

**Theory of Mind and Perspective-taking in Young Children**

by

Hannah Mohun  
B.Sc. (Hons), University of Victoria, 2014

A Thesis Submitted in Partial Fulfillment  
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Psychology

© Hannah Mohun, 2017  
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

**Supervisory Committee**

Theory of Mind and Perspective-taking in Young Children

by

Hannah Mohun

B.Sc. (Hons), University of Victoria, 2014

**Supervisory Committee**

Dr. Ulrich Müller, Supervisor  
Department of Psychology

Dr. Mauricio A. Garcia-Barrera, Departmental Member  
Department of Psychology

## Abstract

### Supervisory Committee

Dr. Ulrich Müller, Supervisor  
Department of Psychology

Dr. Mauricio A. Garcia-Barrera, Departmental Member  
Department of Psychology

Based on past research that suggests an association between perspective-taking and theory of mind, the present study investigates the relationship between these two constructs by administering three perspective-taking and three theory of mind tasks, along with one measure of inhibition and one measure of expressive language, to eighty-four 3- and 4-year olds. The goals of this study were to (1) evaluate the relative difficulties of the six perspective-taking and theory of mind tasks, (2) assess whether the theory of mind and perspective-taking tasks scale onto the same dimension, (3) examine the distinctions proposed within the construct of perspective-taking, (4) investigate the relationship between inhibition and verbal ability with theory of mind and perspective-taking, and (5) examine the relationship between perspective-taking and theory of mind. The results show that the six theory of mind and perspective-taking tasks develop along a single continuum and produce a scale of the understanding of subjectivity for preschoolers. Results also demonstrate that the ability to take someone else's perspective without directly confronting it with one's own perspective is the least challenging, and understanding that others' beliefs can be false is the most challenging task. Finally, results indicate that inhibition and verbal ability account for the relationship between almost all of the perspective-taking and theory of mind tasks.

## Table of Contents

Supervisory Committee .....	ii
Abstract .....	iii
Table of Contents .....	iv
List of Tables .....	vi
List of Figures .....	vii
Acknowledgements .....	viii
Introduction .....	1
Development of Perspective-taking .....	1
Development of Theory of Mind .....	6
Relationship between ToM and PT .....	11
Current Study .....	16
Hypotheses .....	21
Methods .....	23
Participants .....	23
Measures .....	23
Inhibition Task .....	23
Grass/Snow task .....	23
Verbal Ability Task .....	24
K-ABC .....	24
Theory of Mind Tasks .....	24
Diverse desire task .....	24
Diverse belief task .....	25
Explicit false belief task .....	25
Perspective-taking Tasks .....	25
Turtle task .....	25
Perspective-taking task .....	26
Perspective-confronting task .....	26
Procedures .....	26
Daycares .....	27
Child Development Lab .....	27
Results .....	29
Data Preparation .....	29
Descriptive Statistics and Correlations .....	30
Correlations with expressive vocabulary subtest partialled out .....	32
Correlations with grass/snow task partialled out .....	32
Correlations with expressive vocabulary subtest and grass/snow task partialled out .....	33
Linear and Logistic Regressions .....	36
Item Difficulty .....	39
Rasch analysis .....	41
Guttman scaling .....	46

Discussion .....	48
Task Difficulty .....	48
Subjectivity Scale.....	50
Inhibition and Verbal ability.....	52
Relationship between PT and ToM.....	53
Distinctions Among Constructs .....	53
Limitations of the Current Study .....	56
Future Directions .....	57
Conclusion .....	59
References.....	61
Appendix A.....	71
Appendix B .....	72
Appendix C .....	73
Appendix D.....	74
Appendix E .....	75
Appendix F.....	77
Appendix G.....	79
Appendix H.....	80

## List of Tables

<i>Table 1.</i> Mean, standard deviation and range of 8 tasks and age in months, with control questions included.....	31
<i>Table 2.</i> Mean, standard deviation and range of 3 ToM tasks, with control questions excluded. 31	
<i>Table 3.</i> Zero-order and partial correlations among PT tasks, ToM tasks, and K-ABC, control questions included.....	34
<i>Table 4.</i> Zero-order and partial correlations among PT, ToM tasks, and grass/snow, control questions included.....	35
<i>Table 5.</i> Zero-order and partial correlations among PT, ToM tasks, K-ABC and grass/snow task, control questions included .....	36
<i>Table 6.</i> Linear regressions .....	37
<i>Table 7.</i> Logistic binary regressions .....	38
<i>Table 8.</i> Proportion of children correct on each task, control questions included.....	39
<i>Table 9.</i> Proportion of children correct on each task, with control questions excluded.....	40
<i>Table 10.</i> Mean, standard deviation and range of tasks converted to dichotomous scores .....	41
<i>Table 11.</i> Item and Person Summary and Fit Statistics for the Rasch Model with control questions included.....	44
<i>Table 12.</i> Item and Person Summary and Fit Statistics for the Rasch Model with control questions excluded.....	45
<i>Table 13.</i> Guttman scalogram for six tasks.....	47

