both true-belief (matching mean = 3.76 s, mismatching mean = 2.20 s; t (13) = 10.63, $p < .001$) and false belief conditions (matching mean = 4.06 s, mismatching mean = 1.68 s; t (14) = 12.15, $p < .001$). All children showed the same pattern.

Children's looking time over the 8 s window in test trial was submitted to an ANOVA test with condition as a between-subject factor and picture (target picture showing Tim searched the container with the new toy, and the other picture showing Tim searched the other container with the old toy) as within-subject factor. Only the interaction between condition and picture was significant, $F(1, 27) = 4.52$, $p < .05$, $\eta^2 = 0.144$ (see Fig.9, top panel), suggesting that children had different predictions of which toy Tim would search for in the true- and false-belief conditions.

Because of the large variances in children preferential looking in each condition, to better investigate the interaction, we calculated differential looking scores (DLS) by subtracting children's looking time towards the target picture from that to the other picture, and divided by total looking time to both pictures. Therefore, DLS varied from −1 to +1, with positive number indicating a preference of the target picture. The differences between true-belief and false-belief conditions were significant ($F(1, 27) = 6.87$, $p < .05$, $\eta^2 = 0.203$). A planned t-test showed that children in the true-belief condition preferred looking at the target picture (mean DLS = 0.19; $t(13) = 1.85$, $p < .05$, one-tailed), whereas children in the false-belief condition showed the opposite preference of the other picture (mean DLS = −0.16, $t(14) = -1.85$, $p < .05$, one-tailed) (see Fig.9, bottom panel).

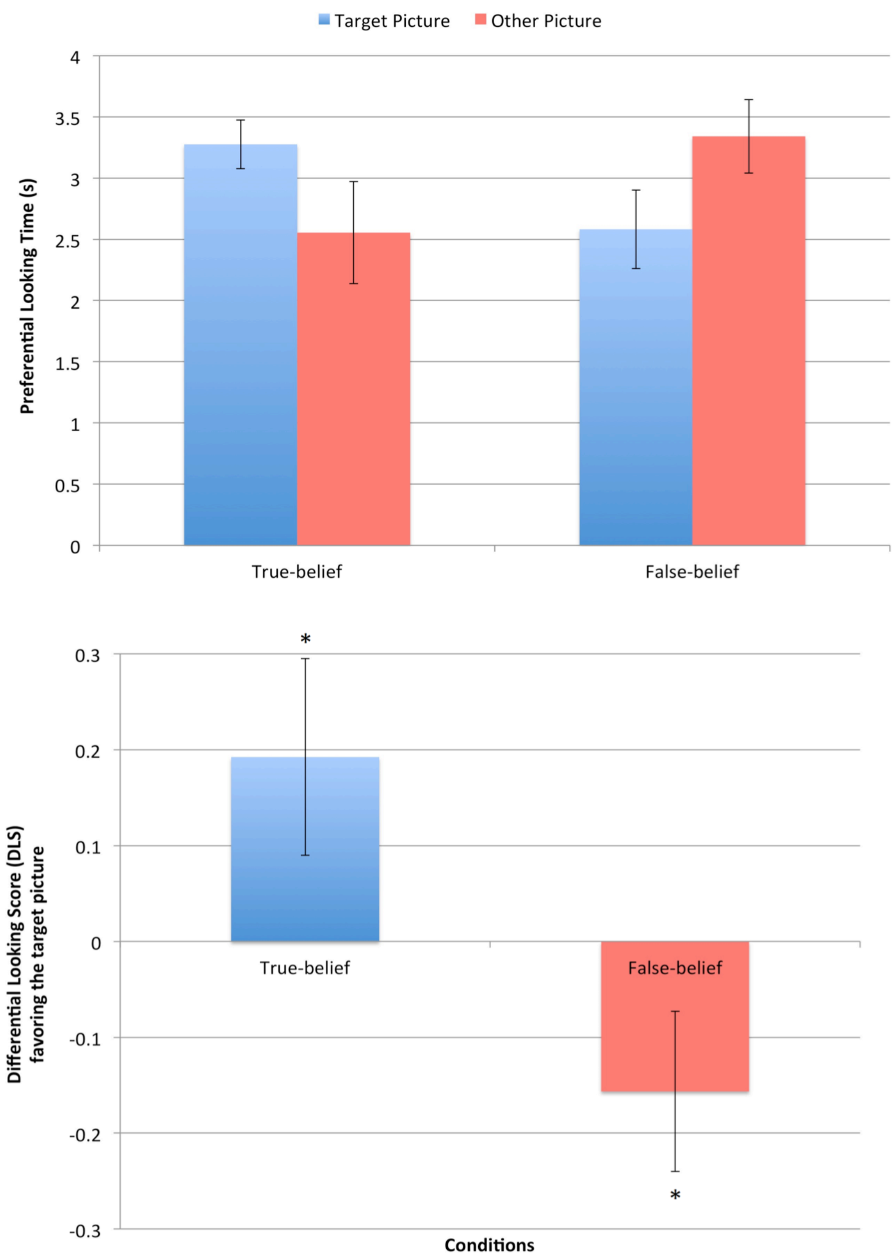


Figure 9. Children's preferential looking time (top panel) and differential looking scores (DLS, bottom panel) over the 8-s window in test trial, for true- and false-belief conditions, respectively.

Discussion

Do children understand second-order false belief? Measuring children's preferential looking time in a second-order belief reasoning task, we established the first evidence that 39-month-old children are sensitive to second-order false belief spontaneously.

We introduced a communication scenario and found that children correctly inferred another person's inference of a novel label given that person's belief about the speaker's epistemic status, even when that belief was false. Critically, to search for Emily's blicket, Tim had to uncover the referent of Emily's communication given his understanding of her epistemic states. In other words, the content of Tim's belief was what Emily had in mind when she labeled the toy inside the container, "her blicket." Studies have shown that infants as young as 17 months can pragmatically disambiguate a speaker's intended referent by considering the speaker's first-order belief even when it was false (Southgate et al., 2010). Our results extend previous findings, and suggest that 39-month-old children can pragmatically represent another person's inference of a speaker's communication giving that person's second-order belief of the speaker's epistemic states, even when the second-order belief was false.

The design of the false-belief condition closely matched the verbal second-order false belief task designed by Perner & Wimmer (1985): in each story, both characters' first-order beliefs of the situation are true, while one character's belief of the other character's belief is false. As discussed by Perner (p 443, Perner & Wimmer, 1985), "a correct answer results only when the second-order belief structure is taken into account. Any of the other strategies which are based on lower level structures or reality lead to wrong answers." Reflecting the claims in our study, had the children in the false-belief conditions followed any first-order belief or the reality, they would have predicted that the boy would correctly interpret the reference of the novel label, and to look for it in the target location, giving rise to the same preferences in true-belief and false-belief conditions. Specifically, in both true-and false-belief conditions, Emily discovers the identity of the toy in the target location, and Tim has a correct belief about which toy is hidden in the target location. Therefore, based on Emily's first-order belief, children could correctly infer that the toy hidden the in target location is her blicket; and based on Tim's first-order belief, they would expect him to search in the target location in both true-and false-belief conditions. Similarly, reasoning about the reality will give rise to the same predictions, because the reality is the same in both true-and false-belief conditions. Therefore, our findings – namely, that 39-month-olds expected the boy to find the "blicket" in the true-belief condition but to err in the false-belief condition – strongly imply an understanding of second-order false belief.

