EXECUTIVE FUNCTIONING AND THEORY OF MIND DEVELOPMENT: A TRAINING STUDY

by

JEANNETTE ELIZABETH BENSON

A thesis submitted to the Department of Psychology

in conformity with the requirements for

the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

September, 2007

Abstract

This study was conducted to investigate the processes underlying the relation between executive functioning and false belief knowledge. We explored the hypothesis that children with advanced executive functioning skills are better equipped to capitalize on the experiences that are necessary to learn how to reason about others' mental states. To examine this possibility, we recruited 3.5-year-old children with age-appropriate variability in executive functioning skills to participate in a training study designed to promote their performance on false belief tasks. We found that individual differences in executive functioning task scores strongly and consistently predicted the extent to which children benefited from false-belief training. Importantly, the relation between executive functioning and false belief improvement remained significant after controlling for age, initial performance on mental state reasoning tasks, language skills, and executive functioning improvement across the testing period. Thus, our results support the hypothesis that executive functioning skills influence the extent to which children are able to capitalize on relevant experience to better predict and understand others' false-belief-based behaviour. This claim is discussed with respect to possible alternative explanations for our findings, and ensuing implications for understanding the interplay between neuromaturational factors and experience.

Acknowledgements

First and foremost, I would like to extend my sincerest thanks to the families who generously volunteered their time to participate in this study. I would also like to thank the undergraduate students who assisted with data collection: Katie Gamble, Ashley Hobb, Grace Liao, Anna Lom, and Linda van der Meer. Their enthusiasm and helpful insights were a valuable addition to the project.

I am extremely thankful to my advisor and mentor, Dr. Mark Sabbagh, for giving me the opportunity to work under his guidance. His support, feedback and direction have been invaluable, not only in the progression and completion of this project, but also in the advancement of my research abilities, writing style, and critical thinking skills. I would also like to extend my thanks to the Developmental Psychology department at Queen's University for providing a stimulating, supportive and positive environment in which to learn. In particular, I would like to thank the members of my thesis committee, Dr. Valerie Kuhlmeier and Dr. Tom Hollenstein, for their helpful feedback in the planning stages of this project, and their support and enthusiasm throughout my time as a graduate student.

I would also like to thank the Early Experience Lab members for providing such an optimistic, friendly working environment. Special thanks to Annette Henderson, Jennie Ito, Eric Bulmash, Sarah Bannoff and Lindsay Bowman for their advice, support and friendship over the past two years. I feel extremely lucky to have been in the company of such a great group of people.

I also wish to thank my family and friends for their continuing love and encouragement. In particular, I would like to thank my grandparents, Roy and Elsie

Gorewich, my father, David Benson, and my stepfather, Dr. Stan Hamstra, who have all played central roles in my life and continue to provide me with their much-appreciated love, support, and guidance. I would additionally like to thank Stan for providing me with thoughtful, constructive advice whenever I had project-related questions (and, also, for all those friendly statistics debates!) I also wish to extend many thanks to my boyfriend and best friend, Rob Nieckarz. I cannot thank him enough for his unfailing love, encouragement, humour, and support. I am also extremely grateful to the Nieckarz, and Dess families for welcoming me into their homes and providing me with their love and support throughout many work-filled weekends over the past 2 years.

Lastly, I would like to thank my mother, Dr. Nancy Benson, for all her love, support, enthusiasm, and her ability to see the humour in almost any situation. Throughout my life she consistently encouraged me to reach for my dreams, and to strive for excellence. Above all, she taught me the importance of optimism, kindness, perseverance and integrity. She has had a profound impact on almost every aspect of who I am as a person, and continues to be a true inspiration in my life. This thesis is dedicated to her loving memory.

TABLE OF CONTENTS

CHAPTER PAGE

CHAPTER PAGE

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

Skillful negotiation of the social world relies in large part on having a 'theory of mind'- the capacity to understand that others' actions are motivated by internal mental states (Wellman, 1992). Theory of mind has become a central focus of research within social, cognitive, comparative, developmental, and clinical psychology. Perhaps some of the most intriguing findings of late have come from developmental psychology, where cross-cultural research has shown that children in all cultures that have been investigated go through major transitions in their theory-of-mind reasoning during the preschool years (Callaghan et al., 2005). The apparent universality of the developmental trajectory is perhaps even more strongly highlighted by the rare cases in which theory-of-mind development and functioning seem to be disrupted, as have been seen in instances of brain injury (Apperley, Samson, Chiavarino, & Humphreys, 2004) or neurodevelopmental disorders such as autism (Baron-Cohen, 2001).

Much is known about the developmental trajectory of theory of mind. The cumulative findings of this research suggest that theory-of-mind understanding involves knowledge of multiple constructs, including beliefs, desires, intentions and emotions (Wellman & Liu, 2004). Moreover, these constructs typically come to be understood through a progressive series of advances in knowledge that begin during infancy and continue well into childhood and adolescence (Lalonde & Chandler, 2002; Wellman, 2002). Although there is some evidence to suggest that there might be subtle cross-cultural variations (e.g., Wellman, Fang, Liu, Zhu, & Liu, 2006), studies conducted on typically developing children in Western cultures have identified an invariant progression of conceptual advances across the preschool period. Specifically, children successively come to understand diverse desires, diverse beliefs, the origins of knowledge, the capacity to hold false beliefs, and the ability to regulate emotional expressions (Peterson, Wellman, & Liu, 2005; Wellman & Liu, 2004).

Within this developmental trajectory, considerable research has been placed on children's understanding that beliefs can be false. In what has come to be a standard test of false belief knowledge, children are asked to predict where a character will look for an object that he previously hid in one location when, unbeknownst to the character, the object was moved to an alternate location in his absence (Wimmer & Perner, 1983). While 3-year-old children typically fail this task, most 5-year-olds correctly report that the character will search in the out-of-date location. False belief tasks such as this explicitly test children's knowledge of the representational nature of mental states; in order to conceive of false beliefs, children must understand that mental states are representations of reality, and as such do not necessarily reflect the true state of the world (Wellman, Cross, & Watson, 2001). Thus, false belief understanding is considered to be the most compelling indicator of children's theory-of-mind knowledge.

Despite all that is known about the developmental timetable of theory-of-mind knowledge, comparatively less is known about the factors that promote developmental changes in understanding mental states. However, a burgeoning number of studies have identified a relation between experience and the developmental trajectory of theory-of-mind knowledge. For example, in lab settings, significant advances in children's false belief performance have been induced through the administration of training paradigms (e.g., Hale & Tager-Flusberg, 2003; Slaughter, 1998; Slaughter & Gopnik, 1996; Swettenham, 1996). These paradigms typically involve presenting children with modified theory-of-mind tasks where mental state concepts are highlighted and explained. A number of additional studies have identified more natural environmental factors that appear to be causally related to the progression of theory of mind, including parents' tendency to use mental state terms (e.g., Ruffman, Slade, & Crowe, 2002), number of older siblings in the home (Perner, Ruffman, & Leekam, 1994; Ruffman, Perner, Naito, Parkin, & Clements, 1998), and socio-economic status (Holmes, Black, & Miller, 1996).

Another line of research has examined the neuropsychological factors that may contribute to theory of mind. In particular, several studies have identified a significant relation between preschoolers' theory-of-mind development and the cognitive abilities associated with frontal lobe activity known as 'Executive Functions' (e.g., Carlson & Moses, 2001; Hala, Hug, & Henderson, 2003). Executive functioning is a term that encompasses the processes that underlie goal-directed behaviour, including self-regulation, planning, working memory, response inhibition and resistance to interference (Zelazo & Müller, 2002). In particular, several studies have identified a robust relation between false belief understanding and tasks that index developmental change in children's inhibitory control and working memory (e.g., Wellman et al., 2001; Hala et al., 2003).

While a considerable body of research has found a significant relation between executive functioning and false belief task performance, there is some debate regarding the nature of this relationship. In this regard, two important questions arise. The first concerns the direction of the relationship: does executive functioning contribute to false belief task performance, or do advances in false belief knowledge lead to improved executive functioning skills? Beyond this, a second question concerns the processes underlying the relationship between executive functioning and false belief knowledge: is the association artifactual, or is there an inherent developmental relation between the two constructs? In the following sections, I present evidence from recent research that provides insight into these two important questions.

The Direction of the Relation between Executive Functioning and False Belief

Perner and colleagues (Kloo & Perner, 2003; Perner & Lang, 1999; Perner, Stummer, & Lang, 1999) have argued that advances in theory of mind facilitate executive functioning task performance. That is, children's knowledge that mental states causally influence action enables them to pass both false belief and executive functioning tasks. The relation is direct in the case of false belief tasks: Children must learn that holding a false belief *causes* a character to look in the wrong location despite his or her goal to obtain the object. This knowledge relates to executive functioning skills somewhat more indirectly: An understanding of causally effective mental states enables children to represent the existence of a prepotent response that will prevent them from achieving their goal. It is only once they have a representation of this causally effective action schema that they will know to actively inhibit it. Therefore, Perner asserts that developments in the realm of theory-of-mind knowledge—specifically, an understanding of causally effective mental representations—enable children to conceive of their prepotent responses and consequently exert conscious control over their behaviour to succeed on executive functioning tasks.

To examine their hypothesis, Kloo and Perner (2003) trained children on either false belief or executive functioning measures, and examined the resulting effects of both training paradigms on children's false belief and executive functioning performance. To the extent that false belief knowledge contributes to executive functioning development, one might expect that the false belief training would cause children's executive functioning performance to improve. The study's results indicate that the false belief training group did show significant improvements in executive functioning task performance across the testing period, but their improvements were not significantly greater than that of a control group trained on an unrelated concept. Therefore, while the authors suggest that these findings provide evidence for the proposal that false belief knowledge facilitates executive functioning skills, support for this hypothesis appears to be weak at best.

An alternate, and more widely held view is that executive functioning precedes false belief performance. This suggestion is supported by a number of study findings. First, several studies have found that reducing the executive demands of false belief tasks results in children's improved performance (Freeman, Lewis, & Doherty, 1991; Mitchell & Lacohée, 1991; Moses, 1993; Wellman & Bartsch, 1988; Zaitchik, 1991; Carlson, Moses, & Hix, 1998; Couillard & Woodward, 1999). This finding is incompatible with the notion that false belief knowledge precedes executive functioning; changing executive components of false belief tasks should not influence subsequent performance if false belief knowledge develops prior to executive functioning skills (Carlson & Moses, 2001). Second, studies investigating the sequence of executive functioning and theory-of-mind development in 3- to 4- year olds have found that in most cases, children show competence on executive functioning tasks prior to improving on measures of false belief understanding (Flynn, O'Malley, & Wood, 2004; Flynn, 2007). Third, longitudinal research suggests the existence of a unidirectional relationship between the development of executive functioning and false belief (Carlson, Mandell, & Williams, 2004; Hughes, 1998; Flynn, 2007). For example, Hughes (1998) found that 3- to 4-year old children's executive functioning skills predicted their performance on false belief tasks 1 year later after controlling for age, verbal ability and initial false belief performance. In contrast, early false belief understanding did not predict later executive functioning abilities**.** Similar findings were reported by Flynn (2007), who used a microgenetic approach to examine the concurrent development of executive functioning and false belief knowledge in preschoolers. Her study replicated Hughes' findings on a smaller timescale by showing that early executive functioning predicted false belief performance 5 months later. Again, early false belief performance did not similarly predict later executive functioning. In sum, the results of several studies indicate that executive functioning likely contributes to children's performance on false belief tasks, not that theory of mind is a precursor to executive functioning development.

Processes Underlying the Relation between Executive Functioning and False Belief

Although the cumulative findings of recent research suggest that executive functioning plays an important role in children's false belief task performance, the best way to characterize this role is a subject of considerable debate. According to Moses (2001), the debate centers on two general proposals: the 'Expression' account and the 'Emergence' account. The Expression account suggests that children require a threshold level of executive functioning to solve false belief tasks, which themselves pose significant demands on executive functioning. According to this view, children fail false belief tasks not because they lack a conceptual understanding of false belief, but because the executive demands of the tasks are too high for them to demonstrate their underlying knowledge. One piece of evidence for this assertion is that when 3-year-olds are presented with the standard false belief task described above, they fail the forced choice prediction question not by responding with the correct answer at chance levels, but by consistently reporting that the character will search in the location where the object is currently hidden. The Expression account asserts that children with poor executive functioning skills may be incapable of resisting the urge to reference what they know to be true (in this example, the object's true location) in order to reveal their understanding that beliefs can be false (Carlson & Moses, 2001).

Also consistent with the Expression account, several studies have found that preschoolers reveal more advanced knowledge of mental states when the executive demands of false belief tasks are minimized (Freeman et al., 1991; Mitchell & Lacohee, 1991; Moses, 1993; Wellman & Bartsch, 1988; Zaitchik, 1991; Carlson et al., 1998; Couillard & Woodward, 1999). For example, when the standard false belief task is changed so that the hidden object is no longer present when the target question is asked, (e.g., while the character is gone, his chocolate is used up), 3-year-old children's performance improves significantly (Wellman et al., 2001). Thus, when the object's true location is no longer as salient a response option, children show better performance on the tasks.

A final piece of evidence to support the Expression account comes from research by Friedman and Leslie (2005), which shows that increasing the executive demands of false belief tasks causes a decline in preschoolers' performance. In a modified version of the standard false belief task, children were shown the usual series of events where a character hides an object, leaves the scene, and the object is then moved to an alternate location. In the modified task version, children are additionally told that the character wishes to *avoid* the hidden object. In this case, then, children must not only inhibit the urge to reference reality in order to consider the character's belief; they must also inhibit the prepotent notion that the character will go towards his or her object, not away from it. Furthermore, the task modification serves to increase the number of facts that must be held in mind when determining where the character will search for the object. Under these circumstances, even 4-year-olds who are able to pass the standard false belief task are unable to correctly predict where the character will look (Cassidy, 1998; Leslie, German, & Polizzi, 2005).