## List of Figures

<i>Figure 1.</i> Parameter estimates and error values for the Rasch model with control questions included.....	46
<i>Figure 2.</i> Parameter estimates and error values for the Rasch model with control questions excluded.....	46

## **Acknowledgements**

First, I would like to thank my supervisor, Dr. Ulrich Müller, for his guidance, support and invaluable feedback, without which this thesis would not have been possible. I would also like to thank Dr. Mauricio Garcia-Barrera for giving his time and thoughtful contributions and suggestions to this project.

I am grateful to the members of the Child Development Lab, whose time and effort to schedule appointments, test children, and enter data is incredibly appreciated. Additionally, thank you to all of the staff, parents and children from the daycares that participated in this study.

Finally, I would like to thank my family and friends for being the best support network out there and for always encouraging and uplifting me.

## **Introduction**

Social skills help children foster positive relationships and understand the behaviours of others. Social competencies also enhance a child's experience in school and have been shown to improve mental health (Cook et al., 2015). For that reason, it is important to learn more about factors that predict social skills. Many aspects of social understanding have been suggested as predictors of social skills, and this paper will focus on two of those: theory of mind (ToM) and perspective-taking (PT). The goal of this paper is to provide an overview of ToM and PT, examine the relative difficulties of several ToM and PT tasks, and evaluate whether or not the two constructs develop along a single continuum.

### **Development of Perspective-taking**

PT is the ability to view a situation or understand a concept from an alternate point-of-view. The ability to take someone else's perspective changes and develops throughout the lifespan, but goes through major changes during the preschool years. The development of PT occurs in a progressive manner, with children acquiring increasingly more difficult PT skills over time. During the first five years of life, joint attention constitutes the basis for the development of PT. Infants learn to jointly attend to objects and events with others, resulting in the sharing of a perception or experience (Moll & Meltzoff, 2011b). Although joint attention begins early in development, infants lack basic PT skills. Infants are unable to understand that a barrier can be used to block an individual's visual perspective (Flavell, Shipstead and Croft, 1978). Flavell and Markman have suggested that infants' inability to take another's perspective stems from a mutual exclusivity rule, or the single constraint hypothesis (see, for a discussion, Perner, Stummer, Sprung, & Doherty, 2002). Infants are not able to see others' perspectives because they believe that each object only has one status and cannot be viewed in any alternative ways.

In the development of PT, different levels have been distinguished. Flavell (1992) divided PT into 2 levels. Level 1 PT is the understanding that others can see different things than you do. Level 1 PT is assessed using percept production, percept deprivation, or percept diagnosis tasks. Percept production tasks require children to intentionally produce a visual percept by actions such as displaying an image for an individual, or by pointing to get the individual to orient towards an object. Percept deprivation tasks require children to hide or block a visual percept by actions such as moving the object or setting up a barricade. Percept diagnosis tasks require children to diagnose an individual's visual experience by demonstrating that they understand what she can and cannot see. One study by Flavell, Shipstead and Croft (1978) assessed percept deprivation and found that children aged 2.5 to 3.5 years were able to correctly move an object so that the other individual wasn't able to see it. These findings demonstrate that young children understand that even though they can see an object, it doesn't mean that the other person can see it as well. In another study that assessed percept production, Liskowski, Carpenter, Striano, and Tomasello (2006) found that infants in their second year of life had level 1 PT abilities. Liskowski et al.'s study showed that 12- and 18-month old infants pointed in order to direct an adult to an object they were looking for. This finding suggests that the infants could take the adult's perspective and understood that the adult could not see something that the infant could.

Level 2 PT is the understanding that not only might others see different things, they might also see the same things in a different way (Flavell, 1992). Level 2 PT is often measured using the turtle task (Masangkay et al., 1978). An image of a turtle is placed between the children and the experimenter. The children are asked how they see the turtle and also how the experimenter sees the turtle (upside down or right side up), forcing them to verbally confront the

differing perspectives. Masangkay et al. found that children aged 4.5 years and older were able to correctly identify that the experimenter saw the turtle differently than they did.