Our findings also cast doubt on the notion that language — in particular, syntactical knowledge of complementation — is required for higher-order belief representations (Hollebrandse, van Hout, Hendriks, 2012). de Villiers and her colleagues proposed that knowledge of sentential complement and embedded propositions ("John thinks that Mary wants him to pretend …") serves as a necessary conceptual framework for children to represent minds (Astington & Jenkins, 1999; de Villiers & de Villiers, 2000; de Villiers, 2005, 2007). This view has been seriously challenged by the new findings of preverbal infants' competence in first-order belief reasoning. However, it is still possible that language is required for higher order belief representations. Echoing the language-first hypothesis, previous studies using verbal or non-verbal stimuli paired with explicit questions suggested that the understanding of second-order false belief did not develop until 5, by which point children have demonstrated the knowledge of sentential complement (de Villiers & Pyers, 2002). However, it seems unlikely that our 39-month-olds' success in second-order false belief reasoning were supported by any syntactic knowledge, given that this ability develops around the end of the fourth years (de Villiers & Pyers, 2002). Instead, their success seems more likely supported by the early competence of ToM, independent from language (Low, 2010). Studies showing spontaneous second-order false belief reasoning with preverbal infants will provide more decisive evidence to this query.

Alternative behavioral explanations?

Perner and his colleagues summarized three non-theory-of-mind mechanisms to explain the early ToM competence (Perner & Ruffman, 2005; Ruffman & Perner, 2005), including (1) a three-way association among agent-object-location, (2) an ignorance-therefore-error strategy (see Ruffman, 1996; Southgate, Senju, & Csibra, 2007; also see a 'irony bias' in Friedman & Leslie, 2004b), and (3) a behavioral rule that people always go to where they last saw or put something to search for it (also see, Perner, 2010; Perner & Roessler, 2012).

Research has provided direct evidence that the first two rules are not adequate in explaining spontaneous ToM competence (e.g., Scott et al., 2010). For example, using looking time measurement, Scott et al. (2010) investigated whether 18-month-old infants understand an actor's behavior in relation to objects' hidden properties. An experimenter introduced a cup that would rattle when shook, and 18-month-old infants saw that another cup with distinct appearance had the same property, while another cup that looked identical to the first one did not rattle when shook. Then an actor appeared, and the experimenter shook the first cup as before, and it made special noise. Then the experimenter encouraged the agent to repeat the rattling effect with one of the other two cups. Infants expected the agent to pick the cup that looked identical to the one demonstrated by the experimenter, assuming that objects that look the same may share the same hidden properties. In this task, there was no association formed between the agent and any of the objects. Therefore, a three-way association is not adequate for explaining infants' success in false belief reasoning.

The authors introduced another condition in which both cups the agent could choose from looked identical to the one used by the experimenter, while only one of them would rattle when shook. Again, infants knew which cup had the hidden property, yet they predicted the actor who was absent during the demonstration would pick randomly between these two similar cups. That is, infants did not expect the actor who was ignorant to make mistakes (choosing the wrong cup that would not rattle when shook), but to behave randomly. Therefore, previous research suggests that neither an association nor an ignorance-therefore-error strategy can explain infants' success in false-belief tasks.

Children's spontaneous reasoning of second-order belief provides further evidence that spontaneous ToM capacity cannot be explained by the behavioral rule either. Because both actors in our second-order false belief task had a true belief about which toy was in the target location, the rule that "a person will search for an object where she last saw it" would predict that the boy would search in the target location.

However, children may have attributed a second-order behavioral rule to the boy. In other words, children not only learn a set of behavioral rules from experience to predict people's behavior, they also expect others to apply similar behavioral rules to interpret a third party's action. In our false-belief condition, for example, the girl put toy A in the target location at the beginning. The boy replaced toy A with a new toy, B, and put toy A in the other location. The girl discovered the new toy B in the target location, therefore, when she labeled the object in the target location her "blicket," she meant to refer to toy B. If children expected the boy to use a behavioral rule that the girl would look for something where she last saw it, and therefore, when labeling an object at a particular location, she was naming the object she last saw in that container, then: (1) the boy did not know that the girl last saw toy B in the target location, but he had evidence that the girl last left toy A at that location; and therefore (2) when the girl referred to the object in the target location her "blicket", the boy should have interpreted the "blicket" as toy A; (3) the boy last saw/put toy A in the other location; thus, (4) given a similar behavioral rule "the boy would look for something (toy A) where he last saw/put it", children expected the boy to search in the other location in the false-belief condition.

The problem with the second-order behavioral rule is that (1) is precisely the boy's second-order false belief: (1) is true if and only if the boy's knowledge that the girl last saw toy A in the target location is false. In other words, implicitly implied in this account is an assumption that children keep track of the information the boy has access to, and expect the boy to behave consistently with that information even when it is outdated. Thus, extending the strategy to a second-order level is not a viable alternative to second-order false belief reasoning.

In sum, together with studies revealing infants' first order false belief reasoning, an early understanding of second order false belief suggest that none of the behavioral rules proposed by Perner and his colleagues (Perner & Ruffman, 2005; Perner, 2010; Ruffman & Perner, 2005) or extension to the rules can explain the early ToM capacity.

Conclusion

We found that 39-month-olds can spontaneously reason about second-order false beliefs. The early competence cannot be reduced to any lower level behavioral rules; instead, it reflects a mentalistic understanding of mind.

V. General Discussion

For many years, it has been widely assumed that false belief reasoning first develops after four years of age (Perner, 1991; Gopnik & Wellman, 1994; Wellman, Cross, Watson, 2001). However, results reported in this dissertation support more recent findings that younger children understand true- and false-belief in the second year of life when measured by various spontaneous behaviors (Onishi, Baillargeon, 2005; Southgate, et al. 2007; Buttelmann, et al., 2009; Scott, et al., 2010; Kovács, et al. 2010). The early competence in infants suggests that later failure in children is likely caused by performance limitations. Nevertheless, the wealth of findings from false belief studies in the past does not simply disappear – it leads to two broad questions: (1) What is the nature of the earliest competence? (2) How does this early competence relate to the later developed abilities? Projects in the dissertation shed light on these questions.

First, studies in chapter 2 show that, like verbal responses, children's spontaneous belief reasoning is subject to task processing demands. When task processing demands were reduced by undermining the child's confidence in his/her own knowledge, he/she was more likely to anticipate an actor's behavior following from a false belief; however, when the tasks' inhibitory demands increased, his/her anticipatory eye gaze was balanced between the true-belief and false-belief positions. This pattern of responses was not caused by a lack of competence in children, because adult subjects showed exactly the same pattern of eye gaze. Our findings provide the first direct evidence for the proposal that the theory-of-mind mechanism actually computes two candidates, a default 'true belief’ and a context-contingent ‘false belief’. The eye gaze pattern we observed shows that both contents are equally attractive and thus about equally uncertain.

Second, navigating the social world typically requires us to track multiple agents and their mental states at the same time. Agents’ mental states are non-stationary, and therefore, tracking multiple agents’ distinct mental states requires a frequently updated working memory, rather than a passive short-term store. It is largely unknown what the working memory capacity is for us to bind mental states to the correct agents. In chapter 3, we established the first direct evidence that three- to four-year-olds can track two people’s distinct false beliefs verbally, no later than their ability to understand a single false belief. When they erred in action prediction, three-year-olds failed by attributing their own knowledge to the agents, while they rarely made any binding mistakes by attributing one agent’s belief to another. Four-year-olds were neither making systematic binding mistakes nor attributing their own true belief to the agents. Furthermore, two-year-olds can do so nonverbally when their anticipatory eye gaze was measured. Together, studies in chapter 3 show that working memory is not a critical limitation in children’s belief-desire reasoning (c.f., Davis & Pratt, 1995; Gordon and Olson, 1998); the capacity to solve the binding problem in ToM extends beyond a single agent.

Finally, in chapter 4, we investigated whether spontaneous ToM can readily represent second-order false belief. We found that 39-month-old children's preferential looking revealed an early sensitivity to an actor’s second-order false belief. This ability cannot be explained by first-order belief attribution or non-theory-of-mind processes, suggesting that spontaneous belief-desire reasoning is mentalistic, rather than behavioral, in nature.

The Competence: is the early understanding of mind real theory of mind?