The implication that follows from these findings taken together is that children's executive functioning skills place a limit on the extent to which they are able to demonstrate their underlying knowledge of mental states on standard false belief tasks. Nonetheless, the results of a number of additional studies suggest that the relation between executive functioning and false belief extends beyond the role that executive functioning plays in negotiating task demands. First, a meta-analysis by Wellman et al. (2001) showed that lowering the executive demands of false belief tasks significantly improves preschoolers' performance, but doing so does not appear to bring young preschooler's performance above chance levels. Thus, children do not appear to demonstrate systematic understanding of false beliefs, even when the executive demands of the tasks are low.

A second line of evidence that would appear to run counter to the Expression account comes from studies showing that executive functioning skills correlate with performance on false belief tasks that do not seem to require preschoolers to overcome a prepotent response (e.g., Perner & Lang, 1999; Perner, Lang, & Kloo, 2002). For instance, Perner et al. presented children with a version of the standard false belief task that ended with the character searching for his object in the out-of-date location. Children were subsequently asked to explain the character's actions, first on their own, and then by choosing from forced choice options (one correct, one incorrect). Unlike the standard false belief task, the children in this study were not required to inhibit the urge to reference reality; they were instead asked to explain the story's occurrences after the fact. Although the executive demands of the task were therefore reduced, performance on the measure was still strongly correlated with executive functioning skills. This is compelling evidence to suggest that the relation between executive functioning and false belief performance is not driven by the executive demands of standard false belief tasks.

A final piece of evidence against the Expression account comes from research that shows cross-cultural dissociations in executive functioning and false belief development. For instance, Sabbagh, Xu, Carlson, Moses, and Lee (2006) compared Chinese (Beijing) and U.S. preschoolers on extensive batteries of both executive functioning and false belief tasks. Results showed that although the Chinese children were advanced on executive functioning relative to their U.S. counterparts, the two groups showed very similar performance on the false belief battery. Nevertheless, within the group of Chinese preschoolers executive functioning *was* predictive of performance on false belief tasks. These findings show that attaining a particular level of executive functioning is not alone sufficient for enabling children to negotiate the demands of false belief tasks and show strong performance.

Similarly intriguing findings were noted in a study by Melinder, Endestad, and Magnussen (2006) examining the relation between executive functioning and false belief performance in Norwegian preschoolers. Although these authors did not perform a direct comparison, children in the Norwegian sample obtained executive functioning scores that were comparable with those of U.S. children tested using comparable tasks, but their false belief task performance was notably more advanced. Specifically, 48% of the 3-year-olds in the Norwegian sample passed a standard contents change false belief task, whereas 3-year-olds in North American samples typically obtain much lower scores (See Callahan, et al., 2005; Carlson, et al., 2004; Carlson & Moses, 2001). Thus, the Norwegian sample showed evidence of superior false belief skills that were not matched by similar advances in executive functioning. These results provide more evidence that the absolute level of executive functioning skills does not predict children's performance on false belief tasks as the Expression account would predict.

In contrast to the Expression account, the 'Emergence' account stipulates that executive functioning skills contribute to the development of children's *understanding* of beliefs. On a fundamental level, children must require a certain degree of inhibition to disengage from current stimuli in order to consider any form of mental representation (Carlson & Moses, 2001). Beyond this, advances in executive functioning may facilitate children's false belief understanding in three important ways. First, executive functioning skills may enable children to suppress their own perspectives in order to consider and reflect upon the mental states of others (Carlson & Moses, 2001). Second, children with advanced executive functioning skills may be better equipped to competently engage in social interactions that in turn provide them with experiences relevant to theory of mind development (Flynn, 2007; Hughes, 1998). Finally, having the capacity to ignore irrelevant stimuli may increase children's ability to use information from both discussions with others and observations of people's actions in order to develop the understanding that beliefs can be false (Moses & Sabbagh, 2007).

The Emergence account offers a plausible explanation for some of the study results described above. First, it accounts for the finding that executive functioning abilities correlate with performance on false belief tasks that do not require the inhibition of a prepotent response. If executive functioning skills contribute to children's *understanding* that beliefs can be false, then performance on all valid measures of false belief understanding should be related to executive functioning skills, regardless of the extent to which the surface characteristics of the tasks require executive inhibition.

The Emergence account also provides some insight into the results from cross-cultural research on the relation between executive functioning and false belief understanding. According to this theory, executive functioning abilities enable children to capitalize on the types of experiences that provide them with information on others' mental states, thus fostering the development of a theory of mind. While the Chinese preschoolers were advanced in executive functioning, it may be that they had less exposure to the kinds of experiences that are necessary for theory-of-mind development. There is evidence to suggest that there are indeed substantial cultural differences in levels of exposure to experiential factors related to theory of mind. One such factor that is positively related to theory-of-mind development is the number of older siblings living in a child's household (Perner et al., 1994; Ruffman et al, 1998). Number of older siblings is thought to be related to theory-of-mind development because children with older siblings have more opportunities to discuss mental states with others (Sabbagh et al., 2006). Because the Chinese government enforces a law that prohibits more than one child per household, Chinese children may not have as many opportunities to engage in mental state discourse as children in the U.S.. It is thus possible that the advantages afforded to Chinese children as a result of superior executive functioning abilities are offset by their relatively limited exposure to important experiential factors.

Relative to Chinese children, the Norwegian sample showed the opposite effect; while they evidenced comparable levels of executive functioning skills as compared to U.S. children, they were advanced in their theory-of-mind performance. The Emergence account states that a child's theory-of-mind understanding develops as a function of *both* executive functioning skills and relevant experience. Thus, while the Norwegian children were not advanced in executive functioning, it could be that they had more exposure to the types of experiences that contribute to an understanding of mental states. Although no research has directly examined this possibility, it is interesting to note the significant cultural differences that exist in socio-economic status. Parents, and in particular single mothers, are much less likely to be poor in Norway than in other Western counties (Curtis & Phipp, 2004). This difference is largely a result of Norwegian government policies that provide considerable financial support to parents, and encourage mothers to further their education by providing them with tuition bursaries (Curtis & Phipp, 2004; Haavind & Magnusson, 2005). Given the positive relation between socio-economic status and children's performance on false belief tasks (Holmes et al., 1996), differences in social status and the resulting experiences that are associated with them may be an important factor contributing to differences in relative theory-of-mind knowledge.

Goals of the Present Study

The primary goal of our research was to examine whether the Emergence account accurately describes the relation between theory of mind and executive functioning in order to gain a better understanding of how the two are related. Despite indirect evidence for the Emergence account, no studies to date have examined the predicted interaction between executive functioning skills and the kinds of experiences that are important for theory-of-mind development. We posited that one way to investigate this interaction would be to expose children with varying executive functioning abilities to similar levels of experiential factors related to false belief knowledge in a controlled setting. To accomplish this, we conducted a training study to determine whether children with relatively advanced executive functioning skills were better able to learn from false belief training to further develop their understanding of mental states. Providing children with the same types of experiences in this way allowed us to examine whether children with high executive functioning are in fact better able to capitalize on relevant experiences to further develop their knowledge of other minds.

During an initial testing session, we measured children's understanding of mental states using a theory-of-mind scale developed by Wellman and Liu (2004), and a battery of tasks assessing false belief. Children then took part in 2 false belief training sessions over a 2- to 3-week period. The false belief battery was subsequently re-administered during a final testing session to assess for improvements in performance across the testing period. To examine 3-year-old children's executive functioning abilities, a number of well-established measures of executive functioning were administered throughout the testing period, and an aggregate executive functioning score was calculated based on children's performance. In addition, a language test was administered during the testing period so that scores on this measure could be used to control for children’s verbal abilities. In our analyses, we planned to investigate whether children's executive functioning skills predicted improvements in performance on the false belief tasks after controlling for relevant variables, including age, receptive language, and initial theory-of-mind knowledge. Findings in support of this relation would provide evidence in favour of the Emergence account; if executive functioning facilitates children’s false belief understanding, then children who have more advanced executive functioning abilities should be better equipped to benefit from false belief training.

Chapter 2: Method

Participants

Participants were 32 preschool-aged children (16 males, 16 females). At the time of the initial testing session, children ranged in age from 3 years, 6 months to 3 years, 11 months (*M* = 3 years, 9 months, *SD* = 1.57 months). Three additional children were recruited but excluded from the study's analyses because they refused to complete one or more of the study sessions (*n* = 2), or the testing environment was deemed inadequate due to high noise levels (*n* = 1). Eight of the study's participants were included in preliminary but not focal analyses because they showed strong performance on the false belief battery prior to training. These children scored above 50% on the initial false belief battery, indicating that they were already beyond the cognitive developmental milestones of interest. These children were, however, included in some of the preliminary analyses that examined the reliability of the measures, because these analyses depended on having a more typically distributed pattern of performance. Participants were recruited through local-area daycares, and through a database that was formed based on information from birth announcements and from parents who signed up at fairs, festivals, and other events in the area. The sample was representative of the primarily white, middle-class community in Southeastern Ontario from which they were recruited.

Materials and Measures

The measures were a false belief battery, a set of executive functioning tasks, false belief training tasks, a theory-of-mind scale, and a language test. Each measure is described below:

False Belief Battery

The scripts for the 4 tasks included in the false belief battery are presented in Appendix A. The battery tasks measure children's ability to identify either their own or others' false beliefs, and predict resulting behaviour. All tasks were presented to children using toy figurines, puppets and picture props. Children received the battery at both the initial and final testing sessions, and changes in performance were used to gauge improvements in false belief understanding. To avoid presenting each child with the same scenarios twice, the surface characteristics of the tasks (i.e., the characters and objects involved) were changed across testing sessions. Although a contents change false belief task is included in both the theory-of-mind scale and the false belief battery, Contents Change was administered only once during the initial testing session. A brief description of the tasks included in this battery follows:

Contents Change (Gopnik & Astington, 1988). Children were shown a container that would normally contain something familiar (e.g., a Smarties box) and asked to state what it contains. The experimenter went on to show them that in actuality, the box contained something unexpected (e.g., pencils). The children were then introduced to a character (e.g., Tigger) and asked to identify what the character, who had never seen inside the box before, would think it contained. They were then asked the memory control question: "Did [character's name] see inside the box?" Children were given a score of 1 if they answered both the test and memory questions correctly, and a score of 0 if they responded incorrectly to one or both questions. In order to succeed on this task, children had to recognize the character's false belief regarding the container's contents.

Location Change (Wimmer & Perner, 1983). A character was shown to hide an object in one location and leave the scene. In his/her absence, a second character moved the object from its original location and hid it somewhere else. The first character was then shown to return, and the children were asked the control question: "Did [character's name] *see* the [object] get moved?" If the children responded incorrectly by saying that they did see the object get moved, the story was re-enacted and the question was asked again. After the children responded correctly to the control question, the test question was asked: "When [character's name] comes back, where will he look for the [object]?" Children were given a score of 1 if they correctly answered the test question, and a score of 0 if they were unable to do so. In order to succeed on this task, children needed to recognize the initial character's false belief that the object still remained in the location where it was originally hidden.

Appearance-Reality (Flavell, Flavell, & Green, 1983; Flavell, Green, & Flavell, 1986). Children were shown 2 objects that had misleading appearances. The first was an object that appeared to be one thing, but was actually another (e.g., a sponge that looked like a rock). The second was a picture that appeared to be one colour when covered by a transparent screen, but was actually another (e.g., an orange castle that looked black under a screen). Children were asked first how the object looks "when you look at it right now", and second, what the object is "really and truly". On each trial, children received a score of 1 if they correctly responded to both test questions.

Deceptive Pointing (Carlson et al., 1998). Children saw an experimenter place a toy into one of two closed boxes. As the experimenter pointed to the box, she explained to the children that you can indicate an objects' whereabouts by pointing to where it is hidden. The children were then asked to place the toy in a box and point to it in order to further demonstrate this concept. Following this exercise, children were introduced to a puppet (e.g., Elephant), and they were told that "Now we can put the toy in a box and see if [puppet's name] can find it." Prior to having the child hide the object, the experimenter placed the puppet in a closed bin where he was reported to be unable to see. After the children hid the object in one of the boxes, they were told to play a trick by pointing to make the puppet look in the wrong box. In order to succeed at this task, children had to point to the empty box upon the puppet's return, thus conveying a misleading clue regarding the toy's whereabouts. During each session, children participated in 2 trials that each involved playing a trick on a different character. On each trial, children were awarded a score of 1 if they successfully used deception, and a score of 0 if they were unable to do so.

Executive Functioning Tasks

The complete scripts for these tasks are given in Appendix B. All executive functioning tasks required children to inhibit their initial impulses in order to respond in alternate ways. The Whisper, Card Sort and KRISP tasks were administered on only one occasion, during the second training session. The Bear/Dragon and Grass/Snow tasks were administered at both the initial and final testing sessions. Executive functioning abilities have been shown to develop at a gradual pace over the preschool years (Zelazo, Müller, Frye & Marcovitch, 2003). Therefore, we did not expect to find a significant change in children's executive functioning performance across the 2- to 3-week testing period. Despite this, administering the same executive functioning tasks at the initial and final testing sessions allowed us to ensure that improvements in children's false belief performance could not be explained by concurrent advances in executive functioning.