Moll, Meltzoff, Merzsch and Tomasello (2013) suggest a more fine-grained distinction within level 2 PT tasks. Specifically, they introduce the distinction between taking and confronting perspectives. The taking of perspectives was required in a study by Moll and Meltzoff (2011a) in which 3-year olds understood how an adult saw two different objects, demonstrating the ability to take the adult's visual perspective. Moll and Meltzoff argue that the 3-year olds were successful at taking someone else's perspective because they weren't required to confront the adult's perspective with their own. In other words, the children were only asked to acknowledge the adult's perspective, and weren't required to recognize both their own and the adult's perspective.

Taking someone else's perspective requires an individual to identify what another person can see, whereas perspective-confronting requires an individual to not only identify what another person can see, but also understand how it conflicts with what the individual him- or herself can see. Perspective-confronting involves the ability to explicitly acknowledge the differing views of yourself and another. Moll *et al.* (2013) found that perspective-confronting was more difficult for children than only taking someone else's perspective.

Perner et al. (2002) also argued that different levels of PT need to be distinguished. Perner et al. believe there is a difference between switching perspectives and confronting perspectives. Switching perspectives occurs when you take different perspectives at different times, and confronting perspectives occurs when you represent two perspectives simultaneously (Perner et al.). For example, once children learn that an apple is a fruit, they can switch from considering it an apple, to considering it a fruit. Confronting perspectives requires children to

understand that it can be considered an apple as well as a fruit. The idea of switching perspectives appears similar to Moll et al.'s (2013) description of perspective-taking. Both definitions require children to be able to take one perspective at a single time. Confronting perspectives proposed by Perner et al. and Moll et al. also appear to be analogous, but there are some discrepancies between the two descriptions. Moll et al. specify that the different perspectives must be your own and another's (between-person PT), but Perner et al. suggest that the different perspectives can both be your own (within-person PT; e.g. understanding that an object is both an apple and a fruit; understanding that you feel sad now but will feel happy at some point later). Moll et al.'s definition of perspective-confronting also requires that the individual overtly acknowledges the differing perspectives (e.g. verbally or by using gestures), whereas Perner et al. don't specify such a requirement (e.g. could be mental confrontation). Moll et al. (2013) and Perner et al. (2002) agree that confronting perspectives is more challenging than taking a single perspective, and both claim that only children around 4- and 4.5-years old are able to successfully confront perspectives.

Even though the focus of the current paper is on the development of PT in preschoolers, it should be noted that there is considerable development in PT after the preschool period. One famous task that has often been used to capture these changes is Piaget and Inhelder's (1956) 3-mountain task. In this spatial PT task the child sits in front of a model of three mountains that are different sizes and have different identifiers (one mountain has snow, one has a red cross on top, one has a hut on top). The child is then shown photographs and is asked to select which photograph best reflects his/her own view of the mountain. Next, a doll is placed at different vantage points relative to the child and the child is shown 10 photographs. The child is asked to select which of the 10 photographs best reflects the doll's view. Since children are required to

choose photos that reflect both their own and the doll's perspective, one could argue that this is measuring perspective-confronting. Piaget and Inhelder found that children were not able to successfully complete this task until about 9 years old.

However, Borke (1975) tested children on similar spatial PT tasks and found that 3- and 4-year old children were generally able to see displays from someone else's perspective. Borke argued that it was the nature of the displays she used that allowed younger children to succeed. The scenes that Borke used included more differentiated objects that children could use as cues to understand how the scene was oriented for someone else. Borke found that when she used a replica of Piaget and Inhelder's 3-mountain task, the 3- and 4-year olds who were successful with her scenes found this display more challenging. Borke suggested that it was the lack of objects in the 3-mountain task that made it more difficult for children to succeed. It is relevant to note that Borke's study did not require children to demonstrate knowledge of their own perspective, suggesting that this task only required perspective-taking as opposed to perspective-confronting, which perhaps played a role in making the task less challenging than Piaget and Inhelder's (1956).

Borke (1971) also argued that young children are often unsuccessful at PT tasks because of the limitation in their response capabilities, which affects their ability to demonstrate understanding. Borke found that children as young as 3-years old could accurately anticipate the emotional responses of others when they were asked to respond behaviourally rather than verbally. Borke concluded that children are able to take another's point of view much earlier than Piaget had suggested, and that it is cognitive capabilities such as verbal responses that impeded children's success in other PT tasks. Chandler and Greenspan (1972) argued that although Borke's study accurately demonstrated how young children are aware of others' emotions, the

ability to accurately take someone else's perspective, which is the more common understanding of PT, is a later development and was not captured by Borke's task.

The development of PT continues to extend into later childhood and adolescence. Selman's (1973) developmental theory of PT goes beyond the objective-subjective distinction required to take someone's perspective that emerges during the preschool years. Selman's theory includes the later developing ability to adopt a third-person perspective and the ability to understand that taking someone's perspective is vulnerable to misunderstandings due to the complexity of individuals' psychological states (Martin, Sokol, & Elfers, 2008).