Investigations on the nature of early ToM is critical in reconciling infants' success and children's failures in false belief reasoning. Some researchers suggest that infants' early success indicates a genuine understanding of mind, and it reflects the same competence as later developed abilities (Baillargeon et al., 2010; Carruthers, 2013). We name this the *continuity view* of infants' data. Nevertheless, other researchers hypothesize that the early reasoning that is present is not ToM *per se*, but some knowledge of behavioral regularities (Perner & Ruffman, 2005; Perner & Roessler, 2012; Ruffman & Perner, 2005), or inflexible "belief-like" registration with "signature limitations" (Apperly & Butterfill, 2009). The later two differ in their explanations of infants' success, yet both agree that only the later-developed verbalized ability reflects a mentalistic and conceptual understanding of mind, and it emerges around the fourth birthday. We name this a *discontinuity view* of infants' and young children's success. I will explain the discontinuity view first, and by leveling arguments against them, I will then defend the continuity view.

Behavioral regularities?

As been stated earlier, Perner and his colleagues have proposed three behavioral principles to explain babies' success in false belief reasoning (Perner & Ruffman, 2005; Ruffman & Perner, 2005): (1). an agent-object-location three-way association; (2). an ignorance-therefore-error strategy that when a person has ignorance, babies do not expect the person to do something random but something wrong, in other words, he/she will fail to fulfill his/her goal; (3). statistical regularities learned from observation of people's behavior, that a person will go to the place where he/she last saw/put an object to retrieve it (for a more detailed account of statistical learning, see Ruffman Taumoepeau, & Perkins, 2012). Studies presented in this dissertation contribute to the on-going investigations of these behavioral accounts.

In terms of the association account, increasing evidence suggests that associations are not sufficient to explain infants' social cognition (e.g. Luo, 2010; Scott, et al., 2010; Song & Baillargeon, 2008). For example, after seeing an agent choose object A over object B repeatedly in familiarization trials, infants expect the agent to choose object A in the test trial even when the spatial positions of the two objects are swamped (e.g. Woodward, 1998). However, if the agent selected object A repeatedly in the familiarization trials without knowing the existence of object B (even though the infant could see both objects), infants did not expect her to select object A again when provided with both A and B in the test trial (Luo, 2010; Luo & Johnson, 2009). Therefore, infants' expectations were based on a calculation of the agent's preference given the agent's knowledge of the candidates, rather than a simple association between the agent and an object.

With respect to the ignorance-therefore-error strategy, as reviewed in chapter 1, Southgate et al (2007) showed toddlers that an actor falsely believed a toy was in location A, while it had been removed, completely out of the scene. Given the ignorance interpretation, children should have expected the actor to search randomly between two empty locations in the scene, because both locations were wrong. However, 25-month-olds consistently expected her to search in location A, congruent with her false belief.

Studies in chapter 2 and 3 adapted a similar design, especially the low-demand false belief task in chapter 2 and the anticipatory looking task in chapter 3. Our findings that children expected ignorant actors to behave consistently with their belief provided further evidence against the ignorance interpretation (also see Friedman and Petrashek, 2009; Scott et al., 2010).

Finally, our findings in chapter 4 that 39-month-old children were able to reason about second-order false belief cannot be explained by behavioral rules. In the study, a character Tim removes another character, Emily's toy from location A to B when Emily is away, and he puts a new toy in location A. However, unbeknownst to Tim, Emily discovers the new toy in location A. In the end, Emily tells Tim the toy in location A is her blicket. Tim wants to play with Emily's blicket, but he falsely believes that Emily does not know about the switch of toys, and therefore, Emily must be referring to the toy currently in location B. This result cannot be explained by behavioral rules, because without understanding Tim's second-order false belief, children can still infer that Emily's blicket is the toy currently hidden in location A based on Emily's true belief; however, given the behavioral rule that a person will go to the place where he/she last saw an object to search for it, the behavioral rule would predict that Tim will search for Emily's blicket in location A, where he last put it. Our findings that 39-month-olds correctly predicted Tim to search location B for the blicket speak against the behavioral rule.

Furthermore, theoretically, some fundamental flaws with the behavioral account make it less likely to be true. First of all, behavioral rules are ad-hoc by their nature. The number of rules posited to explain babies' success on false-belief tasks is dependent on the number of scenarios that have been used in experiments that yield babies' success. This becomes more and more problematic when more studies using non-searching paradigms such as the one designed by Scott and her colleagues (2010) come out. Secondly, there are virtually no constrains on what rules can be learned from sensory/perceptual experiences. Why are these rules learned but not others? How do infants acquire these rules? Without explanations to these questions, the behavioral account is hardly an adequate alternative.

A similar criticism has been advanced by others (Carruthers, 2013; Surian & Geraci, 2012). For example, Carruthers summarized that "this account (behavioral rule account) makes no predictions of its own, however. Indeed, it is entirely parasitic on positive results provided by the infant-mindreading hypothesis to generate the proposed set of rules." Surian & Geraci (2012) pointed out that, "their viability would depend on the plausibility of a number of additional assumptions required to specify the relevant environmental input, learning mechanisms, and experiences that allow young infants to acquire all these rules." Without any learning mechanism to accompany the behavioral rules, it is hardly adequate in explaining infants' early competence in ToM.

"Belief-like" states?

Apperly and colleagues (e.g. Apperly & Butterfill, 2009; Apperly, 2010) suggested a dual-system account to explain the tension between efficiency and flexibility in mind-reading, and the differences between an early ability present in infancy and late-developed flexible ability in adults: a low-level modular system develops early in infancy that is efficient at the cost of signature limitations, and a non-modular system develops later in childhood and adulthood that is flexible at the cost of efficiency.

Although the early modular system is critical for ToM, it is conceptually different from the ability underlying children's ToM measured in verbal tasks (see also Rakoczy, 2012). Specifically, infants possess a "belief-like" registration system that functions like a belief, but is not a belief. The registration is similar to a memory trace - when an actor sees X, such as an object, infants represent the actor's registration of X; if any changes happen to X when the actor is away, the representation of his/her registration remains the same and it "fails to be correct"; when the actor comes back, his/her registration is used to predict his/her action, because "registrations resemble beliefs… one can understand registration as an enabling condition for action, so that registering an object and location enables one to act on it later…" (p. 962). However, the registration system is sharply limited and inflexible, and it suffers "signature limitations on the nature and complexity of theory-of-mind processing that can be achieved" (p. 963). On the contrary, children's ToM reasoning, supported by the later development of language and executive function, is a conceptual system that is flexible and cognitively demanding.

According to Apperly & Butterfill (2009), the signature limitations of the non-conceptual system may include: "restrictions on the kind of input it can take and the kinds of belief-like states it can ascribe." (e.g. Low & Watts, 2013); it cannot "use all cognitively available facts to ascribe any belief that the subject can themselves entertain"; and "any beliefs that could be represented in the late-developing system could never be mapped back into the early-developing system" (p. 964).

Apperly's proposal seems appealing, especially because it highlighted two conflicting requirements that any belief-desire reasoning theory should explain, that is, efficiency and flexibility. We agree with Apperly that perhaps more than one system is involved in theory of mind, yet we disagree on the nature of the early-present modular system (as discussed in the competence).

First of all, Apperly and Butterfill (2009) suggested that unlike the conceptual understanding of mind that is flexible, the efficient modular system suffers from signature limitations. However, given its modular nature, the early belief-desire reasoning system is likely to be limited in certain ways, but these limitations alone do not deem the modular system non-conceptual. In particular, Fodor hypothesized that certain systems operate automatically and mandatorily whenever relevant information in their domain appears; these systems have limited interactions with other systems or with conscious knowledge, that is, operations of these systems are encapsulated. The efficiency of these systems is guaranteed by the properties above, so is their limitations. The assembly of these properties is the hallmark of modularity (Fodor, 1983). Fodor focused on several low-level systems, such as visual processing and language processing, which he named perceptual analyzer. Later, various scholars have extended the idea of modularity to conceptual domains (Leslie, 1986; Sperber, 2005; Carey, 2011), such as the core knowledge system of object and agency (Carey, 2009). Leslie (1994a, 1994b) applied the modular theory to explain ToM reasoning (discussed below), which can readily explain the efficiency and flexibility of ToM processing.