Bear/Dragon (Kochanska, Murray, Jacques, Koenig, 1996). This task is a modified version of a "Simon Says" game. The experimenter first asked children to perform 10 simple, clearly defined actions (e.g., touch your head, touch your tummy, touch your feet, etc.). They were then introduced to a bear puppet and a dragon puppet, and instructed to do what they were told by the "nice bear", but not what they were told by the "mean dragon". The puppets, which were both controlled by the experimenter, then took turns telling the child to act out the actions that were performed earlier in the game. Practice trials where children were given feedback on their answers were administered until the children correctly responded to 1 bear trial and 1 dragon trial. This task measures executive functioning by examining the extent to which children are successful at ignoring the mean dragon's instructions. Responses on the dragon test trials were coded as follows: 0 = full correct movement, 1 = partial correct movement, 2 = incorrect movement (e.g., touches feet when instructed to touch head), 3 = strategic movement (e.g., shakes head no, moves hands behind back), and 4 = no movement. Responses on 30% of the Bear/Dragon trials were re-coded by a second experimenter to determine inter-rater reliability. Inter-rater reliability was 94.7% on the practice trials, and 97.9% on the test trials. Standardized scores were calculated based on children's average performance across trials. An aggregate Bear/Dragon score was calculated by combining standardized total scores on the dragon practice and test trials from both testing sessions.

Grass/Snow (Carlson & Moses, 2001). A black board with a white card attached to the upper left-hand corner and a green card attached to the upper right-hand corner was placed in front of the children. The experimenter asked the children to point to the colour of grass and then to the colour of snow. After establishing that the children knew the corresponding colours, the experimenter instructed them to play a silly game by pointing to the green square anytime she said the word "snow", and to the white square anytime she said the word "grass". Practice trials where children were given feedback on their answers were administered until the children correctly responded to 1 grass trial and 1 snow trial. Following this, 16 test trials were administered where no feedback was provided. This task measures children's ability to inhibit the impulse to point to the colour that corresponds to the spoken word in order to provide the correct response. On each of the test trials, children were given a score of 1 for initially pointing to the correct square and 0 for either initially pointing to the incorrect square or pointing to both squares at the same time. Responses on 30% of the Grass/Snow trials were re-coded by a second experimenter to determine inter-rater reliability. Inter-rater reliability was 94.7% on the practice trials, and 98.7% on the test trials. An aggregate Grass/Snow score was calculated by combining standardized total scores on the practice and test trials from both testing sessions.

Card Sort (Zelazo et al., 2003). Before beginning, two boxes were placed in front of the children, one attached to a picture of a red rabbit and the second attached to a picture of a blue boat. Children were then sequentially handed cards that varied across two dimensions: colour (i.e., red or blue) and shape (i.e., rabbits or boats). For the first 5 cards, children were asked to sort according to shape (i.e., rabbit cards go in the box with the rabbit card on the front, boat cards go in the box with the boat card on the front). For the last 5 cards, children were asked to switch rules and sort based on colour (i.e., red cards go in the box with the red card on the front, blue cards go in box with the blue card on the front). Switching rules requires executive functioning to inhibit the urge to continue sorting by the first rule in order to sort by the second rule. Standard scores were calculated based on the number of correct post-switch trials.

Whisper Task (Kochanska et al., 1996). To begin, children were first asked to whisper their own names. They were then presented with a series of cards depicting both familiar cartoon characters (e.g., Dora the Explorer) and unfamiliar cartoon characters (e.g., Huckle), and instructed to whisper the names of the characters whose names they knew. This task involved executive functioning because children had to inhibit the impulse to blurt out the names of the characters they knew, which became particularly challenging after they were shown some that they were unable to name. Responses to familiar characters were coded in the following manner: 0 = shout, 1 = normal or mixed voice, and 2 = whisper. Responses on 30% of the Whisper trials were re-coded by a second experimenter, and inter-rater reliability was 93.4%. Standardized scores were calculated based on children's average performance across trials.

Kansas Reflection-Impulsivity Scale for Preschoolers (KRISP, Wright, 1971). Children were sequentially shown target pictures and asked to identify an exact match to the target from an array of four to six similar pictures. This task was designed to measure children's ability to refrain from responding impulsively in order to consider all the response options. An aggregate KRISP score was calculated based on standardized average scores for children's accuracy (total # of errors, reverse scored), and the time taken to respond by pointing to a picture (the first choice on each trial).

False Belief Training

A sample false belief training script is presented in Appendix C. The training procedure administered was modeled after a training paradigm developed by Hale and Tager-Flusberg (2003). During each training session, the experimenter enacted 4 location change stories for the children using toy figurines and props. Incorrect responses were given corrective feedback with a simple explanation and story re-enactment, while correct responses were confirmed and a shortened explanation was provided with no reenactment. Characters and surface features of the stories were changed across all false belief tasks administered. In addition, the direction of object transfer (from the left side of the scene to the right, or vice versa), the characters' positions, and the direction of the characters' departures from the scene were counterbalanced across training session trials to prevent children from learning a simple rule to pass the tasks based on the location or movement of characters or objects.

During the second training session, children were asked to provide explanations for their answers on each of the 4 training trials. Children's responses were classified into the following categories: Relevant Situational Explanation, Irrelevant Situational Explanation, Uninterpretable Response, and Don't-Know Response (see Table 1). Inter-rater reliability across all coded responses was 94%, with the two raters coding 47 of the 50 explanations in the same way. Of the 3 explanations that were coded differently across raters, only one involved a situational explanation; the remaining two were instances where the child provided the explanation “Because.” This explanation was classified as Uninterpretable by one coder, and No-Response by the second.

Table 1

Types of False Belief Explanations

Explanation	*Examples*
Situational - relevant	*“That’s where it was when he left.”* *“Cause that’s where he left it when he leaved.”*
Situational – irrelevant	*“It’s in there.” (pointing to alternate location)* *“He won’t find his seashell.”* *“Cause the man moved [the object] in [the other location].”*
Uninterpretable	*“Because his shell.”* *“Because he should.”* *“Because, he’ll look in there and there and there (pointing to multiple locations).”*
Don’t Know (or no response)	*“I don’t know.”*

This particular training paradigm was chosen for 3 reasons. First, Hale and Tager-Flusberg (2003) found that their training method had a significant impact on children’s false belief task performance; children’s mean percent correct score on false belief measures went from 12.5% at pretest to 73.9% at post-test. Second, the training, which took the form of location change tasks, appeared to influence children’s performance on a variety of structurally different false belief tasks; children in the false belief training group had significantly higher post-test scores on location change, contents change, and appearance-reality tasks in comparison to those of a control group who received training on an unrelated concept. Finally, while the training did result in significant improvements in false belief task performance, not all children benefited from the training. This was important, as it allowed us to examine whether executive functioning predicted the extent to which children's false belief performance improved.

Theory of Mind Scale - Wellman and Liu (2004)

The scripts for each task are presented in Appendix D. The tasks included in the scale assess children's understanding of desires, beliefs, knowledge, false beliefs, and discrepancies between real and apparent emotions. Toy figurines, puppets and picture props were used to illustrate the task scenarios in an attempt to maximize the children's engagement in the tasks. The tasks included in the scale were administered in a set order, becoming progressively more difficult. For each task, correct responses were awarded a score of 1, while incorrect responses were awarded a score of 0. Scores obtained on this scale were used in conjunction with scores on the initial false belief battery to control for children's theory-of-mind understanding at the beginning of the testing period. A brief description of each task included in the scale is provided below:

Diverse Desires. Children were shown pictures of two types of snacks (an apple and a cookie), and asked to state which snack they would prefer to eat. After they provided an answer, they were informed that a character (Big Bird) prefers the alternate snack. Children were subsequently asked to state which snack Big Bird would choose to eat. In order to answer the question correctly, children had to understand that Big Bird would choose the snack that was in accordance with his own preferences, not their own.

Diverse Beliefs. Children were shown a puppet (Cookie Monster), and then presented with a picture of two locations: a garage and a tree. They were told that Cookie Monster wants to find his cat, and that his cat might be hiding in either location. Children were subsequently asked to state where they thought the cat was. After they provided an answer, they were informed that Cookie Monster thinks the cat is in the alternate location. Children were then asked to state where Cookie Monster would look for his cat. In order to answer the question correctly, children had to understand that Cookie Monster would look in the location that was in accordance with his beliefs, not their own.

Knowledge Access. Children were shown a jewelry box and asked what they thought was inside. After they responded with an answer or stated that they did not know, the box was opened to reveal a small toy frog. The jewlery box was closed again, and children were introduced to a character, Kitty. They were told that Kitty had never seen inside the box, and then asked: “Does Kitty *know* what’s inside the box?” In order to answer the question correctly, children had to understand that Kitty did not have the same knowledge about the box’s contents that they themselves had.

Contents Change. See description of the contents change false belief task included in the false belief battery.

Real-Apparent Emotion. Children were first shown pictures of a happy face, a sad face, and neutral face and asked to label them. This exercise was included to ensure that children had knowledge of facial expressions that were necessary to correctly respond to the task's target questions. The children were then told a story about a character named Matt who tried to hide how he felt when his friend told a mean joke about him in front of his peers. Children were asked 1) how Matt really felt when everyone laughed, and 2) how Matt tried to look on his face when everyone laughed. To respond correctly, the children had to answer the question about how Matt *felt* more negatively than the target question about how Matt *looked*. This task requires children to understand that one's facial expressions may not be in accordance with one's internal feelings.

Language Measure

The Peabody Picture Vocabulary Test, Revised (PPVT-R, Dunn & Dunn, 1981) is a standardized measure of children's receptive language abilities. This measure was included to control for general language skills when examining individual differences in children's performance on the study tasks. In each trial, children were shown a page with four line drawings. Children were then asked to point to the line drawing that corresponded with a vocabulary word. Children continued through progressively more difficult items until they incorrectly answered more than 8 in a set of 12, or until they completed all 9 sets of words in the task. Children's scores on the measure were initially calculated by subtracting the total number of errors made from the highest item answered, as outlined in the instruction manual (Dunn & Dunn, 1981). As a result of experimenter error, the administration of the PPVT-R was terminated early for two children in the study sample (administration of the test was terminated at set 7). In the interest of keeping these two children in the study sample, new scores for all children were calculate based on their performance up to set 7. Correlations of original scores with scores up to set 7 were very high, $r = .998$, $p < .001$. Therefore, children's PPVT-R scores up to set 7 were used for all further analyses.

Procedure

Children were tested individually across 4 videotaped sessions, including an initial testing session, 2 training sessions, and a final testing session. Each session took approximately 20 to 30 minutes, and the same female experimenter administered all study sessions. The 4 sessions were conducted within a 2- to 3-week period, with a minimum of 2 days lapsing between each session. Based on each family's preference, the study sessions were held in the children's daycare ($n = 4$), in the laboratory ($n = 4$), or in a quiet room in the family's home ($n = 24$). During the initial testing session, children were presented with the false belief battery, two executive functioning tasks, and the theory of mind scale. The second session consisted of the first 4 false belief training tasks and the PPVT-R. Over the course of the third session, the remaining 4 false belief training tasks were administered along with 3 additional executive functioning measures. During the final testing session, the false belief battery and executive functioning tasks from the initial testing session were re-administered. The tasks were administered in a fixed order for all children (see Table 2). Although presenting the tasks in an invariant sequence may have produced order effects, it was important to keep the task order consistent across participants to ensure that individual differences in children's performance on the tasks could not be attributed to variance in task administration.

Table 2

Task Administration Sequence

Session 1	Session 2	Session 3	Session 4
Contents Change	FB Training 1	FB Training 5	Contents Change
Location Change	FB Training 2	Card Sort	Location Change
Appearance/Reality	FB Training 3	FB Training 6	Appearance/Reality
Deceptive Pointing	FB Training 4	Whisper	Deceptive Pointing
Bear/Dragon	PPVT-R	FB Training 7	Bear/Dragon
Grass/Snow		KRISP	Grass/Snow
Diverse Desires		FB Training 8	
Diverse Beliefs			
Knowledge Access			
Real-Apparent Emotion			

Chapter 3: Results

The primary goal of this study was to determine whether executive functioning skills would predict the extent to which children benefited from false belief training. We will begin by reporting results from preliminary analyses on age and sex. We will then report findings from analyses on the dependent measures independently, followed by a discussion of focal analyses examining the relation between executive functioning and false belief improvement. It should be noted that some of the analyses on the internal and external reliability of the batteries and measures depended on normal, age-appropriate variability in task performance and thus were conducted on an extended sample of children ($n = 32$). In contrast, all analyses involving indexes of children's false belief improvement and their relation to executive functioning were conducted on the focal sample of children who showed little to no evidence of false belief knowledge at the onset of the testing period (i.e., those who failed more than 50% of the Time 1 false belief tasks, $n = 24$).

Preliminary Analyses

Preliminary analyses were carried out to compare boys' and girls' scores on each of the individual dependent measures. No significant differences were found, so all further analyses were conducted collapsing across sex.

Further analyses were conducted to determine whether performance on the dependent measures was related to age. Because of the sample's narrow 6-month age range, we did not expect to find age-related trends in task performance. Results of correlation analyses confirm that performance on the dependent measures did not show relations with age, with one exception: age was significantly correlated with

Time 1 false belief battery performance, r (30) = .460, p = .008. Because false belief battery performance is one of our key dependent measures, we elected to control for age in our focal analyses.

Control Variables

To ensure that any relations between executive functioning and children's false belief improvement could not be explained by a common underlying relation to either language development or initial theory-of-mind knowledge, we measured these and controlled for them statistically in the focal analyses.

Language Measure

Raw scores on the PPVT-R ranged from 27 to 75 (M = 51.7, SD = 12.9). Performance within this range is expected of children in this age group. This provides further confidence regarding the typicality of our sample, and that the measure was administered appropriately.

Theory-of-Mind Scale

Mean scores and standard deviations for the tasks included in the scale are presented in Figure 1. The first 3 tasks of the scale are of particular interest because they measure understandings that reliably emerge prior to false belief understanding (i.e., diverse desires, diverse beliefs, and knowledge access). In line with previous research (e.g., Wellman & Liu, 2004), children's scores on the tasks comprising the theory-of-mind scale declined with each successive task. Thus, children's total score from the first 3 tasks was used in conjunction with initial false belief battery scores to control for initial theory-of-mind knowledge.