To summarize, extant empirical evidence suggests that PT tasks vary in difficulty depending on the specific ability required. PT tasks include a wide range of skills such as being able to alter someone else's perspective, inhibit your own perspective, verbally confront two different perspectives, or understand what someone else is able to see. Differentiations have been made within the ability of PT (Moll et al., 2013; Perner et al., 2002), but more research needs to be conducted to evaluate whether these refinements are valid and are identifying independent skills. In addition, research by Piaget and Inhelder (1956) and Borke (1971; 1975) emphasizes how task methodology plays a large role in the difficulty level of the task and can alter children's performance.

### **Development of Theory of Mind**

ToM is the ability to attribute mental states such as beliefs, desires, knowledge, and intents to other people (Wellman, Cross, & Watson, 2001). ToM as a construct was first mentioned in the 1978 issue of *Behavioural and Brain Sciences* by Premack and Woodruff, who raised the question of whether chimpanzees had what they coined a 'theory of mind'. False belief understanding is considered to be a hallmark in the development of ToM. Successfully

completing a false belief task requires children to understand that others can have a false belief, or a belief that doesn't match reality. Wimmer and Perner (1983) created the first false belief task, called the Maxi task, which was later adapted by Baron-Cohen, Leslie, and Frith (1985), and called the Sally-Anne task. In this task, children are told or shown a story about two characters, Sally and Anne. Sally has a basket and Anne has a box. Sally has a marble and puts the marble into her basket, and then leaves the room. While she is gone, Anne moves the marble into the box. Sally returns, and children are asked where Sally will look for the marble. Children pass the task if they say Sally will look in the basket, and fail the task if they say that Sally will look in the box. The Sally-Anne task assesses whether children understand that someone else's belief can be different from reality, and whether they can predict their behaviour based on that understanding.

In the 1980's and 1990's, a number of studies examined false belief understanding in preschool children and found inconsistent results in regards to the age at which children are able to pass false belief tasks. To investigate these empirical inconsistencies, Wellman et al. (2001) conducted a meta-analysis of 178 studies that used false belief tasks. The meta-analysis demonstrated a clear developmental progression of false belief understanding during the preschool years, with conceptual changes occurring between 2.5- and 5-years old, though the majority of studies have found that false belief understanding develops around age 4 (Gallagher, 2015b). Wellman et al.'s results were inconsistent with the theory that developmental changes in false belief understanding are due to task artifacts. Although the meta-analysis revealed four task manipulations that improved children's performance, since both younger and older children found these manipulations helpful and no manipulations boosted performance to above-chance, the conceptual developmental changes are still relevant above and beyond those task

manipulations. In fact, the meta-analysis found that many potentially relevant task manipulations (e.g. type of task, type of question) had no effect on children's performance, suggesting that different false belief tasks can be equitably compared.

The studies that found major changes in false belief understanding around the age of 4 have used explicit (verbal) measures of false belief (Gallagher, 2015b). More recently, implicit, looking time measures of false belief have been introduced. Based on the violation of expectation paradigm, these implicit tasks use infant's looking time to evaluate whether the actions of an agent were in accordance with the infant's expectations or not. For example, Onishi and Baillargeon (2005) found that infants looked longer when an agent who had a false belief about the location of a toy looked in the location where the toy was. Onishi and Baillargeon suggest that these findings indicate that the infant expected the agent to look where he/she thought the toy was hidden, and not where it was actually hidden. Studies using implicit measures show that much younger children (infants as young as 6 months of age) already understand false belief. However, this is a controversial issue that will not be further discussed in this paper (see Buttelmann, Carpenter & Tomasello, 2009; Onishi & Baillargeon; Perner, 2010).

The focus on false belief tasks as a measure of ToM has not been without criticism. It is generally assumed that if children can understand that someone can have a false belief, then they understand that others have mental states; however, this assumption is not necessarily accurate. Dennett (1978) differentiates between having expectations about someone else's behaviour because you have a belief about their belief, versus having expectations about someone else's behaviour based on a habit or a belief about some other aspect of the world. If a child completes a false belief task because of the latter reason, there is not necessarily an understanding that others have mental states. More recently, Bloom and German (2000) argue that the false belief

task is not an accurate measure of ToM for two reasons: there is more required to pass a false belief task than just ToM, and there is more to ToM than just passing the false belief task. Finally, the focus on false belief tasks leads people to overlook that false belief understanding does not come out of nowhere and that ToM develops before and continues to develop after false belief understanding.