Secondly, it is not entirely clear what might be "beliefs that could be represented in the late-developing system could never be mapped back into the early-developing system" (Apperly & Butterfill, 2009, p. 964). Nevertheless, given that even five-year-olds have difficulty in the verbal second-order false belief tasks, we may hypothesize that ascribing second-order false belief is one of the complicated beliefs that "could never be mapped back." Our findings that even 2.5- to 3-year-old children were sensitive to second-order false belief as revealed in their spontaneously preferential looking contradicted this speculation.

Rakoczy (2012) also explained that infants' subpersonal level understanding of mind is conceptually different from the personal level understanding in older children and adults, because the 'real' concept is not just realized by the module, but also via the connections between the module and other functions in the 'right' way: "in normal adults some ToMM representation might well be the core realizers of beliefs about beliefs because they are functionally related to other subdoxastic states in the right ways (playing their role in theoretical and practical reasoning and the rational control of action, etc.) so that the whole system is a **total realizer** of beliefs about beliefs. But that simply does not mean that these representations already realize beliefs about beliefs in infants. Rather, the infants might have the subdoxastic parts, but not the right kind of functional connections yet, that constitute the very competence in question." (P65, Rakoczy, 2012).

We do not deny the enrichment of ToM reasoning or building "the right kind of functional connections" along development. In fact, a lot needs to be learned from experiences, such as links between various mental state concepts (e.g. German and Leslie, 2001). However, it seems arbitrary to grant infants or children ToM concepts if and only if all enrichment or functional connections have been completely established. Whether the origin of ToM reasoning is conceptual and mentalistic, and whether infants' understanding of belief is real are empirical questions, rather than graduation with some selected requirements. Therefore, instead of listing the functional connections that have to be established for a concept to be real, my approach is to investigate the properties of infants' ToM reasoning, namely, the operative principles of the original state that can be enriched and elaborated into the full state of ToM in adults. What inferences, then, can the early ToM competence afford?

The competence

First of all, I would like to argue that preverbal representation can be conceptual, and those preverbal concepts can be knowledge based (e.g. representational theory of mind, as advocated by Perner, 1991, 1995), or embodied in the operative principles of cognitive mechanisms. Implicit in the discontinuity view is the suggestion that preverbal representation is sensory and perceptual in nature (Piaget, 1954; Quine, 1960). However, as thoroughly elaborated by Carey (2011), perceptual reasoning cannot explain babies' capacity in various cognitive domains. For example, babies as young as three months understand object permanence (e.g. Baillargeon & DeVos, 1991; Baillargeon, Wasserman, & Spelke, 1985), an abstract concept that the same object keeps existing even when it is not perceived. Moreover, infants in their first year of life represent numerical information of a set of objects (Starkey & Cooper, 1980; Strauss & Curtis, 1981; Starkey, Spelke, & Gelman, 1990; Wynn, 1992, 1995), and they represent number cross sensory modalities (Jordan & Brannon, 2006; Jordan, Suanda, & Brannon, 2008; Kobayashi, Hiraki, & Hasegawa, 2005; Starkey, Spelke, & Gelman, 1990). Again, numerical information is not directly presented in sensory/perceptual inputs; yet infants not only represent numerical entities, they also make inferences of these representations across modals. Therefore, Carey suggested that some preverbal representations are conceptual in nature: (1) they cannot be reduced to sensory/perceptual primitives or behavioral regularities. (2) they can play rich inferential role in thoughts. Critically, sensory/perceptual representation has limited inferential power. For example, not much can be inferred from being "red"; however, rich inferences can be made from being an "agent" (Carey, 2011, p157-213); these rich inferences may be not grounded by infants' knowledge, but are carried on implicitly in the operative principles of corresponding cognitive mechanisms.

Is preverbal representation in ToM conceptual on Carey's metric? First of all, can early ToM representations be reduced to sensory/perceptual primitives? Clearly not. Discussions in the previous sections suggest that the early understanding mind is exactly one of the kind that cannot be explained by sensory/perceptual, or behavioral vocabularies. Secondly, can early ToM representations play a rich inferential role in thought? To answer this, we need to specify the inferential role played by the representations generated from a fully developed, conceptual ToM. ToM enables us to make predictions and generate explanations about an agent's mental states and his/her action if we assume (based on folk psychology and the principle of rationality) that a person's action is the most efficient way to satisfy his/her desire in light of his/her belief of the environment (Dennett, 1989). Three distinct inferences can be derived from the trio-relations between belief, desire, and action (Wang and Leslie, under review), including a prediction of action given belief and desire, an inference of desire given belief and action, and an inference of belief given action and desire. All three inferences are yoked under ToM and constrained by the principle of rationality, forming a hallmark of conceptual ToM reasoning.

There is evidence that infants in their second year of life are able to make all three inferences. First of all, most studies measured infants' action predictions given information of an agent's belief and desire, and they provide compelling evidence that preverbal infants and toddlers can predict an agent's action given information of her desire and false belief (Onishi & Baillargeon, 2005; Scott & Baillargeon, 2009; Scott et al., 2010; Scott et al., 2012; Song & Baillargeon, 2008; Southgate, Senju, & Csibra, 2007; Southgate & Vernetti, 2014; Surian, Caldi, & Sperber, 2007; Träuble, Marinovic, & Pauen, 2010; for a review, see Baillargeon, Scott, & He, 2010). The studies in this dissertation (chapter 2 and 3) add further evidence to this picture.

Second, some studies also show that infants can make inference of agent's desire given a representation of her belief and observed action (Buttelmann, et al., 2009; Buttelmann, Over, Carpenter, & Tomasello, 2014; Luo, 2011). For example, Buttelmann et al. (2009) showed that 18-month-old infants were able to infer different desires out of the same action, depending on an actor's belief. In their study, an actor hid a toy in container A; then an experimenter moved the toy and hid it in container B, all the while the first actor was either present (true-belief condition) or absent (false-belief condition). In the end, the actor tried to open container A where he had put the toy initially. Infants were encouraged to help the actor by opening either container A or B. In the false-belief condition, the actor still believed the toy was hidden in container A, and thus, his action indicated a desire to retrieve the toy. In the true belief condition, however, the actor knew the toy's whereabouts (in container B); therefore, his action can only be justified by a desire to open container A. Appropriate helping can only be achieved via inference of the actor's desire from his action and his belief. That was exactly what they found: in the false-belief condition, infants were more likely to open container B and retrieve the toy for the actor; while in the true-belief condition infants were more likely to open container A. Again, it suggests that the early competence of ToM can afford an inference of desire given belief and action.

Third, one study also shows that infants are able to make the third inference, namely, an inference of the content of belief given an understanding of an actor's desire and observed action (Southgate, et al., 2010). In the study, 17-month-olds saw that an actor hid two novel toys into two locations respectively. While the actor was away, an experimenter switched the toys' locations. In the end, the first actor returned, and she pointed to one location and labeled the toy inside a "sefo." Infants were encouraged to retrieve the "sefo" for her. To succeed, infants have to infer the content of the actor's belief (what she thinks the "sefo" is), based on her ostensive action (pointing to one location) and her desire to play with whichever toy the "sefo" is. Infants as young as 17 months were able to make this inference. Studies presented in chapter 4 further extend the finding above, and show that 39-month-olds not only make such inferences themselves, but also expect others to infer the content of belief given a third party's desire and action (second-level).

ToM can be thought of as computing the values of a function that relates belief, desire, and action, such that given any two of the arguments the third can be inferred. That is, three distinct inferential processes are yoked together under ToM. Interestingly, infants provide evidence for making inference from belief-and-desire to action prediction at the same age as they provide evidence of inferring desire from belief-and-action, and inferring belief from desire-and-action. These integrated inferences can hardly be explained by any non-conceptual behavioral generalizations. The evidence that early ToM reasoning not only draws together different aspects of belief-desire-action inferences, but also integrate mental state representations with rational control of one's own behavior strongly favors the continuity view of infants' and toddlers' ToM competence.

The performance: are different performance measures of theory of mind real measures?