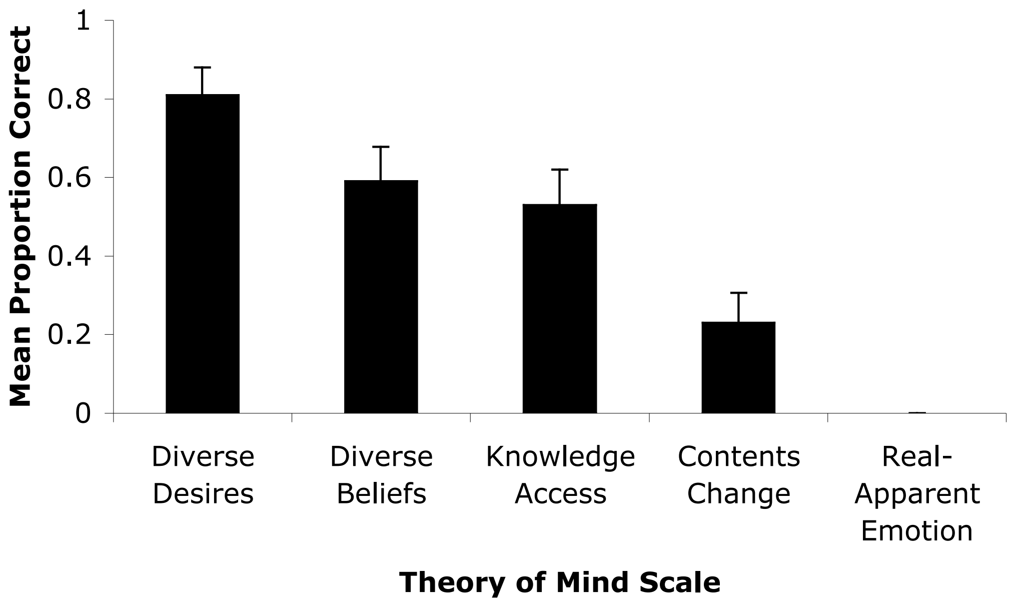

Figure 1. Mean Scores and Standard Errors on the Theory-of-Mind Scale

Executive Functioning Aggregate

The executive functioning battery was comprised of 5 tasks: Bear/Dragon, Grass/Snow, Card Sort, Whisper and KRISP. Bear/Dragon and Grass/Snow were administered twice, once during the initial testing session and again during the final testing session.

Bear/Dragon

Both administrations of Bear/Dragon involved 2 dependent measures: number of tries taken to pass the practice trials (Time 1: $M = 2.34$, $SD = 2.41$, Time 2: $M = 1.80$, $SD = 2.07$), and percentage score on the test trials (Time 1: $M = 74.1\%$, $SD = 35.0\%$, Time 2: $M = 76.1\%$, $SD = 34.0\%$). As predicted, children who passed the Bear/Dragon practice trials after fewer tries did significantly better on test trials (Time 1: r (30) = .916, $p < .001$, Time 2: r (28) = .895, $p < .001$). Moreover, aggregate time 1 and 2 Bear/Dragon scores were strongly related, r (28) = .839, $p <$ .001. Thus, overall Bear/Dragon scores were calculated by combining the 4 standardized Bear/Dragon dependent measures: initial and final practice scores (reverse-scored), and initial and final test scores.

Grass/Snow

Both administrations of Grass/Snow involved 2 dependent measures: number of tries taken to pass the practice trials (Initial: M = 3.28, SD = 2.05, Final: M = 2.54, SD = 1.11), and percentage score on the test trials (Initial: M = 59.3%, SD = 32.8%, Final: M = 67.7%, SD = 23.8%). Children who passed the Grass/Snow practice trials after fewer tries did significantly better on test trials during the initial testing session, r (30) = .377, p = .036. However, the relation dropped below significance during the final testing session, r (28) = .046, n.s.. This may have been a result of decreased variance in performance at time 2, with 75.0% of children passing both practice trials on the first try, as compared to 53.1% at time 1. Aggregate time 1 and 2 Grass/Snow scores were strongly related, r (27) = .453, p < .018. Thus, overall Grass/Snow scores were calculated by combining the 4 standardized Grass/Snow dependent measures: initial and final practice scores (reverse-scored), and initial and final test scores.

Card Sort

Children were scored on 5 post-switch trials where they were asked to sort cards according to colour after previously having to sort according to shape. All children responded correctly on the 2 compatible post-switch trials where following either rule (sort by shape/sort by colour) resulted in the same sorting outcome. Percentage scores were relatively low on the remaining 3 incompatible post-switch trials where children had to inhibit their urge to sort by shape in order to successfully sort by colour (*M* = 35.6%, *SD* = 47.9%).

Whisper

Children were scored on the extent to which they were able to maintain a whisper while naming familiar and unfamiliar characters' names. The sample's performance was near ceiling on the measure (*M* = 88.8%, *SD* = 18.5%). Overall, 63% of children obtained perfect scores.

KRISP

This task involved 2 dependent measures: average time taken to respond on the test trials (*M* = 4.79, *SD* = 1.70), and total number of errors made across trials (*M* = 5.64, *SD* = 3.24). As predicted, children who took more time to give their responses were also significantly more accurate, *r* (28) = .426, *p* = .019. The error totals were reverse scored, and standardized performance on the 2 dependent measures was combined to form overall KRISP aggregate scores.

Relations among Tasks and Scale Aggregating

Initially, we had intended to aggregate children's performance on all 5 of the executive functioning tasks into a single scale score. However, inspection of the internal reliability of the 5-item executive functioning battery revealed that it was low (alpha = .419). Analyses by item showed that removal of the KRISP and Whisper tasks resulted in a sharp increase in the reliability of the measure (alpha = .688). With respect to Whisper, as reported above, children's performance was near ceiling, which may have affected the extent to which it could correlate with other tasks. With respect to KRISP, closer inspection of previous research that has used this task has found only moderate (and not always significant) relations with other executive functioning tasks (e.g., Carlson & Moses, 2001). Given these factors, and the desirability of having a highly reliable measure given our relatively small sample size, we elected to proceed with the more reliable 3-item aggregate that included Grass/Snow, Bear/Dragon, and Card Sort.

Improvements in False Belief Performance

Improvements in false belief understanding were characterized with two goals in mind. The first was to determine whether executive functioning predicted performance on the final administration of the false belief battery while controlling for performance in the first session. The second was to determine whether children with stronger executive functioning skills started showing improved performance on false belief tasks earlier in the testing session relative to children with poorer executive functioning skills. We will address each of these focal questions in turn.

False Belief Battery Improvements

The false belief battery—including the Contents Change, Location Change, Deceptive Pointing, and Appearance/Reality tasks—was administered during both the initial and final testing sessions. Proportions on each task (0 to 1.0) were used instead of sums when forming initial and final false belief battery scores to give equal weight to each of the four tasks. The internal reliability of the scale was respectable at the initial testing session (alpha = .743) and at the final testing session (alpha = .600). Figure 2 shows children’s performance on the tasks comprising the battery at both time points.

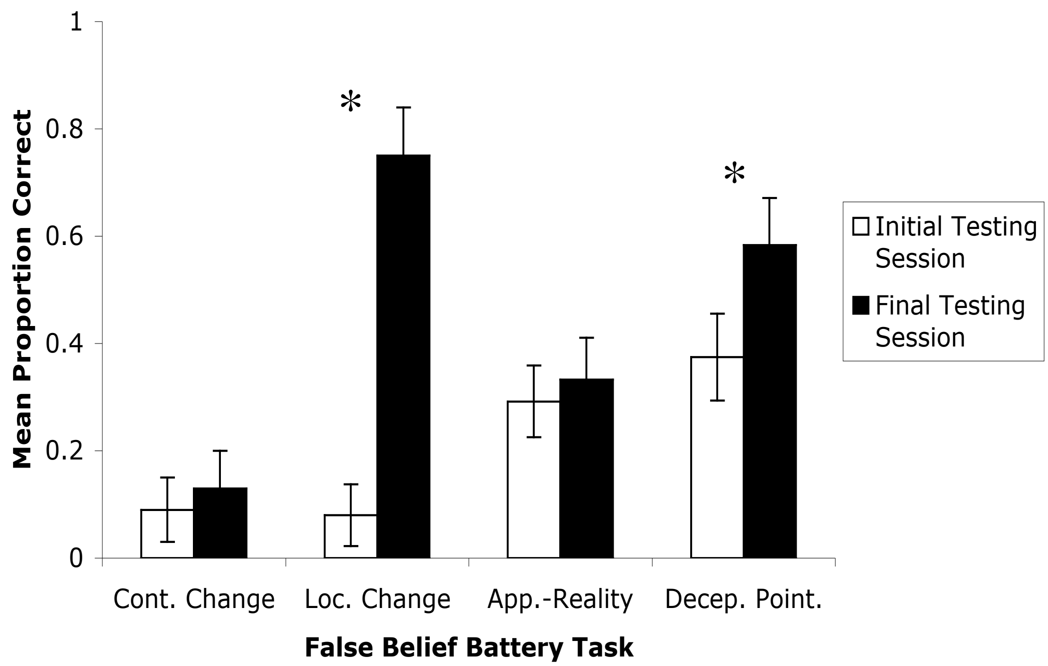

Figure 2. Mean Proportion Scores for Initial and Final False Belief Battery Tasks

Results of a repeated measures t-test on the focal sample comparing initial and final false belief battery scores showed that children performed significantly better on the second administration of the battery in comparison to their initial performance, $t\,(23) = 4.20$, $p < .001$, thus showing that the training was effective in general. However, closer inspection of the results showed that training was only effective in significantly improving performance on two of the tasks: Location Change, which was the focus of training, $t\,(23) = 6.78$, $p < .001$, and Deceptive Pointing, $t\,(23) = 2.46$, $p = .020$. The significant improvement on Deceptive Pointing was particularly notable because it provided some evidence that the training had an effect on children's underlying understanding of false beliefs in general and not just in the circumscribed context of the location change task.

Our main question was whether false belief battery improvement varied with children's performance on the 3-item executive functioning aggregate. To test this, we first performed a partial correlation analysis measuring the relation between executive functioning performance and performance on the final false belief battery, controlling for performance on the initial false belief battery. This analysis showed a strong, significant relation, r (21) = .526, p = .010. The relation held up even after adding age, PPVT-R and theory-of-mind scale performance to the list of control variables in the partial correlation analysis, r (18) = .462 p = .040. Thus, executive functioning predicted the extent to which children benefited from training.

Given that the bulk of the battery improvement was owed to children's improvement on the location change false belief task, we re-ran the same analyses with only improvement on Location Change as the dependent measure. The pattern of results was the same, only stronger. The first partial correlation analysis controlling for only initial location change task scores showed a strong relation between final location change task performance and executive functioning aggregate scores, r (21) = .576, p = .004. Furthermore, controlling for age, PPVT-R and theory-of-mind scale scores did not diminish the magnitude of the relation, r (18) = .582, p = .007.

Improvements across Training

Training performance of the focal sample is presented in Figure 3. Visual inspection of this graph suggests that overall, children's performance steadily improved across the two training periods. Results of a repeated measures ANOVA confirmed this trend: The main effect of training was significant, F (7, 161) = 5.61, $p < .001$, as was a follow-up linear contrast, F (1, 23) = 27.9, $p < .001$.

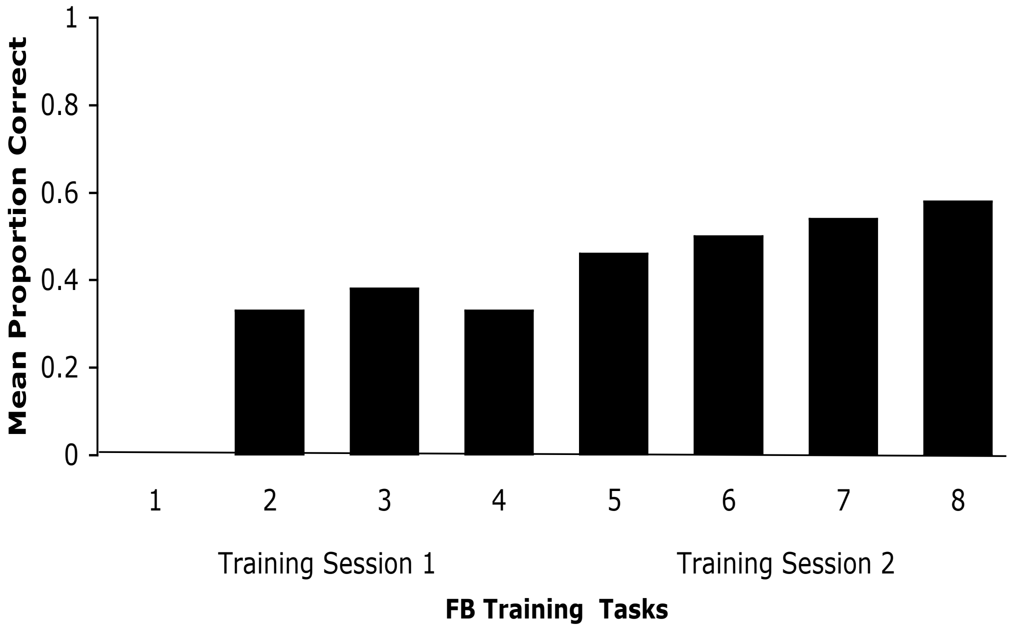

Figure 3. Histogram Depicting Children's Performance on False Belief Training Tasks.