Wellman et al. (2001) argue that false belief tasks should not be used in isolation but should be used alongside other measures of mental state understanding and should be interpreted in the context of a developmental sequence. For that purpose, Wellman and Liu (2004) used five different tasks to assess various aspects of ToM and false belief understanding. They found a consistent developmental progression demonstrating that some aspects of ToM are reliably more challenging than others. In addition, their findings suggest that there is progressive conceptual development underlying ToM understanding. Flavell (1992) predicted this finding when he stated that ToM abilities “emerge in fixed sequence, suggesting an orderly developmental succession of such insights” (p. 127).

I will elaborate on some of the tasks below to demonstrate the range of abilities measured in Wellman and Liu’s study. I will discuss them in the order of the tasks that were generally found easiest, to the tasks that were generally found to be hardest for children. The series of tasks assessed five different types of understanding: diverse desire, diverse belief, knowledge access, false belief, and hidden versus apparent emotions. Tasks involved used comparable formats, similar materials and questioning style so that differences in performance could be attributed to conceptual differences rather than less pertinent task-related differences.

Diverse desire/belief tasks require children to correctly identify their own versus others’ desires and beliefs. These tasks examine whether children understand that others can have

different desires and beliefs from themselves. Wellman and Liu (2004) found diverse desire was the easiest for children, followed by diverse belief as the second easiest. The knowledge access task was the next easiest for children. It requires children to understand that not all individuals possess the same knowledge they do, and children were fairly successful at this task, with 73% of the 3-, 4- and 5-year olds passing this task.

Wellman and Liu (2004) used explicit, verbal false belief tasks in their study. Within explicit false belief tasks, there are unexpected transfer tasks, unexpected location/identity tasks, and explanation tasks. Wellman and Liu (2004) tested children on both an unexpected transfer task and an unexpected identity task. In the unexpected identity task, which Wellman and Liu called contents false belief, children are shown a Smarties box and are asked what is inside. Once they answer 'chocolate' or 'candy', the box is opened and they are shown that it actually contains something unexpected, like pencils or buttons. Children are asked what they first thought was in the box, and also what they think their friend will think is in the box (Carpendale & Lewis, 2015). The unexpected transfer task is very similar to the original Sally-Anne task (Baron-Cohen, Leslie & Frith, 1985). A character named Scott is looking for his mittens and thinks that they are in his closet. Children are told that the mittens are actually in Scott's backpack. Children are then asked where they think Scott will look for the mittens. Wellman and Liu (2004) found that these two tasks were very similar in difficulty level, with 59% of 3-, 4-, and 5-year old children successfully completing the unexpected content task, and 57% of those children successfully completing the unexpected transfer task. The similarity in performance between the two false belief tasks suggests that results were due to conceptual similarity between the two tasks and were not affected by differences with materials and procedures.

Another type of task that requires ToM but also ties in emotions is the belief-emotion task. The belief-emotion task is similar to the unexpected content task, but in this task the child judges how a person will feel, given a belief that is mistaken. The real-apparent emotion task is a related task that involves understanding emotions. The child must understand that a person can feel one thing but display a different emotion. These tasks that involve understanding emotions and emotion regulation are more challenging for children compared to other ToM tasks (Wellman & Liu, 2004). Other studies have found similar results in regards to ToM tasks that involve emotion. One task in which children listened to a story about Little Red Riding Hood demonstrated that some 4-year olds have trouble understanding false belief when emotions are involved. Four-year olds generally understood that Little Red Riding Hood had a false belief because she did not know her grandmother was actually the wolf. Even so, some children mistakenly believed that Little Red Riding Hood would be frightened by her grandmother (Bradmetz & Schneider, 1999). The ability to understand emotions in relation to false beliefs seems to develop after the age of 4 for most children.

### **Relationship between ToM and PT**

PT had been studied long before ToM was introduced, and there is a rich history to PT research. Although there is general consensus that the two constructs are related (Harwood & Farrar, 2006), their exact relationship is unclear. PT demonstrates one's ability to attribute specific perceptions to other people; therefore, one point of view is that PT is in fact a particular type of ToM (Apperly & Butterfill, 2009; Barnes-Holmes, McHugh, & Barnes-Holmes, 2004). For example, with respect to visual PT, Laranjo, Bernier, Meins and Carlson believe that "visual perspective taking is an important aspect of children's ToM" (2013, p. 49). Despite this, these

two constructs are traditionally studied separately, utilizing different tasks. Recent evidence, however, suggests that these skills *are* related.

Bigelow and Dugas (2008) found that children's performance on visual PT tasks was predicted by their performance on ToM tasks (and vice versa) while controlling for children's age and language abilities. Using fMRI, Schurz et al. (2015) also found that showing participants a different visual perspective engaged areas of the brain that are typically linked to ToM.