Infants' conceptual reasoning of ToM is readily explained, in fact predicted, by the theory-of-mind mechanism (ToMM) advocated by Leslie and his colleagues (Leslie, 1987, 1994a; Leslie, German, & Polizzi, 2005; Roth & Leslie, 1998). As reviewed in the introduction to this thesis, ToMM is a domain-specific, probably innate, module that underlies belief-desire reasoning. As a module (Fodor, 1983; more particularly, a conceptual module, Carey, 2011), ToMM processes information about people's mental states automatically whenever relevant information in the environment is present (Cohen & German, 2009, 2010; Kovács et al., 2010). It directs infants' attention to information about agents and their mental properties, embodying an agent-centered representation of the world (Leslie, 1994a).

If ToMM is already present in infancy, why do three-year-olds fail in standard false belief tasks? One striking feature of the new wave of findings on early success in theory of mind is that they are derived from very different measures than those on which the old wave was based. Six types of experimental tasks have been developed, measuring infants' and children's action prediction given belief-desire reasoning. These tasks can be roughly categorized depending on whether or not language was used in (1) stimuli presentation and (2) elicitation of the response (Table 2). For example, on one extreme, the standard Sally-and-Anne task uses verbal stimuli and measures children's deliberate responses to a verbal question. On the other extreme, Onishi & Baillargeon (2005)'s study used non-verbal stimuli and measured infants' spontaneous looking time.

It has previously been suggested (Baillargeon, Scott, & He, 2010; Scott & Baillargeon, 2009) that the discrepancy between infants' early success and children's later failure can be largely explained by the mode of response elicitation used by the study. Put another way, it was hypothesized that the ToM tasks can be divided into two categories: those in which responses were elicited by direct questions (as in standard Sally-and-Anne task) and those in which responses were tapped indirectly (as in Onishi & Baillargeon, 2005). In both kinds of tasks, ToMM computes appropriate content for belief attribution. However, here is where (so the hypothesis goes) things differ from one mode of response elicitation to the next. Spontaneous tasks only draw on the belief representations provided by ToMM, without posing demands on other cognitive processes. On the other hand, when children's belief reasoning is elicited by direct questions, two additional processes beyond belief representation are required. In tasks that involve direct questioning, a child has to access his/her belief representation and select a response among the available choices (the first additional process: response selection), and inhibit the response that matches the child's own knowledge (the second additional process: response inhibition). Younger children with limited executive resources fail the standard tasks because they are not yet able to simultaneously execute a belief representation, response-selection and response-inhibition processes (for a similar and more general discussion of the difference between conceptual competence and successful performance in a task, see Gelman, 1993).

Our findings, however, suggest that different measures of belief might engage more than two performance processes, each with its own characteristics (processing demands or strength of connections with the ToMM (e.g. German and Cohen, 2012)). Particularly, ToM reasoning measured by anticipatory looking in a non-verbal task is subject to tasks' processing demand produced by the true-belief default (e.g. studies presented in chapter 2; Schneider, Bayliss, Becker, & Dux, 2012; Schneider, Lam, Bayliss, & Dux, 2012; c.f. Low & Watts, 2013). So is children's performance in standard verbal false-belief tasks. Yet it is still unclear whether other measures of belief reasoning are also subject to this demand. Table 2 summarized the six types of experimental tasks, and these measures seem to have different sensitivity to the processing demand produced by a true belief. Yet the effect of inhibitory demand in priming task, looking-time study, and verbal preferential looking task are untested.

Table 2. Summary of available data for different measures of false belief performance relative to selection demand made by the true-belief bias.

Measure	Predicted demand effect?	Evidence
Priming (Kovács et al., 2010)	No	Untested
Looking time	Perhaps	Not directly tested (but see Scott, et al., 2010; Yott & Poulin-Dubois, 2012)
Non-verbal anticipatory eye-gaze	Yes	e.g., Schneider, Lam, Bayliss, & Dux, 2012; the present studies
Verbal spontaneous anticipatory eye-gaze	Yes	e.g., Wang B. et al., 2012
Verbal preferential looking	Perhaps	Untested
Verbal standard Sally & Anne task	Yes	e.g., Carpenter, et al., 2002; Wang & Leslie, 2013

For example, in a priming task, Kovács and her colleagues (2010) showed that infants and adults automatically represented another agent's belief. The belief representation subsequently gave rise to a priming effect, such that subjects were more readily detecting a target as long as the target was expected by the agent (not necessarily expected by the subjects themselves). It has not been investigated whether the demand produced by the true-belief default will affect the priming effect. According to ToMM (Leslie, 1994a; Leslie, Friedman, & German, 2004), the true-belief content, whenever it is available, will be favored in a selection between the true-belief and false-belief contents. However, priming effect is initiated by an active false-belief representation. It happens as long as the false belief content is available, regardless of the selection process. Therefore, we hypothesize that the demand effect will not affect the priming effect.

Is false belief performance measured by looking time (as in Onishi & Baillargeon, 2005) subject to the demand produced by the true-belief default? So far, the evidence suggests that it is possible, yet untested. For example, as introduced in chapter 2, when appropriate comparisons could be made, Scott et al. (2010, p.383) found evidence that infants looked longer when viewing behavioral violations of false beliefs than when viewing behavioral violations of true beliefs. Furthermore, Yott & Poulin-Dubois (2012) used the same non-verbal false-belief task as in Onishi & Baillargeon (2005), and they found that infants' looking time when the actor's behavior violated her false belief was positively correlated with a measure of their executive function. That is, the higher the infant scored in the executive function task, the longer the infant looked at the unexpected outcome in the false-belief task. These finding suggest that even VOE looking time can measure the increased processing demands of tracking an actor's false belief. However, this hypothesis has not been tested directly.

Furthermore, Garnham (neé, Clements) and her colleagues did a series of experiments using verbal anticipatory looking tasks in which children's anticipatory looking was triggered by a verbal prompt embedded in a verbal story (Clements & Perner, 1994; Clements, Rustin, & McCallum, 2000; Garnham & Perner, 2001; Garnham & Ruffman, 2001; Ruffman, Garnham, Import & Connonlly, 2001; for recent development, see Low, 2010; He, Bolz, & Baillargeon, 2012; Wang B, Low, Jing, Qinghua, 2012). The verbal anticipatory looking task is similar to the standard Sally-and-Anne task, except that the direct question, "Where will Sally look for her marble?" is replaced by a self-addressed narration, “I wonder where she is going to look.” Children’s anticipation of the main protagonist’s reappearing after the verbal prompt is measured. Children older than three and a half robustly succeeded in anticipating what the protagonist would do given his/her false belief. Spontaneous anticipatory eye-gaze was also revealed in continuous eye-tracking while a story unfolded. Rubio-Fernádez (2013) showed pictures of the standard Sally-and-Anne story accompanied by narrations to adults and measured their continuous gaze pattern. While listening to, “When Sally comes back the next day, she goes to look for her doll”, before the utterance of “in the basket,” adults already shifted gaze to the empty location in consistent with Sally’s false belief. Interestingly, in the two scenarios above, the target object is typically hidden in the scene, result in a salient true belief. A recent study showed that when a low-demand task was used, that is, children’s belief about the target object’s whereabouts was uncertain, three-year-olds’ anticipatory looking following the verbal prompt was even better than that in the standard high-demand condition (Wang B, et al., 2012). Therefore, spontaneous anticipatory looking elicited by a verbal prompt seems to be sensitive to the increased demands of a certain true belief, even though children typically succeeded even in a high-demand condition.

Additionally, Scott et al. (2012) designed a novel task measuring spontaneous preferential looking embedded in a verbal story (as adapted in chapter IV; also see Barrett, Broesch, Scott, & He, et al., 2013). In their task, children sat in front of a large storybook, and each page of the book presented two scenarios, one consistent with the story narration while the other was not. Children heard a story similar to the Sally-and-Anne task (Emily and her apple), and in the end, the narration said, “Emily is looking for her apple." while children saw two pictures, one showing Emily searched for her apple in its original position (consistent with her false belief), while the other showed her reaching toward the apple's real position. Children as young as 2.5 years preferred looking at the picture consistent with the girl's belief states, without being influenced by their own true-belief. Again, however, the effect of inhibitory demand on preferential looking within verbal tasks is so far untested.