To address the question of whether stronger executive functioning affected children's ability to capitalize on the training, we conducted several analyses. Our first looked simply at how many tasks children responded correctly to over the training sessions. Raw correlations showed that performance on the 3-item executive functioning aggregate predicted the number of test questions children answered correctly during training, r (22) = .620, p = .001. This relation held up in a partial correlation analysis that controlled for age, theory-of-mind scale scores and PPVT-R, r (19) = .528, p = .014. [1]

[1] This analysis was re-computed using performance on both the initial false belief battery and the theory-of-mind scale to control for initial theory-of-mind knowledge. Under these conservative conditions, the relation between false belief training and executive functioning aggregate scores remained strong and significant, r (18) = .560, p = .010.

Perhaps more intriguing was the inspection of individual children's performance across the training sessions. Figure 4 shows individual children's performance in each of the training sessions, with asterisks denoting correct performance on the task. Children are listed in order of their executive functioning aggregate scores (highest at top). In Figure 5, we divided children into high and low executive functioning groups (median split) and graphed each group's mean performance across the 8 training trials. What seemed clear from these figures is that children with strong executive functioning showed more robust performance earlier in the training sessions than did children with poorer executive functioning. Results of a 2 (training session) x 2 (high/low executive functioning) repeated measures ANOVA recapitulated the significant main effects of training, F (1, 22) = 14.7, p = .001, and executive functioning, F (1, 22) = 5.28, p = .032, but the interaction between the two factors was not significant, F (1, 22) = .024, *n.s.*, thereby suggesting that the extent of improvement was similar across executive functioning groups.

Although there was no significant interaction, it was of particular interest whether there was evidence that one group improved more quickly than the other. To examine this possibility, we compared the high and low executive functioning groups on mean number of items passed (out of 4) in the first and second training sessions. For the first session, children in the high executive functioning group outperformed the low executive functioning group, t (22) = 2.36, p = .029. For the second session, however, the difference in performance between the high and low executive groups dropped below significance, t (22) = 1.70, n.s. These findings suggest that children with stronger executive functioning skills were better able to capitalize on the training early on in the training period and were able to apply their knowledge on subsequent trials. In contrast, children with poorer executive functioning were eventually able to benefit from training, but not until later stages.

EF Rank Score	Session 1				Session 2			
	1	2	3	4	5	6	7	8
High EF Group								
1	-	*	*	*	*	*	*	*
2	-	*	-	-	-	*	*	-
3	-	*	*	*	*	*	*	*
4	-	-	*	*	*	*	*	*
5	-	*	*	*	*	*	*	*
6	-	-	*	*	*	*	*	*
7	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	*
9	-	*	*	*	*	*	*	-
10	-	*	*	-	-	-	*	*
11	-	-	-	-	-	-	-	*
12	-	-	-	-	*	-	-	*
Low EF Group								
13	-	-	-	-	*	*	*	*
14	-	-	*	-	-	-	-	-
15	-	-	-	-	-	-	-	*
16	-	-	-	-	-	-	-	-
17	-	*	-	-	*	*	*	*
18	-	-	-	-	*	*	*	-
19	-	-	-	*	-	-	*	*
20	-	*	*	*	-	*	-	-
21	-	-	-	-	-	-	-	*
22	-	-	-	-	*	*	-	-
23	-	-	-	-	-	-	*	-
24	-	-	-	-	-	-	-	-

Figure 4. Individual Children's False Belief Training Performance, Listed in Order of their Executive Functioning Aggregate Scores (Highest at Top, Division between High and Low Executive Functioning Marking the Median Split).

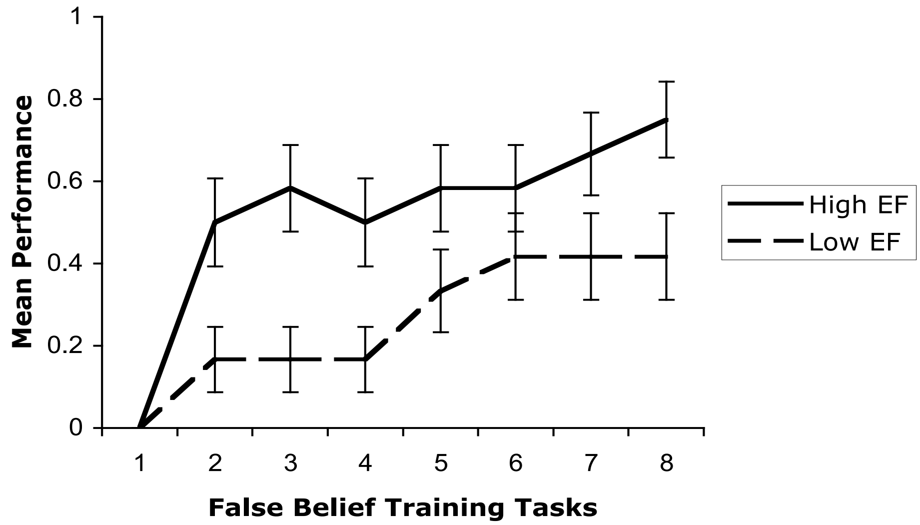

Figure 5. High and Low Executive Functioning Groups' Performance Across Individual False Belief Training Tasks

An important question concerns whether improvement over the training session reflects the actual acquisition of knowledge relevant to passing the false belief task. One way of exploring this is to look at how children justified their correct answers. Recall that children were asked to justify correct responses only on the four trials of the second session. Across all subjects there were 50 correct answers that could be justified – 31 from 11 out of 12 children in the high executive functioning group, and 19 from 9 out of 12 children in the low executive functioning group. Explanations were coded into the following categories: Relevant Situational, Irrelevant Situational, Uninterpretable, and Don't Know/No Response. None of the children in either group justified their correct responses with reference to mental states. The mean proportion of responses coded into each category across subjects is summarized in Figure 6. As can be seen from inspection of the graph, justifications of children in the high executive functioning group most commonly included reference to relevant situational factors (e.g., he'll look there "cause that's where it was when he left"). This is not surprising given that relevant situational information was precisely what the experimenter provided in the course of the training (see Appendix C). In contrast, children in the low executive functioning group tended to justify their responses by stating irrelevant situational facts (e.g.,"it's not there").

To account for the fact that some children provided more than 1 explanation (depending on how many test questions they responded correctly to), each of these children received a score that reflected the proportion of relevant situational explanations provided. An independent samples *t*-test on these scores showed that children in the high executive functioning group were more likely to provide relevant situational explanations overall, t (18) = 2.37, p = .030. These findings give some confidence that the reason children with high executive functioning were doing so well on these tasks was because they were capitalizing on the information present in the training, and using that information to guide their responses to the test questions.

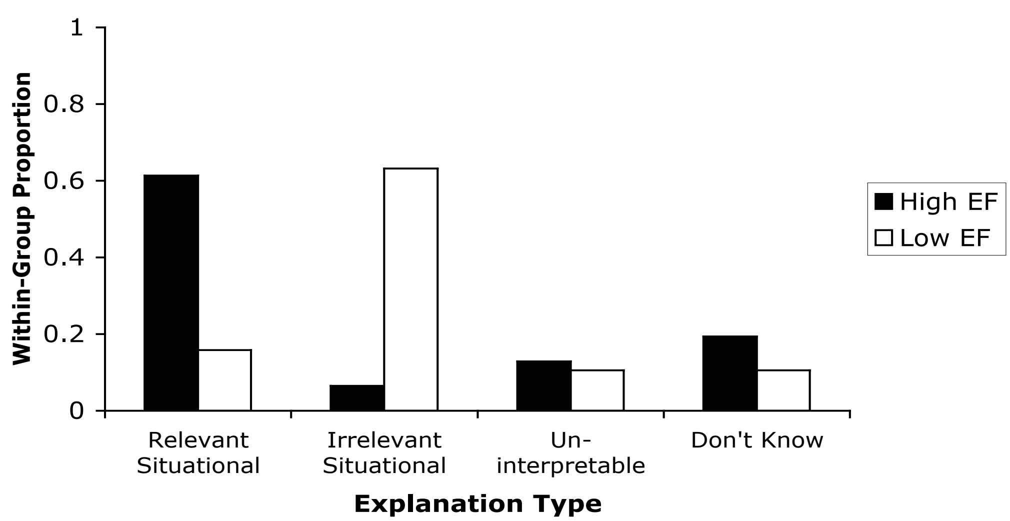

Figure 6. Proportion of Explanation Types Provided by High and Low Executive Functioning Groups Respectively on False Belief Training Tasks

False Belief Improvement Aggregate

Children's final false belief battery scores and their scores on the false belief training tasks were strongly correlated, r (22) = .756, $p < .001$. Thus, the two sets of scores were aggregated to form a third dependent measure indexing children's overall false belief improvement across the testing period. We reasoned that this analysis might provide the most robust characterization of the extent to which variation in executive functioning affected children's ability to capitalize on training.

Following the same general strategy as above, we conducted a first analysis investigating the correlation between executive functioning and the false belief improvement aggregate, controlling only for performance on the false belief battery at time 1. This analysis showed a strong relation, r (21) = .612, p = .002. When age and children's scores on the theory-of-mind scale and PPVT-R measures were additionally controlled, the relation remained strong and statistically significant, r (18) = .554, p = .011.

Controlling for Executive Functioning Improvement

Although our analyses were consistent in showing that individual differences in executive functioning predicted the extent to which children capitalized on the training, there was one potential confound. Specifically, participation in the false belief training might have improved children's executive functioning across the testing period. Individual differences in executive functioning improvement, then, might have accounted for individual differences in false belief improvement.

To test this possibility, we investigated the extent to which children's performance improved on the two executive functioning tasks that were administered at both the initial and the final testing sessions: Bear/Dragon and Grass/Snow. Both tasks had two dependent measures (number of practice trials before one correct performance, and total performance on the test trials). Paired sample t-tests showed that performance did improve on the practice trials for both measures, but not on the total performance measures, though the trends were in the right direction (see Table 3). A likely explanation for this pattern of results is that children were familiar with the tasks during the second testing session and thus took less time to show that they understood the rules of the game. Nonetheless, to ensure that whatever improvement in executive functioning that occurred over the training session cannot account for children's false belief improvement, we computed an aggregate executive functioning improvement score by computing the difference between initial and final practice and test scores on the Bear/Dragon and Grass/Snow tasks independently. The 4 resulting improvement scores were then standardized and combined to form an executive functioning improvement aggregate score (alpha = .632).

Table 3

Repeated Measures t-Tests Examining Executive Functioning Improvement

Comparison	*df*	*t*	Sig (2-tailed)
Bear/Dragon Practice	23	2.26	.034
Bear/Dragon Test	23	.998	.329
Grass/Snow Practice	21	2.71	.013
Grass/Snow Test	22	1.22	.237

All of the above analyses investigating the relation between executive functioning and false belief improvement were re-run adding the executive functioning improvement variable to the list of controls. The findings are summarized in Table 4. In each case, the magnitude of the correlation was similar, and the *p*-values remained significant.

Table 4

Partial Correlation Analyses Additionally Controlling for Scores on Executive Functioning Improvement Aggregate (n = 24)

False Belief Improvement Measure	Control Variables Included	Correlation with Executive Functioning Aggregate
Final False Belief Battery	Age Initial FB Battery Initial ToM Battery PPVT-R EF Improvement Aggr.	r (17) = .457, p = .049
Final Location Change Score	Age Initial Location Change Initial ToM Battery PPVT-R EF Improvement Aggr.	r (17) = .589, p = .008
Total False Belief Training Score	Age Initial ToM Battery PPVT-R EF Improvement Aggr.	r (18) = .539, p = .014
False Belief Improvement Aggregate	Age Initial Location Change Initial ToM Battery PPVT-R EF Improvement Aggr.	r (17) = .566, p = .012

Chapter 4: Discussion

Previous research has identified a relation between executive functioning and false belief reasoning (e.g., Carlson & Moses, 2001; Hala et al., 2003). This study aimed to provide insight into the nature of this relation by testing whether executive functioning was related to the extent to which children's understanding of mental states improved as a result of false belief training. The study's results showed that executive functioning was strongly and consistently related to false belief task improvement across the testing period. Moreover, these relations remained significant after age, initial theory of mind knowledge, receptive language skills and executive functioning improvement were controlled for. In addition to showing greater improvement overall, children with high executive functioning skills took less time to show training improvements, and the explanations they provided for their answers suggested that they were more successful at using the training feedback to better their performance.

This pattern of results is consistent with the Emergence account of the relation between theory of mind and executive functioning. As noted above, proponents of this account suggest that executive functioning may be necessary for capitalizing on the experiences necessary to develop an understanding of concepts related to theory of mind, especially false belief. Thus, a critical question concerns the extent to which we can confidently claim that children with advanced executive functioning truly did capitalize on the experiences we provided in the service of acquiring concepts that are critical for passing the false belief task. In the next sections, we raise three possible concerns with this strong claim and discuss the relevant evidence.

Did Children Truly Benefit from Training?

One limitation of the present study is that we did not compare the false belief training group's improvement to that of a control group trained on an unrelated concept. As a result, we were unable to statistically test whether improvements were more substantial than we might have expected over the same period had we not provided them with relevant theory-of-mind experiences. Establishing this would be important because the Emergence argument hinges on demonstrating the effects of experience on false belief improvements.

While this is a concern, there are several findings from the extant research literature that support the conclusion that children truly did benefit from our training, per se. First, the research on which we modeled our training paradigm showed similar rates of improvement on the location change training tasks in comparison to what we saw in our own training (Hale & Tager-Flusberg, 2003). More important, that previous research study did use a control group, which did not show similar improvements over the short testing period. These findings give confidence that improvements were related to the training. Second, the striking trajectory of improvement we observed over the course of the training period is not typical of that seen in general false belief development. Usually, false belief development is slow – microgenetic studies show that when children are given no training, individual children show meager improvements from 3.5- to 4-years-old (e.g., Amsterlaw & Wellman, 2003). In contrast, we saw dramatic improvements, especially in the high executive functioning group, over just 2 weeks. It is therefore difficult to argue that this striking improvement is not related directly to our training regimen.