Harwood and Farrar (2006) also investigated the relationship between affective PT and ToM, and found that the two constructs were significantly positively correlated. The relationship between the two constructs was the strongest when the PT scenario involved conflicting perspectives (e.g., different emotional responses), and there was no relationship between the two constructs when the affective perspective scenario involved corresponding perspectives, and age and language ability were accounted for. Harwood and Farrar concluded that it is the ability to represent conflicting mental states that connects PT and ToM, and that the two constructs may not be related if the perspectives are compatible. Although Harwood and Farrar looked exclusively at ToM tasks that involved false beliefs, ToM does not always require a conflicting perspective (e.g. you could attribute the same desire to yourself and to someone else and that would still entail ToM). PT also does not always require a conflicting perspective. Although most tasks that are used to measure PT skills use conflicting perspectives to evaluate an individual's ability to take someone else's perspective, simply understanding that someone else views the same thing that you do also requires PT skills, but does not involve conflicting perspectives.

Another line of empirical evidence that supports the similarity between ToM and PT is the research that examines the link between ToM, PT and executive function. Executive function

is an umbrella term that refers to higher level cognitive control processes such as working memory, task shifting, and inhibition (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). These various cognitive processes work together to handle novel situations and to allow flexible and complex goal-directed behaviours. Numerous studies have examined the relation between ToM and executive function, with relatively few studies having been conducted on PT and executive function. A meta-analysis by Devine and Hughes (2014) reviewed 102 studies and found a moderate association between executive function and false belief understanding in 3 – 6 year olds. This association was not impacted by different executive function tasks but was impacted by different types of false belief tasks. Devine and Hughes found that explicit false belief tasks (e.g. unexpected transfer tasks) were substantially more strongly correlated with EF tasks than implicit/indirect (e.g. looking time tasks) measures of false belief. There are multiple explanations that have been proposed to explain the relationship between executive functions and false belief understanding. However, the discussion of these is beyond the scope of this paper (see Müller & Kerns, 2015 for a summary).

Another line of evidence suggests that the relationship between false belief understanding and executive function *is* dependent on the type of executive function measure. Carlson, Moses and Breton (2002) argue that ToM tasks require both inhibition and working memory and therefore believe that tasks that measure both inhibition and working memory (e.g. conflict inhibition tasks) should exhibit stronger correlations with ToM than tasks that measure more pure inhibition (e.g. delay tasks). Conflict inhibition tasks require children to both inhibit a prepotent response as well as activate a contradictory response, using both inhibitory skills and working memory. Delay tasks require children to inhibit responding until there is some sort of signal, arguably only using inhibitory skills (Carlson et al.). Carlson et al. found that conflict

inhibition tasks were significantly related to working memory whereas delay inhibition tasks were not. Moreover, results from a study by Carlson, Claxton and Moses (2015) support the hypothesis that there is a stronger relationship between conflict inhibition tasks and ToM tasks than between delay inhibition tasks and ToM tasks.

Carlson et al. (2015) administered two ToM tasks with high executive demands and two ToM tasks with low executive demands. They also administered measures of conflict inhibition, pure inhibition, working memory and verbal ability. Results did not clearly indicate whether tasks that measure inhibition alone or working memory alone are more highly related to ToM tasks. However, they did find that the more pure measure of inhibition (a delay task) was only significantly related to the sources of knowledge ToM task (low executive demands); whereas, the more pure measure of working memory was significantly related to the location false belief task (high executive demands). Findings demonstrated that the conflict inhibition task was highly related to ToM tasks with both high and low executive demands. Carlson et al. suggest that conflict inhibition abilities are a prerequisite skill for both successful ToM task performance and conceptual advances in ToM.

A study by Fizke, Barthel, Peters, and Rakoczy (2013) that investigated the link between executive function and other measures of ToM aside from false belief understanding also found a relationship between the two constructs. Fizke et al. found that executive function is related to both incompatible belief and incompatible desire tasks in which children are required to understand that others can hold beliefs and desires that differ from their own. Fizke et al. merge the concepts of PT and ToM and suggest that it is the ability to coordinate different perspectives during these ToM tasks that involve executive function. Fizke et al. found that it was self-perspective inhibition in particular that was significantly associated with PT and ToM. Inhibiting

one's own perspective is an important ability when trying to take someone else's perspective, especially when the two perspectives differ.

As previously mentioned, relatively few studies have been conducted that look at the relationship between PT and executive function. Research that looked at children's communicative PT abilities found that children who had more established inhibitory skills were better at inhibiting their own perspective and subsequently more easily assumed the perspective of someone else (Nilsen & Graham, 2009). Nilsen and Graham found that working memory was not predictive of children's ability to use another's perspective for interpreting a statement. However, Wardlow (2013) administered a PT task that evaluated a speaker's ability to take into account another's perspective when naming target objects and found that both working memory and inhibition were predictive of success on the task. Overall, executive function, especially inhibition and working memory, appear to play a significant role in co-ordinating perspectives and are an essential component of both ToM and PT. More research is needed to tease apart the relative influence of different aspects of executive function on ToM and PT.