In sum, variations to false belief tasks and measures may tap on different performance systems that employ the ToM competence. The fact that the sensitivity to false belief tasks' inhibitory demands extends beyond the verbal performance system to the implicit system that controls anticipatory eye gaze suggests that the inhibitory demands might reflect a property of ToM competence (attribution of true belief derived from one's own knowledge as a priori). Systematic investigations of the sensitivity to this particular demand in other performance measures will be informative.

Final Remark

Experiments in this dissertation outlined the properties of spontaneous ToM reasoning. Studies in chapter 2 show that subjects' spontaneous anticipation measured by anticipatory eye gaze is sensitive to false belief tasks' processing demands. In chapter 3, two studies explored the capacity for children to track two agents distinct false beliefs both verbally and spontaneously, and they suggest that the working memory capacity to bind distinct false beliefs to the correct agent extends beyond one. Studies in chapter 4 show that children are able to represent the recursive nature of mind spontaneously, further suggesting that early competence in ToM is mentalistic, and it cannot be explained by behavioral regularities.

Built on these findings, we proposed two competence theories and a performance account - (1) early ToM ability is conceptual and it embodies three inferential processes; (2) true-belief default is a rational prior for belief-desire reasoning, and it may be a part of ToM competence. (3) The success-and-failure patterns in belief-desire reasoning may be caused by distinct performance systems with various characteristics and sensitivity to tasks' processing demands.

These findings have several implications for future research. I shall outline a few:

(1) We proposed that conceptual ToM reasoning entails three distinctive inferential processes. The three inferences yoked together form the hallmark of ToM competence. Most research on infants' belief-desire reasoning has focused on one of the inferences, namely, from desire and belief to action prediction. More empirical work exploring the other two inferential processes in infants is critical.

(2) Is the demand produced by true-belief default also present in other measures of infant's belief reasoning? Table 2 summarized some measures of belief, and these measures seem to have different sensitivity to the processing demand produced by a true belief. Yet the effect of inhibitory demand in priming task, looking-time study, and verbal preferential looking task are untested. Future work shall investigate both the nature of competence and the performance systems that employ that competence.

(3) Studies in chapter 3 proposed a new line of research, namely, the working memory capacity to solve the binding problem in ToM. We established first evidence of such capacity. More questions can be asked, for example, what produces the working

memory load in ToM reasoning? Does working memory system interact differently with different performance systems of ToM?

(4) In chapter 4 we found that 39-month-olds are able to represent second-order false belief as measured by their preferential looking time. If the early present ToM ability can support second-order belief, can we observe it in younger infants? Furthermore, representing second-order false belief requires an understanding of two agents' mental states and the recursive nature of mind. What is the working memory capacity for higher-order belief reasoning? More specifically, can children track two second-order false beliefs when they have the ability to represent one?

In sum, it is an empirical question of whether the origin of ToM is perceptual or conceptual in nature. Evidence over the last 30 years favors the hypothesis that we possess a conceptual module specialized for understanding agents' mental world. It emerges early in infants (e.g. 10 months as in Kovács, et al., 2010; 8 months as in He, et al., 2009; 6 months as in Southgate and Vernetti, 2014) and continues to operate throughout life. With the development of natural language and other cognitive abilities, the early competence of ToM might be elaborated and enriched by a non-modular system with more available power. Studies in this dissertation contribute to the on-going investigation of the nature of early theory of mind ability and how this relates to the later developed ability in preschoolers and adults. After 30 years, we still hunger for more insights on how the brain makes sense of another brain.

VI. References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106-111.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953-970.

Apperly, I. A., Samson, D., Carroll, N., Hussain, S., & Humphreys, G. (2006). Intact first- and second-order false belief reasoning in a patient with severely impaired grammar. *Social Neuroscience*, *1*(3-4), 334-348.

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: neuropsychological evidence from a false-belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience*, *16*(10), 1773-1784.

Apperly, I. (2010). *Mindreaders: The Cognitive Basis of "Theory of Mind"* (1 ed.). Hove, UK: Psychology Press.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, *35*(5), 1311-1320.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622-628.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Science*, *14*(3), 110-118.

Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development*, *62*(6), 1227-1246.

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*(3), 191-208.

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.

Bartsch, K., & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development*, *60*, 946-964.

Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge, MA: MIT Press.

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

Bloom, P. (2001). Roots of word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 159-184). Cambridge University Press.

Buresh, J. S., & Woodward, A. L. (2007). Infants track action goals within and across agents. *Cognition*, *104*(2), 287-314.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*(2), 337-342.

Callaghan, T., Rochat, P., Lillard, A., Claux, M. L., Odden, H., Itakura, S., .Singh, S. (2005). Synchrony in the onset of mental-state reasoning evidence from five cultures. *Psychological Science*, *16*(5), 378-384.

Carey, S. (2011). *The Origin of Concepts* (Reprint ed.). New York, NY: Oxford University Press, USA.

Carlson, S., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032-1053.

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

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*(2), 595-616.

Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*(4), 299-319.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, *69*(3), 672-691.

Carpenter, M., Call, J., & Tomasello, M. (2002). A new false belief test for 36-month-olds. *British Journal of Developmental Psychology*, *20*(3), 393-420.

Cassidy, K. W. (1998). Three- and four-year-old children's ability to use desire- and belief-based reasoning. *Cognition*, *66*(1), B1-11.

Castelli, F., Happe, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage*, *12*(3), 314-325.

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-scale deceit: Deception as a marker of two-, three-, and four-year-olds' early theories of mind. *Child Development*, 1263-1277.

Chen, M., & Leslie, A. M. (2009). Multiple object tracking in infants: Four (or so) ways of being discrete. In B. Hood & L. Santos (Eds.), *The origins of object knowledge* (pp. 85-107). New York, NY: Oxford University Press, USA.

Cheng, M., Wang, L., & Leslie, A. M. (2013). Preschoolers track multiple minds. Eighth Biennial Meeting of the Cognitive Development Society (CDS), *October 18-19*,Memphis, TN.

Clements, W. A., Rustin, C. L., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: A training study of false belief. *Developmental Science*, *3*(1), 81-92.

Cohen, A. S., & German, T. C. (2009). Encoding of others' beliefs without overt instruction. *Cognition*, *111*(3), 356-363.

Cohen, A. S., & German, T. C. (2010). A reaction time advantage for calculating beliefs over public representations signals domain specificity for 'theory of mind'. *Cognition*, *115*(3), 417-425.

Coull, G. J., Leekam, S. R., & Bennett, M. (2006). Simplifying second-order belief attribution: What facilitates children's performance on measures of conceptual understanding? *Social Development*, *15*(2), 260-275.

Davis, H. L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology*, *47*(1), 25-31.

de Villiers, J. (2005). Can language acquisition give children a point of view? In J. Astington & J. B. (Eds.), *Why Language Matters for Theory of Mind* (pp. 186-219). Oxford, UK: Oxford University Press.

de Villiers, J. (2007). The Interface of Language and Theory of Mind. *Lingua*, *117*(11), 1858-1878.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs.

de Villiers, J. G., & de Villiers, P. A. (2003). Language for thought: Coming to understand false beliefs. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 335-384). Cambridge, MA: MIT Press.

de Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development*, *17*(1), 1037-1060.

Dennett, D. (1978). *Brainstorms*. Cambridge, MA: Bradford Books/MIT Press.

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development*.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage*, *55*(2), 705-712.

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, *12*(6), 1127-1133.

Feigenson, L. (2007). The equality of quantity. *Trends in Cognitive Science*, *11*(5), 185-187.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: evidence from infants' manual search. *Developmental Science*, *6*(5), 568-584.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*, 109-128.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Friedman, O., & Leslie, A. M. (2004a). A developmental shift in processes underlying successful belief-desire reasoning. *Cognitive Science*, *28*(6), 963-977.