Did Children Learn about False Beliefs?

A second part of the Emergence account is that as executive functioning improves, children are better able to capitalize on experience to acquire the concepts necessary for passing false belief tasks. An important question, then, concerns whether we have any evidence to suggest that they did indeed acquire these relevant concepts.

Our evidence is somewhat mixed on this question. The clearest evidence for children having acquired a generalized conceptual understanding of false belief would be if good performance transferred from the location change tasks on which they were trained to other false belief tasks, such as Contents Change, Appearance Reality, or Deceptive Pointing. We did find some transfer effects – children showed significant post-test improvement on the deceptive pointing version of the false belief task. However, unlike previous research that administered this training regimen (e.g., Hale & Tager-Flusberg, 2003), we did not index improvements on the contents change or appearance reality tasks. Thus, our transfer evidence was modest, which may weaken our strong claim that the children in our study acquired conceptual understandings relevant to passing false belief tasks.

The alternative to the conceptual development account of our findings is that children with high executive functioning skills were better able to use their training experience to make use of effective situation-specific rules for passing the false belief tasks. For example, children might have learned a rule like "the character will look opposite to where the object is." Of course, it is important to note that for a "rule" explanation to work, it would have to be formulated to account for the transfer improvement we saw on the deceptive pointing task, and not be useful on appearance-reality or contents change. Nonetheless, insofar as such a formulation might be possible, it is important to consider how executive functioning skills might have related to children's acquisition of such rules.

There are two plausible explanations for how rule-based learning might have been related to executive functioning. The first is that children with high executive functioning were better able to capitalize on experience with training in order to learn the situation-specific rules that served to improve their performance on subsequent tasks. This theory is comparable to the Emergence account, with one important exception: children are capitalizing on experience to acquire a *situation-specific rule* that allows them to pass future false belief tasks, rather than using experience to develop their *conceptual knowledge of beliefs* for the same end. A second possibility is that children with high executive functioning are generally better able to adhere to rules, likely in part because they are better able to resist the prepotent urge to respond otherwise. Thus, the high executive functioning children may have been better able to follow a behavioural rule to pass the false belief tasks, just as they were better able to follow the executive functioning task rules.

Children's response explanations provide some evidence to suggest, at the very least, that this second alternative is not the sole contributor to the relation between executive functioning and false belief improvement. Children in the high executive functioning group commonly provided response explanations that could be construed as situation-specific rules (e.g., he'll look there "cause that's where he left it when he leaved."), while children in the low executive functioning group often provided irrelevant facts, or statements that could be construed as low-level rules at best (e.g., he'll look there "cause it's not there"). Moreover, children with high executive functioning abilities typically gave false belief task explanations that closely matched the experimenter feedback, whereas this was not the case for the low executive functioning group. These findings suggest that the difference between high and low executive functioning was not simply the extent to which children were able to *employ* relevant rules. Instead, a possible explanation for our results is that children with advanced executive functioning skills were better able to use the information presented over the course of training to *derive* rules that helped them on subsequent tasks.

In sum, our data do not conclusively show whether the training had an impact on children's conceptual grasp of false beliefs, or instead on their acquisition and use of relevant situation-specific rules. Although the latter alternative may not necessarily reflect an explicit understanding of false beliefs, it may still represent a conceptual shift towards this knowledge; noticing behaviour in others that conflicts with what you know to be true and picking up on relevant situational factors that help to predict that behaviour may very well represent a meaningful step towards a conceptual understanding of false beliefs.

In any event, it is clear that executive functioning skills affected the extent to which children were able to make use of the training feedback to improve their performance on subsequent tasks. Thus, our findings lend preliminary support to the Emergence account, which stipulates that children with well-developed executive functioning skills are better able to make use of relevant experiences in order to better understand and predict others' behaviour.

Can a Case Be Made for the Expression Account?

In suggesting that our data are consistent with the Emergence account of the relation between theory of mind and executive functioning, it is also important to show that our data are not consistent with the competing Expression account. Recall that the Expression account posits that executive functioning and false belief performance are superficially related in that children require a certain level of executive inhibition to demonstrate their underlying mental state knowledge on standard false belief tasks (Moses, 2001). Thus, the Expression account predicts that advances in performance on standard false belief tasks rely on concurrent advances in executive functioning skills.

There are two ways in which one might attempt an Expression account for our data. The first is that perhaps the false belief training actually affected some children's executive functioning skills, and those children became the children who went on to do well in the post-test. We were able to rule this possibility out with our data. There was a small suggestion that executive functioning skills might have improved over the course of the study in that as a group, children needed fewer practice trials on the post-test executive functioning tasks than they did in the pre-test. However, there was no evidence from the test trials that executive functioning skills actually improved, and the improvement in practice trials could simply have been related to increased familiarity with the tasks. Even assuming that there was underlying improvement in executive functioning skills, partial correlation analyses revealed that executive functioning skills remained significantly related to children's false belief improvement after controlling for improvements in executive functioning across the testing period.

A second way in which an Expression account might be attempted is by suggesting that practice with the false belief tasks in the training session led to a decrease in the executive demands inherent to the tasks. If this were true, then false belief improvement in children with well-developed executive functioning skills may have occurred because their pre-existing executive functioning skills became sufficient to overcome the gradually lowered executive task demands. In contrast, lowered executive task demands from practice with the tasks may not have been sufficient to improve the performance of children with less-developed executive functioning skills to the same degree. Given children's pattern of performance across the testing period, we think this is unlikely. First, it seems unlikely that practice would have decreased the executive demands of the tasks to cause such a dramatic improvement in performance after even the first training trial, which is where we saw a great degree of improvement for our high executive functioning group. Second, it is unclear how this explanation could account for our findings involving children's false belief task explanations. Children with high executive functioning skills not only provided more correct responses to the false belief test questions; they also typically provided relevant explanations for their correct false belief task responses. These findings are not predicted by the Expression account - explaining correct responses to false belief questions after the fact does not require the inhibition of a prepotent response. Therefore, children with less-developed executive functioning skills should have been equally as capable of explaining their correct answers as children with well-developed executive functioning skills. Our findings overall contribute to the growing number of studies suggesting that the relation between executive functioning and false belief task performance goes beyond the inhibitory demands inherent to false belief tasks.

Future Directions

Our findings show that executive functioning skills predict the extent to which children are able to absorb and make use of experience with false belief training to improve their performance on subsequent tasks. While these results offer preliminary support for the Emergence account, future research is necessary to directly test the extent to which our findings reflect what occurs in real-world settings. Does executive functioning predict the extent to which children benefit from everyday experiential factors that are known to influence false belief knowledge, such as parental use of mental state terms? Do our findings reflect how executive functioning and experience interact to contribute to more natural developmental trajectories of false beliefs knowledge? Answering these questions would provide conclusive evidence on the extent to which the Emergence account accurately characterizes the relation between executive functioning and false belief knowledge.

Another important avenue for future research involves examining the neurological changes that underlie theory-of-mind development, and how these changes might themselves be related to both relevant experiential factors and executive functioning skills. The impact of experience on brain structures has been observed in a number of carefully designed studies (e.g., Maguire et al., 2003).

However, no research has directly linked the development of brain regions associated with theory of mind to experience, or to the neuropsychological changes that underlie advances in executive functioning. While our study represents a first step in understanding how executive functioning and experience might work together to foster advances in theory-of-mind knowledge, future research should aim to better understand how these factors inter-relate on a neurological level.

Conclusions

In sum, our training study is an important addition to the existing cross-sectional, cross-cultural, and microgenetic research on the processes underlying the relation between executive functioning and false belief reasoning. Through the use of a false belief training paradigm, we were able to examine how children's executive functioning skills interact with experience to facilitate changes in performance on false belief tasks. Our data provide strong evidence that executive functioning skills predicted the extent to which children capitalized on experience in the form of false belief training. While it is unclear whether our training induced a full-fledged understanding of false beliefs in our sample, our results nevertheless suggest that children with well-developed executive functioning appear to be better suited to make use of the types of experiences that are relevant to advances in false belief knowledge. Moreover, our findings add to the growing number of studies that provide evidence against the Expression account. This study makes an important contribution to our understanding of the factors that may facilitate, or conversely impeded, healthy social development.

References

Amsterlaw, J., & Wellman, H. M. (2006). Theories of mind in transition: A microgenetic study of the development of false belief understanding. *Journal of Cognition and Development, 7*, 139-172.

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. Special Issue: Social Cognitive Neuroscience, 16*, 1773-1784.

Baron-Cohen, S. (2001). *Theory of Mind and Autism: A Review*. San Diego, CA: Academic Press.

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

Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology, 40*, 1105-1122.

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

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

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

Couillard, N. L., & Woodward, A. L. (1999). Children's comprehension of deceptive points. *British Journal of Developmental Psychology, 17*, 515-521.

Curtis, L., & Phipps, S. (2004). Social transfers and the health status of mothers in Norway and Canada. *Social Science and Medicine, 58*, 2499-2507.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised: Form M.* Circle Pines, MN: American Guidance Service.

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive psychology, 15*, 95-120.

Flavell, J. H., Green, F. L., & Flavell, E. R. (1986). Development of knowledge about the appearance-reality distinction. *Monographs of the Society for Research in Child Development, 51*, 1-68.

Flynn, E. (2007). The role of inhibitory control in false belief understanding. *Infant and Child Development. Special Issue: Using the Microgenetic Method to Investigate Cognitive Development, 16*, 53-69.

Flynn, E., O'Malley, C., & Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Developmental Science, 7*, 103-115.

Freeman, N. H., Lewis, C., & Doherty, M. J. (1991). Preschoolers' grasp of a desire for knowledge in false-belief prediction: Practical intelligence and verbal report. *British Journal of Developmental Psychology. Special Issue: Perspectives on the Child's Theory of Mind: I, 9*, 139-157.

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

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearanceeality distinction. *Child Development, 59*, 26-37.

Haavind, H., & Magnusson, E. (2005). The nordic countries - welfare paradises for women and children? *Feminism & Psychology, 15*, 227-235.

Hala, S., Hug, S., & Henderson, A. (2003). Executive function and false-belief understanding in preschool children: Two tasks are harder than one. *Journal of Cognition and Development, 4*, 275-298.

Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science, 6*, 346-359.

Holmes, H. A., Black, C., & Miller, S. A. (1996). A cross-task comparison of false belief understanding in a head start population. *Journal of Experimental Child Psychology, 63*, 263-285.

Hughes, C. (1998). Finding your marbles: Does preschoolers' strategic behavior predict later understanding of mind? *Developmental Psychology, 34*, 1326-1339.

Kloo, D., & Perner, J. (2003). Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. *Child Development, 74*, 1823-1839.

Kochanska, G., Murray, K., Jacques, T. Y., & Koenig, A. L. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*, 490-507.

Lalonde, C. E., & Chandler, M. J. (2002). Children's understanding of interpretation. *New Ideas in Psychology. Special Issue: Folk Epistemology, 20*, 163-198.

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

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S. J., & Burgess, N. (2003). Navigation expertise and the human hippocampus: A structural brain imaging analysis. *Hippocampus, 13*, 250-259.

Melinder, A., Endestad, T., & Magnussen, S. (2006). Relations between episodic memory, suggestibility, theory of mind, and cognitive inhibition in the preschool child. *Scandinavian Journal of Psychology, 47*, 485-495.

Mitchell, P., & Lacohée, H. (1991). Children's early understanding of false belief. *Cognition, 39*, 107-127.

Moses, L. J. (1993). Young children's understanding of belief constraints on intention. *Cognitive Development, 8*, 1-25.

Moses, L. J. (2001). Executive accounts of theory-of-mind development. commentary on "meta-analysis of theory-of-mind development: The truth about false belief.". *Child Development, 72*, 688-690.

Moses, L.J., & Sabbagh, M.A. (2007). Interactions between domain general and domain specific processes in the development of children's theories of mind. In M.J. Roberts (Ed.) *Integrating the Mind: Domain General versus Domain Specific Processes in Higher Cognition* (pp. 275-291). New York: Psychology Press.

Perner, J., Stummer, S., & Lang, B. (1999). *Executive Functions and Theory of Mind: Cognitive Complexity or Functional Dependence?* Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences, 3*, 337-344.

Perner, J., Lang, B., & Kloo, D. (2002). Theory of mind and self-control: More than a common problem of inhibition. *Child Development, 73*, 752-767.

Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of mind is contagious: You catch it from your sibs. *Child Development, 65*, 1228-1238.

Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development, 76*, 502-517.

Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. A. (1998). Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology, 34*, 161-174.

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*, 734-751.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science, 17*, 74-81.

Slaughter, V. (1998). Children's understanding of pictorial and mental representations. *Child Development, 69*, 321-332.

Slaughter, V., & Gopnik, A. (1996). Conceptual coherence in the child's theory of mind: Training children to understand belief. *Child Development, 67*, 2967-2988.

Swettenham, J. (1996). Can children be taught to understand false belief using computers? *Journal of Child Psychology and Psychiatry, 37*, 157-165.

Wellman, H. M. (1992). *The Child's Theory of Mind*. Cambridge, MA: MIT Press.

Wellman, H. M. (2002). *Understanding the Psychological World: Developing a Theory of Mind*. Malden, MA: Blackwell Publishing.

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning and 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*, 655-684.

Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychological Science, 17*, 1075-1081.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*, 523-541.

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

Wright, J. C. (1971). *KRISP (Kansas Reflection-Impulsivity Scale for Preschoolers).* Unpublished manuscript, University of Kansas, Lawrence.

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

Zelazo, P. D., & Müller, U. (2002). *Executive Function in Typical and Atypical Development*. Malden, MA: Blackwell Publishing.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function. *Monographs of the Society for Research in Child Development, 68*, 11-27.