Another line of evidence that supports the close relation between ToM and PT is research that demonstrates their association with language. Early language development has a strong impact on the development of ToM (Ebert, 2015; Milligan, Astington, & Dack, 2007). Language used on a day-to-day basis provides information about different mental states, and using language to talk about perspectives with others helps to build children's understanding of mental states and perspectives (Flavell, 2004). Carpendale and Lewis (2004) also believe that language is a building block for developing ToM, and have suggested that the nature of relationships early in children's lives affects the quality of communication, which in turn affects the development of ToM. The more language use and human interaction a child experiences, and the more

communicative relationships there are in a child's life, the greater understanding the child will have of others' perspectives and how others think.

### **Current Study**

There has been much debate about the development of ToM and PT and when children are able to 'pass' false belief tasks and other measures of social understanding. There are several questions arising from previous research that this study aims to answer. First, investigating the relative difficulties of ToM and PT tasks together is an important follow-up study to Wellman and Liu's (2004) study. Second, Wellman and Liu found that all of the ToM tasks they administered scaled onto a single continuum that they believed measured an increasing understanding of mental subjectivity. With evidence that there is a strong overlap between ToM and PT, incorporating PT tasks into a scaling study allows us to investigate whether or not both ToM and PT measure the same conceptual dimension. Third, although there is evidence that PT and ToM both involve executive function and verbal ability, determining whether all ToM and PT tasks engage these abilities equally or whether there are differences among the strength of these relationships is another area of interest. In addition, the current study will investigate whether Perner et al.'s (2002) and Moll et al.'s (2013) proposed distinctions within the construct of PT are valid. Examining whether the tasks that measure perspective-confronting are similar in terms of difficulty level and whether perspective-taking tasks are easier will provide information about the accuracy of this suggested distinction. Finally, the study will investigate the possibility of a reciprocal relationship between ToM and PT.

With regard to the first and second goals, expanding on Wellman and Liu's (2004) developmental sequence of ToM to include PT tasks will help to conceptualize the relationship between the two constructs. Wellman and Liu suggest that their ToM scale is evidence that all of the tasks that they administered develop along a single dimension, and the first goal of this study

is to examine whether PT tasks fall within this same dimension. Although visual PT does not necessarily involve understanding *mental* subjectivity as Wellman and Liu argue the ToM scale does, it still requires an understanding of subjectivity in terms of visual perspectives. While we have been discussing PT as a visual concept since it is typically assessed using visual tasks, there are theories that discuss perspective-taking from a more mentalistic standpoint (providing more evidence for the connection between the two constructs). Piaget provides a more mentalistic definition of PT and specifically relates children's confusion between objectivity and subjectivity to the failure in PT because they cannot distinguish between what is true and what only appears to be true (Rothbaum, 1979).

Based on Flavell's (1972) taxonomy of developmental sequences, Wellman and Liu (2004) suggest that their ToM task sequence could be either a modification or a mediation sequence. A modification sequence suggests "earlier items represent initial insights that are broadened or generalized to encompass later insights" (Wellman & Liu, p. 536). The ToM scale demonstrates the progression of the understanding of subjectivity in preschool children. Applying to Piaget's notion of ontological egocentrism, the ToM scale is the process of moving away from ontological egocentrism, as children are able to distinguish the subjective from the objective. Some tasks administered in Wellman and Liu's scale involved subjective-subjective distinctions whereas other tasks involved subjective-objective distinctions. For example, understanding that two people can have different beliefs about the same scenario (diverse belief) is a subjective-subjective distinction because the truth is not known; whereas understanding that someone can have a different belief than the objective truth (false belief) is a subjective-objective distinction. A mediation sequence goes beyond the idea of modification and suggests "earlier insights enable or aid in the attainment of later insights" (Wellman & Liu, p. 536). Although it is

not clear whether the ToM sequence by Wellman and Liu is one of modification or mediation, it is possible that the earlier understanding of certain tasks aids in the understanding of more challenging tasks. For example, the subjectivity of diverse desire that appears early in the scale could help in the understanding of the subjectivity of diverse belief that emerges later.