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

Friedman, O., & Petrashek, A. R. (2009). Children do not follow the rule "ignorance means getting it wrong". *Journal of Experimental Child Psychology*, *102*(1), 114-121.

Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, *20*(2), 177-182.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Science*, *7*(2), 77-83.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, *38*, 11-21.

Garnham, W. A., & Perner, J. (2001). Actions really do speak louder than words — but only implicitly: Young children's understanding of false belief in action. *British Journal of Developmental Psychology*, *19*, 413-432.

Garnham, W. A., & Ruffman, T. (2001). Doesn't see, doesn't know: Is anticipatory looking really related to understanding of belief? *Developmental Science*, *4*(1), 91-100.

Gelman, R. (1993). A rational-constructivist account of early learning about numbers and objects. In D. Medin (Ed.), *Learning and motivation* (pp. 61-96). Academic Press: New York.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, *349*(6305), 154-156.

Gopnik, A., & Wellman, H. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257-293). New York, NY: Cambridge University Press.

Gopnik, A., & Wellman, H. (1992). Why the child's theory of mind really is a theory of mind. *Mind & Language*, *7*, 145-171.

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

Grice, H. P. (1957). Meaning. *The Philosophical Review*, *66*, 377-388.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557-559.

Hamlin, J. K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science*, *13*(6), 923-929.

Happé, F., & Loth, E. (2002). 'Theory of mind' and tracking speakers' intentions. *Mind & Language*, 24-36.

Harris, P. (2005). Conversation, pretense, and theory of mind. In J. Astington & J. Baird (Eds.), *Why Language Matters for Theory of Mind* (pp. 70-83). Oxford, UK: Oxford University Press.

Harris, P. L., de, R., Marc, & Pons, F. (2005). Language and Children's Understanding of Mental States. *Current Directions in Psychological Science*, *14*(2), 69-73.

He, Z., Bolz, M., & Baillargeon, R. (2012). 2.5-year-olds succeed at a verbal anticipatory-looking false-belief task. *British Journal of Developmental Psychology*, *30*(Pt 1), 14-29.

He, Z., Bolz, M., & Baillargeon, R. (2011). False-belief understanding in 2.5-year-olds: evidence from violation-of-expectation change-of-location and unexpected-contents tasks. *Developmental Science*, *14*(2), 292-305.

Helming, K. A., Strickland, B., & Jacob, P. (2014). Making sense of early false-belief understanding. *Trends in Cognitive Science*, *18*, 167-170.

Henderson, A. M., & Woodward, A. L. (2012). Nine-month-old infants generalize object labels, but not object preferences across individuals. *Developmental Science*, *15*(5), 641-652.

Hogrefe, G., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, *57*, 567-582.

Hollebrandse, B., Hout, A., & Hendriks, P. (2012). Children's first and second-order false-belief reasoning in a verbal and a low-verbal task. *Synthese*.

Jordan, K. E., & Brannon, E. M. (2006). The Multisensory representation of number in infancy. *Proc Natl Acad Sci U S A*, *103*(9), 3486-3489.

Jordan, K. E., Suanda, S. H., & Brannon, E. M. (2008). Intersensory redundancy accelerates preverbal numerical competence. *Cognition*, *108*(1), 210-221.

Kahneman, D., & Treisman, A. (1984). Changing view of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of Attention* (pp. 29-61). New York, NY: Academic Press.

Káldy, Z., & Leslie, A. M. (2005). A memory span of one? Object identification in 6.5-month-old infants. *Cognition*, *97*(2), 153-177.

Káldy, Z., & Leslie, A. M. (2003). Identification of objects in 9-month-old infants: integrating 'what'and 'where'information. *Developmental Science*, *6*(3), 360-373.

Kibbe, M. M., & Leslie, A. M. (2011). What do infants remember when they forget? Location and identity in 6-month-olds' memory for objects. *Psychological Science*, *22*(12), 1500-1505.

Kibbe, M. M., & Leslie, A. M. (2013). What's the object of object working memory in infancy? Unraveling 'what' and 'how many'. *Cognitive Psychology*, *66*(4), 380-404.

Kobayashi, T., Hiraki, K., & Hasegawa, T. (2005). Auditory-visual intermodal matching of small numerosities in 6-month-old infants. *Developmental Science*, *8*(5), 409-419.

Koós, O., Gergely, G., Csibra, G., & Bíró, S. (1997). Why eating Smarties makes you smart? Understanding false belief at the age of 3. Poster presented at the biennial meeting of the Society for Research in Child Development, Washington, DC.

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, *14*(5), 402-408.

Leslie, A. (1986). Getting development off the ground: Modularity and the infant's perception of causality. In P. van Geert (Ed.), *Theory building in developmental psychology* (pp. 405-437). Amsterdam: Elsevier North-Holland.

Leslie, A. M. (1994a). ToMM, ToBy, and Agency: Core architecture and domain specificity. *Mapping the mind: Domain specificity in cognition and culture*, 119-148.

Leslie, A. M. (1994b). Pretending and believing: issues in the theory of ToMM. *Cognition*, *50*(1-3), 211-238.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind". *Trends in Cognitive Science*, *8*(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., & Káldy, Z. (2007). Things to remember: Limits, codes, and the development of object working memory in the first year. In L. M. Oakes & P. J. Bauer (Eds.), *Short-and long-term memory in infancy and early childhood: Taking the first steps toward remembering* (pp. 103-125). Oxford: Oxford University Press.

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

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: neuropsychological evidence from autism. *Cognition*, *43*(3), 225-251.

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: developing `what' and `where' systems. *Trends in Cognitive Science*, *2*(1), 10-18.

Lewis, C., & Osborne, A. (1990). Three-year-olds' problems with false belief: Conceptual deficit or linguistic artifact? *Child Development*, 1514-1519.

Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: a meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, *44*(2), 523-531.

Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: a training study. *Child Development*, *74*(4), 1130-1144.

Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: Relations with Complex syntactical mastery. *Child Development*, *81*(2), 597-615.

Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science*, *24*(3), 305-311.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *39*(20), 279-281.

Luo, Y. (2011). Do 10-month-old infants understand others' false beliefs? *Cognition*, *121*(3), 289-298.

Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition*, *105*(3), 489-512.

Luo, Y., & Johnson, S. C. (2009). Recognizing the role of perception in action at 6 months. *Developmental Science*, *12*(1), 142-149.

Luo, Y. (2010). Do 8-Month-Old Infants Consider Situational Constraints When Interpreting Others' Gaze as Goal-Directed Action? *Infancy*, *15*(4), 392-419.

Mareschal, D., & Johnson, M. H. (2003). The "what" and "where" of object representations in infancy. *Cognition*, *88*(3), 259-276.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Cambridge, MA: The MIT Press.

Meyer, M. L., & Lieberman, M. D. (2012). Social working memory: neurocognitive networks and directions for future research. *Frontiers in Psychology*, *3*, 571.

Meyer, M. L., Spunt, R. P., Berkman, E. T., Taylor, S. E., & Lieberman, M. D. (2012). Evidence for social working memory from a parametric functional MRI study. *Proc Natl Acad Sci U S A*, *109*(6), 1883-1888.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81.

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin*, *135*(5), 749-773.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*(2), 622-646.

Mitchell, P. (1994). Realism and early conception of mind: A synthesis of phylogenetic and ontogenetic issues. In C. Lewis & P. Mitchell (Eds.), *Children's Early Understanding of Mind: Origins and Development* (pp. 19-45). Hove, UK: Psychology Press.

Mitchell, P., & Lacohée, H. (1991). Children's early understanding of false belief. *Cognition*, 107-127.

O'Neill, D. K. (1996). Two-year-old children's sensitivity to parent's knowledge state when making requests. *Child Development*, 659-677.

Olson, D. R. (1988). On the origins of beliefs and other intentional states in children. In P. Harris & D. R. O. (Eds.), *Developing Theories of Mind* (pp. 414-426). New York, NY: Cambridge University Press.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255-258.