Appendix A

Sample False Belief Battery Tasks

Contents Change

Show child a Smarties box with pencil crayons inside. “Here’s a Smarties box. What do you think is inside the Smarties box?” Next, the Smarties box is opened. “Well let’s see….it’s really pencil crayons inside.” Close Smarties box. “Okay, what’s in the Smarties box?” If the child responds correctly that it contains pencils, then proceed to the following paragraph. If the child responds incorrectly, start the task from the beginning.

Show child a stuffed animal, Tigger. “Here’s Tigger. Tigger’s never seen inside this Smarties box before.” Test question: “So, what does Tigger think is in the box, Smarties or pencil crayons?” Memory question: “Did Tigger *see* inside this box?” To be correct, the child must answer “Smarties” to the test question and “No” to the memory question.

Location Change False Belief

Place a miniature toy box, a miniature set of drawers, and a toy car on the table in front of the child. “Now let me show you what else I have. Here are two kids. This one’s name is Eric,” (make Eric say hello) “and this is Annette” (make Annette say hello). “Eric and Annette are playing with this toy car.” Show the dolls playing together. “After a while, Eric decides to go play outside. He puts the car away in the toy box, and goes outside to play.” Show Eric placing the car in the toy box, and then remove him from the scene. “While he’s gone, Annette wants to play with the car some more, so she gets the car out of the toy box to play.” Show Annette taking the car out and playing with it. “Now it’s time for Annette to go eat lunch. Before she goes, *she* puts the car away, but she puts it in this drawer. Then she goes away to eat.” Show Annette placing the car in the drawer, and then remove her from the scene. “Now here comes Eric again. He wants to play with the car some more.” Memory question: “So, *[child’s name]*, did Eric *see* the toy car get moved?”

If the child responds correctly with “no”, say: “That’s right! He didn’t see it get moved!”, and proceed to the next paragraph. If the child responds incorrectly with “yes”, say: “Actually, he didn’t see it get moved. Let’s watch the story again.” Continue on to repeat the task from the beginning, and ask the memory question again. If the child responds incorrectly a second time, say: “No, he didn’t see it get moved, cause he was outside when Annette moved it, remember?” Proceed to the next paragraph.

Test question: “Alright, so when Eric comes back, where will he look for the car?” To be correct, the child must answer that Eric will look in the toy box for his car.

Appearance-Reality

Candle/Orange. Place an object that looks like an orange, but is actually a candle on the table in front of the child. “Look what I have here. What does this look like?” Wait for the child to respond with ‘orange’. “Yeah, it looks like an orange, doesn't it? But really and truly it's a candle, see?” Encourage the child to touch the object. Test question 1: “When you look at this right now, does it look like a candle or an orange?” Wait for the child to respond. Test question 2: “And what is this really and truly, a candle or an orange?" To be correct, the child must respond with "orange" to the first question and "candle" to the second question.

Red/Black Castle. Place a picture of a red castle in front of the child. "Now look at this. This is a red castle. But look, when I put this cover over it, it looks black, see?" Place a transparent, coloured cover over the castle to make it appear black. "So, really and truly it's red." Uncover the castle briefly. "But we can make it look black." Replace the cover. Test question 1: "When you look at the castle right now, does it look red or black?" Test question 2: "And what colour is the castle really and truly, red or black?" To be correct, the child must respond with 'black' to the first question and 'red' to the second question.

Deceptive Pointing

Practice Trial. Place a toy horse and two boxes on the table in front of the child. Bring out a puppet, Samantha. "Now we're gonna play a game with Samantha. See this horse? And see these boxes over here? I'm gonna put the horse inside one of the boxes like this." Place the horse in one of the boxes. "I can point to the box with my finger like this so we'll know which box the horse is in." Point to the box that contains the horse. "Now you try. Take the horse out of the box." Wait for the child to remove the horse. "Now go ahead and put the horse in the other box." Wait for the child to place the horse in the other box. "Ok, now, point to *that* box so we'll know where the car is." Wait for the child to point to the box that contains the horse. "Good job! See, now we can tell where the horse is by pointing to it!"

Test Trial 1. "Now let's put the horse in a box and see if Samantha can find it. But first, Samantha is going to leave and go in here (toy bin to the side of the experimenter) so she can't see." Place Samantha in the toy bin and close the lid. "Okay, go ahead and put the horse in one of the boxes." Wait for the child to place the horse in one of the boxes. "Hey, I have a great idea! Let's play a funny trick on Samantha. Let's play a trick so she can't find the horse. Maybe we could trick her so she'll look in the wrong box, okay? Now, remember, we're gonna play a funny trick on Samantha. We can play a trick by pointing so she *won't* find the horse. Are you ready?" Wait until the child says he/she is ready. "Here comes Samantha!" Remove Samantha from the bin, bring her up to the boxes, and make Samantha ask: "Where's the horse?"

If the child hesitates at first, say: "Samantha is asking where the horse is. Where do you want to point?" If the child still does not respond, say: "Which box do you want to point to?" If the second prompt is also ineffective in inducing the child to respond, say: "Do you want to point to this box (point) or this box (point)?" Wait until the child responds by pointing to one of the boxes. "Ok, Samantha. You can look now!" Make Samantha open the box that the child pointed to. If the child used deception and thus pointed to the empty box, say: "Oh, she didn't find it! I guess we tricked her! We're so tricky! Go ahead and show her where it really is!" If the child failed to use deception and thus pointed to the box that contained the horse, say: "Oh, she found it! I guess we didn't trick her this time." Place Samantha back in the bin. To be correct, the child must point to the empty box, thus conveying a misleading clue about the whereabouts of the toy.

Test Trial 2. Bring out a new puppet, Puppy. "Now it's time for Puppy to come play with us. Let's put the horse in a box and see if Puppy can find it. But first, Puppy is going to leave and go in here (toy bin to the side of the experimenter) so he can't see." Place Puppy in the toy bin and close the lid. "Okay, go ahead and put the horse in one of the boxes." Wait for the child to place the horse in one of the boxes. "Hey, I have a great idea! Let's play a funny trick on Puppy. Let's play a trick so he can't find the horse. Maybe we could trick him so he'll look in the wrong box, okay? Now, remember, we're gonna play a funny trick on Puppy. We can play a trick by pointing so he *won't* find the horse. Are you ready?" Wait until the child says he/she is ready. "Here comes Puppy!" Remove Puppy from the bin, bring him up to the boxes, and make Puppy ask: "Where's the horse?"

If the child hesitates at first, say: "Puppy is asking where the horse is. Where do you want to point?" If the child still does not respond, say: "Which box do you want to point to?" If the second prompt is also ineffective in inducing the child to respond, say: "Do you want to point to this box (point) or this box (point)?" Wait until the child responds by pointing to one of the boxes. "Ok, Puppy. You can look now!" Make Puppy open the box that the child pointed to. If the child used deception and thus pointed to the empty box, say: "Oh, he didn't find it! I guess we tricked him! We're so tricky! Go ahead and show him where it really is!" If the child failed to use deception and thus pointed to the box that contained the horse, say: "Oh, he found it! I guess we didn't trick him this time." Place Puppy back in the bin. To be correct, the child must point to the empty box, thus conveying a misleading clue about the whereabouts of the toy.

"That was fun! You did a great job! Now remember, it's fun to play tricks sometimes, but we should always tell the truth, okay?"

Appendix B

Executive Functioning Task Scripts

Bear-Dragon

"Ok, I'm going to ask you to do some silly things before we start our next game." Proceed with the following list of commands, modeling each action along with the child. "Stick out your tongue. Touch your ears. Touch your teeth. Touch your eyes. Clap your hands. Touch your feet. Touch your head. Touch your tummy. Touch your nose. Wave your hand. Okay, good job!"

"Now I have a game that we can play with these puppets." Bring out two puppets: a friendly-looking bear, and a mean-looking dragon. Place one on each hand, and focus the child's attention on the bear puppet. "This puppet is a nice bear. When he talks to us, we'll do what he tells us to do." Focus the child's attention on the dragon puppet. "This puppet isn't very nice. This puppet is a dragon. When he talks to us, we won't listen to him. If he tells us to do something, we won't do it. Okay, let's practice one time." Bring the bear towards the child. "This is the good bear. He says, '*Touch your nose'*." Use a mellow, nice voice to say the bear's instructions. If the child does not touch his/her nose, say: "Remember, we listen to the nice bear and do what he tells us to do because that's how we play the game." Continue to repeat the instructions and model the action if necessary until the child succeeds. Wait until the child touches his/her nose. "That's right! Now let's practice with the naughty dragon. In this game, we won't do what the dragon tells us to do because he's not so nice. '*Touch your tummy'*." Use a low, gruff voice to say the dragon's instructions. If the child toches his/her stomach, say: "Remember, we're not going to listen to the mean dragon. We won't do what he tells us to do because that's how we play the game." Repeat the instructions until the child successfully refrains from touching his/her stomach. If necessary, hold the child's hands down on the sixth try. Wait until the child succeeds. "Yeah! Good job! That was fun! So, when the bear tells you to do something, do you do it?" Wait for the child to respond and correct if necessary. "And when the dragon tells you to do something, do you do it?" Wait for the child to respond and correct if necessary. "Okay, let's play some more!"

On the following test trials, alternate between using the dragon and bear voices. Whenever the dragon provides the instruction, bring the dragon slightly towards the child. Likewise, whenever the bear provides the instruction, bring the bear slightly towards the child. Bear: "*Stick out your tongue*." Dragon: "*Touch your ears*." Bear: "*Touch your teeth*." Dragon: "*Touch your eyes*." Bear: "*Clap your hands*." Provide the following reminder, regardless of performance: "Remember the way we play this game. We do what the bear tells us to do because he's nice, but we won't do what the dragon tells us to do because he's not so nice." Dragon: "*Touch your feet*." Bear: "*Touch your head*." Dragon: "*Touch your tummy*." Bear: "*Touch your nose.*" Dragon: "*Wave your hand*."

On each of the 5 bear test trials, code children's behavioural responses as follows: 0 = failure to move, 1 = a wrong movement, 2 = a partial correct movement, 3 = a full correct movement. On each of the 5 dragon test trials, code children's behavioural responses as follows: 0 = full correct movement, 1 = partial correct movement, 2 = incorrect movement (e.g., touches feet when instructed to touch head), 3 = strategic movement (e.g., shakes head no, moves hands behind back), and 4 = no movement.

Grass/Snow

On the table in front of the child, place a board that has a solid whie card attached to the upper left corner, a solid green card attached to the upper right corner (both cards are 15 x 10 cm), and two fabric cut-outs shaped like a child's hands centered below the cards. Place the child's hands on top of the felt child-sized hands on the board. "Now we're going to play a game with this board. Do you know what colour grass is?" Wait for the child to respond with 'green', and correct if necessary. "That's right! And do you know what colour snow is?" Wait for the child to respond with 'white', and correct if necessary. "Very good. We're going to play a silly game. In this game, when I say the word 'grass', I want you to point with your finger to the while card, like this." Demonstrate pointing to the white card. "Can you point to the white card?" Praise the child if he/she pointed correctly, prompt if not. "And when I say the word 'snow', I want you to point with your finger to the green card like this." Demonstrate pointing to the green card. "Can you point to the green card?" Praise the child if he/she points correctly, prompt if not.

Begin the practice trials. "Grass." Wait for the child to respond. If the child hesitates, say: "What card do you want to point to for this one?" Praise the child if he/she responds correctly. If he/she is incorrect, say: "Remember, this is a silly game. When I say 'grass', I want you to point over here to the white card. When I say snow, *that's* when you point to the green card. Grass." Continue reiterating the instructions until the child responds correctly. "Very good! Snow." Wait for the child to respond.

If the child hesitates, say: “What card do you want to point to for this one?” Praise the child if he/she responds correctly. If he/she responds incorrectly, say: “Remember, this is a silly game. When I say ‘snow’, I want you to point over here to the green card. When I say grass, *that’s* when you point to the white card. Snow.” Continue with the practice trials until the child correctly responds to both trials consecutively.

Begin the test trials. If the child hesitates, say: “What card do you point to for this one?” Do not give feedback on the test trials. The order of the 16 test trials is as follows: Grass, Snow, Snow, Grass, Snow, Grass, Grass, Snow, Snow, Grass, Snow, Grass, Grass, Snow, Grass, Snow. On each trial, the child will receive a score of 1 if he/she initially points to the correct card, and a score of 0 if he/she initially points to the incorrect card, or to both cards at once.

Card Sort

Place 2 boxes on the table in front of the child. The box to the left has a red rabbit card pasted on it, while the box to the right has a blue boat card pasted on it. “Now we’re going to play a game. This is the shape game. All the rabbits go in this box.” Point to the left red rabbit box. “And all the boats go in that box.” Point to the right blue boat box. “We don’t put any rabbits in that box. No way! We put all the rabbits over here.” Point to the left red rabbit box. “And only boats go over there.” Point to the right blue boat box. “Okay? So if it’s a rabbit, then it goes here.” Point to the left red rabbit box. “And if it’s a boat, then it goes there.” Point to the right blue boat box. “This is the shape game. I’ll go first. Rabbits go here.” Place a blue rabbit card upside down in the left red rabbit box. “And boats go here.” Place a red boat card upside down in the right blue boat box.

Pre-Switch Trials. "Now it's your turn. If it's a rabbit, then put it here." Point to the left red rabbit box. "But if it's a boat, then put it there." Point to the right blue boat box. "Here is a *blue rabbit*. Where does this go?" Wait for the child to place the card in one of the boxes. If the child is correct, say: "That's right! Very good!" If the child is incorrect, say: "No, that's not right. Remember the rules." Proceed to the next trial.

"If it's a rabbit, then put it here." Point to the left red rabbit box. "But if it's a boat, then put it there." Point to the right blue boat box. "Here's a *red boat*. Where does this go?" Wait for the child to provide an answer, and respond as described above. Proceed to the next trial.

"If it's a rabbit, then put it here." Point to the left red rabbit box. "But if it's a boat, then put it there." Point to the right blue boat box. "Here's a *blue boat*. Where does this go?" Wait for the child to provide an answer, and respond as described above. Proceed to the next trial.

"If it's a rabbit, then put it here." Point to the left red rabbit box. "But if it's a boat, then put it there." Point to the right blue boat box. "Here's a *red rabbit*. Where does this go?" Wait for the child to provide an answer, and respond as described above. Proceed to the next trial.

"If it's a rabbit, then put it here." Point to the left red rabbit box. "But if it's a boat, then put it there." Point to the right blue boat box. "Here's a *blue rabbit*. Where does this go?" Wait for the child to provide an answer, and respond as described above. Proceed to the post-switch trials only after the child has responded correctly to 5 consecutive pre-switch trials.

Post-Switch Trials. "Now we are going to switch. We are not going to play the shape game anymore. We are going to play the *colour* game. When it is red, you have to put it in this box." Point to the left red rabbit box. "But whenever it is blue, then it goes in that box." Point to the right blue boat box. "We don't put red things in that box. No way. We put red things over here." Point to the left red rabbit box. "And only when it's blue does it go over there." Point to the right blue boat box. "If it's blue, then it goes there." Point to the right blue boat box. "If it's red, then it goes here." Point to the left red rabbit box.

"Okay, so remember - if it is red, then put it here," Point to the left red rabbit box. "But if it is blue, put it there." Point to the right blue boat box. "Here is a *red boat*. Where does this go?" Wait for the child's response, and do not provide them with feedback. "Okay, let's do another!"

"If it is red, then put it here," Point to the left red rabbit box. "But if it is blue, put it there." Point to the right blue boat box. "Here is a *red rabbit.* Where does this go?" Wait for the child's response, and do not provide them with feedback. "Okay, let's do another!"

"If it is red, then put it here," Point to the left red rabbit box. "But if it is blue, put it there." Point to the right blue boat box. "Here is a *blue boat.* Where does this go?" Wait for the child's response, and do not provide them with feedback. "Okay, let's do another!"

"If it is red, then put it here," Point to the left red rabbit box. "But if it is blue, put it there." Point to the right blue boat box. "Here is a *red boat.* Where does this go?" Wait for the child's response, and do not provide them with feedback. "Okay, let's do another!"

"If it is red, then put it here," Point to the left red rabbit box. "But if it is blue, put it there." Point to the right blue boat box. "Here is a *blue rabbit*. Where does this go?" Wait for the child's response, and do not provide them with feedback.

"Okay, that was fun! Good job on that game!" The child receives a score of 1 for every post-switch trial that he/she answers correctly (maximum total = 5).

Whisper

"For this next game, I need to see if you can whisper." The experimenter now starts to whisper, and continues to do so for the duration of the task. "Can you whisper your name to me?" Wait for the child to respond. Praise the child if they successfully whisper their name, and proceed to the following paragraph. If they are unsuccessful, say: "Let's try to talk very, very quietly, just like I am. Can you whisper your name to me?" Wait for the child to respond, and continue even if the child does not whisper.

"Ok, let's play this whisper game I know. I have some pictures of some cartoon characters. Let's see if you can tell me their names. Some of them I bet you'll know and some might be kind of hard and you won't know them, and that's okay. Just remember to whisper because that's how we play the game." Proceed to show the child a series of cards that depict cartoon characters. Before each card is placed on the table in front of the child, say: "Do you know who this is?" The first 5 characters shown are: Dora, Blue, Huckle, Bob the Builder, and Elmer Fudd. After these first 5 trials, provide the following prompt, regardless of performance: "You are doing really well. Remember that this is a whisper game where we try to whisper the whole time." The next 5 characters shown are: Winnie the Pooh, Petunia, Big Bird, Fat Albert, and Mickey Mouse.

Across all trials, if the child responds by simply nodding his/her head to indicate that he/she knows the character's name, prompt them with: "Can you tell me his/her name?" In instances where the child says that he/she does not know the name of the character, say: "That's okay. This is a hard one! His/her name is *[character's name]*. Let's try another one." In the event that the child provides a name that is incorrect, say "okay" and move on to the next picture.

KRISP

Practice Trials. Throughout this task, children are sequentially shown target pictures and asked to identify an exact match to the target from an array of four to six similar pictures. Begin by placing a binder that holds the KRISP pictures on the table in front of the child, and open the book to the first page. "Now we have another game to play. Do you see this picture at the top?" Point to the target picture. "Can you find the one picture down here (indicate entire array on lower page) that looks just like this one up here?" Point to the target picture, and wait for the child to respond. If the child is correct, say: "Very good. Let's do that every time. Always point to the one down here (indicate) that is just exactly like the one up here (indicate)." If the child is incorrect, provide him/her with feedback, explaining why the pictures are different. For example: "No, that picture isn't *exactly* the same as the one up here, because the one up here is a *red* circle, and this one down here is a *yellow* circle. Now point to the one that is *exactly* like this one up here." Continue with corrective feedback until the child succeeds in identifying the exact match. Turn to the next page, which is blank.

"Okay, when I turn the page, point to the picture down here (indicate) that is exactly the same as the one you will see up here (indicate). Ready?" Wait for the child to respond, and then turn the page. "Which one down here (point to the array of similar pictures) looks just like this one (point to the target picture)? Can you point to it the first time?" Reinforce the correct response and provide feedback for incorrect responses as with the first practice item. Turn to the next page, which is blank.

"Okay, when I turn the page, point to the picture down here (indicate) that is exactly the same as the one you will see up here (indicate). Are you ready?" Wait for the child to respond, and then turn the page. "Which one down here (point to the array of similar pictures) looks just like this one (point to the target picture)? Can you point to it the first time?" Reinforce the correct response and provide feedback for incorrect responses as with the first practice item. Turn to the next page, which is blank.

"Remember, only one is exactly the same. Always try to find it the first time. Are you ready?" Wait for the child to respond, and then turn the page. Say nothing until the child makes his/her first response. Reinforce the correct response and provide feedback for incorrect responses as with the first practice item. Turn to the next page, which is blank.

"Remember, only one is exactly the same. Always try to find it the first time. Are you ready?" Wait for the child to respond, and then turn the page. Say nothing until the child makes his/her first response. Reinforce the correct response and provide feedback for incorrect responses as with the first practice item. Turn to the next page, which is blank. “Are you ready for the next one?” Wait for the child to respond, and then turn the page.

Test Trials. On all subsequent pages (i.e., test trials) the experimenter gives the child brief and varied reinforcement for correct responses, without providing reminders or instructions. Turn the page immediately after correct responses. When the child is incorrect, say: “No, look up here. Can you find the one that’s *exactly* like this one up here?” After 3 successive wrong choices, say: “Okay, let’s go on to the next one. Remember, only one is exactly the same as the one up here. (point to the target picture) Can you find it the first time?” Turn the page immediately.

In addition to clearly correct and incorrect answers, there are a number of additional potential responses that require specific feedback from the experimenter. If the child responds that none of the items match, say: “Yes, there is *one* that is exactly the same. Keep looking and see if you can find it.” If the child responds that all of the pictures are the same, say: “No, some of them are different. Only one is exactly the same. Can you find it?” If the child says that he/she does not know which option to choose, say: “Keep looking. Try to find the one that’s the same as the one on top.” If the child responds by pointing to an option that they indicate to be *incorrect*, say: “Just point to the one that is exactly the same. Try to find it the first time.” Finally, if the child points rapidly to more than one alternative, stop the clock and interrupt to ask: “Which one did you point to first? Just point to the one that is the same, but don’t point to the others.” On each trial, the experimenter must record 1) the time to first response (regardless of whether it is correct) 2) all choices made by the child, and 3) the number of errors made.

Appendix C

Sample False Belief Training Task Script

On the table in front of the child, place a miniature bush, a miniature log, a rabbit, and two characters, Sonia and Eddie. "This girl's name is Sonia, and this is her brother Eddie." Show the child the two dolls. "Sonia and Eddie have a pet rabbit. They like to play with the rabbit in their back yard. One day while they're playing, their rabbit hops into the bush." Show the rabbit moving into the bush until it is out of sight. "Then Eddie goes inside to get a snack." Move Eddie off the table, out of sight. "While he's inside, Sonia wants to play with the rabbit some more, so she gets the rabbit out of the bush." Show Sonia retrieving the rabbit from the bush. "And then they play some more. When she's done, Sonia puts the rabbit away, but she puts it in the log." Show Sonia placing the rabbit inside the log, out of sight. "Then Sonia goes away to play in the park." Move Sonia off the table, out of sight. "A little while later, Eddie comes back outside and he wants to play with the rabbit some more.
So, *[child's name],* where will Eddie look for the rabbit?" Wait for response. During the second testing session, the experimenter additionally asks: "Why will he look there?"

If the child chooses the correct location, provide the following feedback: "That's right. Eddie *does* look in the bush for his rabbit." Show Eddie looking in the bush for his rabbit. "Eddie was gone when Sonia took the rabbit and put it in the log, so Eddie didn't see the rabbit get moved! When Eddie comes back, he'll look in the bush for his rabbit, cause that's where it was when he left!"

If the child chooses the incorrect location, place all the objects in their original positions and provide the following feedback: “Well, that’s a really good guess, but that’s not right! When Eddie comes back, he’ll look in the bush for his rabbit. Let’s watch again and see why. Eddie saw the rabbit go into the bush.” Place the rabbit back in the bush. “Then he left!” Move Eddie off the table, out of sight. “While he was gone, Sonia got the rabbit and put it in the log.” Show Sonia retrieving the rabbit, and placing it in the log. “Eddie was gone, so he didn’t see that! He didn’t *see* Sonia put the rabbit in the log. So, when Eddie comes back, he’ll look in the bush for the rabbit, cause that’s where it was when he left!” Show Eddie looking in the bush to retrieve the rabbit.

Appendix D

Theory of Mind Scale Task Scripts

Diverse Desires

The child is shown a stuffed animal (Big Bird) and a sheet of paper with an apple and a cookie depicted on it. “Here is Big Bird. It’s snack time, and Big Bird wants a snack to eat. Here are two different snacks: an apple and a cookie. Which snack would you like best? Would you like the apple or the cookie best?”

If the child chooses the apple**,** provide the following response: “Well, that’s a good choice, but Big Bird really likes cookies. He doesn’t like apples. What he likes best are cookies.” If the child chooses the cookie, provide the following response: “Well, that’s a good choice, but Big Bird really likes apples. He doesn’t like cookies. What he likes best are apples.”

“So, now it’s time to eat. Big Bird can only choose one snack, just one. Which snack will Big Bird choose? The apple or the cookie?” To pass the task, the child must respond to the target question with the food item that is in accordance with Big Bird’s desires.

Diverse Beliefs

The child is shown Cookie Monster and a sheet of paper with a garage and a tree depicted on it. “Here’s Cookie Monster. Cookie Monster wants to find his cat. His cat might be hiding in the tree, or it might be hiding in the garage. Where do you think that cat is? In the tree or the garage?”

If the child chooses the tree**,** provide the following response: “Well, that’s a good idea, but Cookie Monster thinks his cat is in the garage. He thinks his cat is in the garage." If the child chooses the garage, provide the following response: "Well, that is a good idea, but Cookie Monster thinks his cat is in the tree. He thinks his cat is in the tree."

"So where will Cookie Monster look for his cat? In the tree or in the garage?" To pass the task, the child must respond to the target question with the location that is in accordance with Cookie Monster's beliefs.

Knowledge Access

Show the child a jewelry box with a toy frog inside a closed drawer. "Here's a drawer. What do you think is inside the drawer?" Open the drawer to show the child its contents. "Let's see. Oh, it's a frog!" Proceed to close the drawer. "Okay, so what's in the drawer?" Wait for the child's response. If he/she responds incorrectly, repeat the task from the beginning until he/she is able to provide the correct answer.

Place a stuffed animal (Kitty) on the table in front of the child. "This is Kitty. Kitty has never ever seen inside the drawer." Test question: "So, does Kitty know what is in the drawer?" Memory question: "Did Kitty see inside the drawer?" To be correct, the child must answer "No" to the test question and "No" to the memory question.

Contents Change

Refer to Appendix A for a sample Contents Change script.

Real-Apparent Emotion

Show the child a sheet of paper with a happy face, a sad face, and a neutral face depicted on it. Check to make sure the child is familiar with each type of face. Then present the child with a second picture of a boy who is depicted so that his facial expression cannot be seen. “This story is about a boy. I’m going to ask you how the boy really feels inside, and how he looks on his face. He might really feel one way inside but look a different way on his face. Or, he might really feel the same way inside as he looks on his face. So remember, I want you to tell me how he really feels, and how he looks on his face. This story is about Matt. One day, Matt’s friends were playing together and telling jokes. One of the older children, Rosie, told a mean joke about Matt and everyone laughed. Everyone thought it was very funny, but not Matt. But, Matt didn’t want the other children to see how he felt about the joke because they would call him a baby, so Matt tried to hide how he felt.”

Memory questions: “What happened when Rosie told a mean joke about Matt? What did the other children do?” Wait for the child to respond. The correct response is that the children laughed and/or thought it was funny. “In the story, what would the other children do if they know how Matt felt?” Wait for the child to respond. The correct response is that Matt thought they would call him a baby or tease him. Test question 1: “How did Matt *really* feel when everyone laughed? Did he feel happy, sad or okay?” Wait for the child to respond. Test question 2: “How did Matt try to look on his face when everyone laughed? Did he look happy, sad or okay?” Wait for the child to respond. To be correct the child must answer the memory questions correctly, in addition to answering the target ‘feel’ question more negatively than the target ‘look’ question.