Perner, J. (1988). Developing semantics for theories of mind: From propositional attitudes to mental representation. In J. Astington, P. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 141-172). Cambridge, England: Cambridge University Press.

Perner, J. (1991). *Understanding the representational mind. Learning, development, and conceptual change.* Cambridge, MA: The MIT Press.

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Science*, *16*(10), 519-525.

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: how deep? *Science*, *308*(5719), 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.

Perner, J. (2010). Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In P. A. Frensch & R. Schwarzer (Eds.), *Cognition and neuropsychology: International perspectives on psychological science (Volume 1)* (pp. 241-262). Psychology Press.

Peterson, C. C., & Siegal, M. (1999). Representing inner worlds: Theory of mind in autistic, deaf, and normal hearing children. *Psychological Science*, *10*(2), 126-129.

Peterson, C. C., & Siegal, M. (2000). Insights into theory of mind from deafness and autism. *Mind & Language*, *15*(1), 123-145.

Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic Books.

Pyers, J. E., & Senghas, A. (2009). Language promotes false-belief understanding: evidence from learners of a new sign language. *Psychological Science*, *20*(7), 805-812.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: a sketch of the FINST spatial-index model. *Cognition*, *32*(1), 65-97.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179-197.

Quine, W. V. O. (1960). *Word and Object* (First Edition ed.). Cambridge, MA: The MIT Press.

Rakoczy, H. (2012). Do infants have a theory of mind? *British Journal of Developmental Psychology*, *30*(Pt 1), 59-74.

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development*, *74*(6), 1807-1822.

Roth, D., & Leslie, A. M. (1998). Solving belief problems: toward a task analysis. *Cognition*, *66*(1), 1-31.

Rubio-Fernández, P. (2013). Perspective tracking in progress: Do not disturb. *Cognition*, *129*(2), 264-272.

Rubio-Fernández, P., & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science*, *24*(1), 27-33.

Ruffman, T. (1996). Do children understand the mind by means of simulation or a theory? Evidence from their understanding of inference. *Mind & Language*, *11*, 388-414.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology*, *80*(3), 201-224.

Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? Response to Leslie. *Trends in Cognitive Science*, *9*(10), 462-463.

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, *73*(3), 734-751.

Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social understanding in children. *British Journal of Developmental Psychology*, *30*(Pt 1), 87-104.

Russell, J. A., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology*, *9*(2), 331-349.

Sabbagh, M., Moses, L. J., & Shiverick, S. (2006). Executive function and preschoolers' understanding of false beliefs, false photographs, and false signs. *Child Development*, *77*(4), 1032-1049.

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, *72*(4), 1054-1070.

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: a case of a selective deficit in inhibiting self-perspective. *Brain*, *128*(Pt 5), 1102-1111.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, *55*, 87-124.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking peopleThe role of the temporo-parietal junction in "theory of mind". *Neuroimage*, *19*(4), 1835-1842.

Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2012). Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General*, *141*(3), 433-438.

Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012). Cognitive load disrupts implicit theory-of-mind processing. *Psychological Science*, *23*(8), 842-847.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, Development and 'theory of mind'. *Mind & Language*, 131-153.

Scholl, B. J., & Leslie, A. M. (2001). Minds, modules, and meta-analysis. *Child Development*, *72*(3), 696-701.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: clues to visual objecthood. *Cognitive Psychology*, *38*(2), 259-290.

Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development*, *80*(4), 1172-1196.

Scott, R. M., Baillargeon, R., Song, H. J., & Leslie, A. M. (2010). Attributing false beliefs about non-obvious properties at 18 months. *Cognitive Psychology*, *61*(4), 366-395.

Scott, R. M., He, Z., Baillargeon, R., & Cummins, D. (2012). False-belief understanding in 2.5-year-olds: evidence from two novel verbal spontaneous-response tasks. *Developmental Science*, *15*(2), 181-193.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*(9), 668-671.

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. *Science*, *325*(5942), 883-885.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *298*, 199-209.

Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge, MA: Cambridge University Press.

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

Slaughter, V., & Peterson, C. C. (2012). How conversational input shapes theory of mind development in infancy and early childhood. In M. Siegal & L. Surian (Eds.), *Access to Language and Cognitive Development* (pp. 3-22). New York, NY: Oxford University Press, USA.

Song, H. J., & Baillargeon, R. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology*, *44*(6), 1789-1795.

Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, *13*(6), 907-912.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587-592.

Sperber, D., & Wilson, D. (1986). *Relavance: Communication and Cognition*. Cambridge, MA: Harvard University Press.

Sperber, D. (2000). Metarepresentations in an evolutionary perspective. In D. Sperber (Ed.), *Metarepresentations: A multidisciplinary perspective* (pp. 117-137). New York, NY: Oxford University Press.

Starkey, P., & Cooper, R. G. (1980). Perception of numbers by human infants. *Science*, *210*, 1033-1035.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition*, *36*(2), 97-127.

Strauss, M. S., & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development*, *52*, 1146-1152.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology*, *30*(3), 395-402.

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

Surian, L., & Geraci, A. (2012a). Where will the triangle look for it? Attributing false beliefs to geometric shape at 17 months. *British Journal of Developmental Psychology*, *30*, 30-44.

Surian, L., & Geraci, A. (2012b). Where will the triangle look for it? Attributing false beliefs to a geometric shape at 17 months. *British Journal of Developmental Psychology*, *30*(Pt 1), 30-44.

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

Tager-Flusberg, H., & Sullivan, K. (1994). A second look at second-order belief attribution in autism. *Journal of Autism & Developmental Disorders*, *24*(5), 577-586.

Tamir, D. I., & Mitchell, J. P. (2010). Neural correlates of anchoring-and-adjustment during mentalizing. *Proceedings of the National Academy of Sciences*, *107*(24), 10827-10832.

Tamir, D. I., & Mitchell, J. P. (2013). Anchoring and adjustment during social inferences. *Journal of Experimental Psychology: General*, *142*(1), 151.

Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 132-158). Cambridge University Press.

Tomasello, M., & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development*, *10*(2), 201-224.

Trabasso, T., & Stein, N. L. (1997). Narrating, representing, and remembering event sequences. In P. v. d. Broek, P. J. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension and representation: Bridging fictional and actual events* (pp. 237-270). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Träuble, B., Marinović, V., & Pauen, S. (2010). Early Theory of Mind Competencies: Do Infants Understand Others' Beliefs? *Infancy*, *15*(4), 434-444.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of Visual Behavior* (pp. 549-586).

Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and 'theory of mind'in an agrammatic aphasic patient. *Current Biology*, *10*(12), 723-726.

Varley, R., Siegal, M., & Want, S. C. (2001). Severe impairment in grammar does not preclude theory of mind. *Neurocase*, *7*(6), 489-493.

Wang, L., & Leslie, A. M. (2013). False Belief and Working Memory: Can Children Represent Two False Beliefs Spontaneously? Paper presented at the biennial meeting of the Society for Research in Child Development (SRCD), April 18-20, 2013,Seattle, WA.

Wang, L., & Leslie, A. M. (under review). The true-belief default in intuitive theory-of-mind: Competence and performance in belief-desire reasoning.

Wellman, H. M. (1990). *The Child's Theory of Mind*. Cambridge, MA: The MIT Press.

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition*, *30*, 239-277.

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

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1-34.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*(6389), 749-750.

Wynn, K. (1995). Infants possess a system of numerical knowledge. *Current Directions in Psychological Science*, *4*(6), 172-177.

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*(2), 343-368.

Yott, J., & Poulin-Dubois, D. (2012). Breaking the rules: do infants have a true understanding of false belief? *British Journal of Developmental Psychology*, *30*(Pt 1), 156-171.

Zaitchik, D. (1990). When representations conflict with reality: The preschooler's problem with false beliefs and "false photographs". *Cognition*, *35*, 41-68.

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

Zelazo, P. D., Carter, A., Reznick, S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, *1*(2), 198-226.

Zhang, W., & Luck, S. J. (2011). The number and quality of representations in working memory. *Psychological Science*, *22*(11), 1434-1441.