Incorporating PT tasks with ToM tasks and creating a new sequence will allow us to re-examine the scale while including a broader scope of social cognitive abilities. Investigating whether the PT tasks fit along the same dimension, and if they do, where they distribute in relation to the ToM tasks will provide information about how the two constructs relate. Flavell (1992) supports the idea of one fixed sequence for PT and ToM, suggesting that PT and ToM are developmentally related and that abilities related to both constructs should emerge in an orderly developmental sequence. Learning more about the differences in task difficulty and the development of children's PT and ToM ability could also help theorists to hone in on the mechanisms that foster their development.

The third objective of the current study is to evaluate the involvement of inhibition and verbal ability in the PT and ToM tasks. Past research has found that ToM tasks rely heavily on both language abilities and executive function, and inhibition more specifically (Call & Tomasello, 1999; Devine & Hughes, 2014; Milligan et al., 2007). Less research has been conducted examining the relationship between PT and language and inhibition, but some researchers do postulate that there is an association (Fizke et al., 2013; Flavell, 2004). This study aims to investigate whether PT and ToM are differentially related to these abilities.

The fourth aim of the current study is to further investigate the different levels of PT that have been suggested. Moll and Meltzoff (2011a; 2011b) argue that, because the findings from many studies demonstrate that children find certain PT tasks significantly more difficult than

others, it is necessary to distinguish in a more fine-grained manner between different levels of PT. Flavell (1992) and Perner et al. (2002) claim that level 2 PT tasks and false belief tasks measure the same ability. Level 2 PT tasks require children to understand that someone else can see the same thing as they themselves but in a different way, and false belief tasks require children to understand that someone else can hold a belief that is different from their own and counter to reality. Furthermore, false belief tasks require perspective-confronting rather than perspective-switching (Perner et al.) or perspective-taking (Moll et al., 2013) because in order to pass this task a person must hold in mind his or her own perspective (true belief) as well as the other's (false belief). This study will investigate whether false belief tasks are in fact more challenging because of the added required understanding that a belief can be false. In addition, examining both tasks that require confronting and tasks that require switching will help to provide information in regards to Perner et al.'s and Moll et al.'s proposed distinctions between the abilities.

Finally, the present study will address what type of relationship exists between PT and ToM and investigate whether PT abilities and ToM reciprocally influence one another. Examining when PT and ToM tasks develop will provide information about how a reciprocal relationship might function. For example, perhaps the completion of an easier ToM task (e.g., diverse desire) then helps to scaffold the understanding of a PT task (e.g., the turtle task), which in turn provides the prerequisite skills for the emergence of a more complex task (e.g., false belief). This idea fits well with Flavell's (1972) mediation sequence, according to which earlier understanding helps in the achievement of later understanding. To elaborate on this idea, while both ToM and PT abilities are developing, taking someone else's perspective could help strengthen ToM in children. A child who is able to take another person's visual perspective may

be better able to understand how that person will perceive their situation, and more specifically, they may be better able to understand certain beliefs the person will hold about their situation. Although there have been many studies conducted on ToM and PT abilities separately, more research needs to be done to investigate how these two abilities are related and how they interact. ToM and PT abilities are overlapping constructs, and learning more about how they are similar will add to our knowledge of the development of social understanding in children.

Following Wellman and Liu's (2004) study that found a consistent developmental progression demonstrating that certain ToM conditions are reliably more challenging than others, the present study will look at both PT and ToM tasks and how they relate. Successful completion of PT and ToM tasks is dependent on the complexity of the task administered and the cognitive abilities required. The current study will investigate the developmental sequence of PT and ToM with six tasks in young children. This study will test 3- and 4-year olds on three ToM tasks and three PT tasks to assess how well children perform on each task and identify whether they measure one underlying dimension. The three ToM tasks used will be the diverse desire task, the diverse belief task, and the explicit false belief task (Wellman & Liu, 2004). The three PT tasks used will be the turtle task (Masangkay et al., 1974), and the perspective-taking and perspective-confronting tasks conducted by Moll et al. (2013). Since PT and ToM both involve executive function, especially inhibition, (Fizke et al., 2013; Nilsen & Graham, 2009), and both are influenced by language abilities (Flavell, 2004), the expressive vocabulary subtest from the Kauffman Assessment Battery for Children (2<sup>nd</sup> edition), as well as the grass/snow task, a measure of conflict inhibition, will also be administered. The grass/snow task was selected as the measure of executive function in this study because past research has found that measures of conflict inhibition show a stronger association with ToM abilities than other measures of

executive function (Carlson et al., 2015; Müller, Liebermann-Finestone, Carpendale, Hammond, & Bibok, 2012).

Evaluating whether the PT and ToM tasks are related along one dimension has important implications in terms of theoretical and practical knowledge about children's social development. Wellman and Liu concluded that their ToM scale confirms that "ToM understandings represent an extended and progressive set of conceptual acquisitions" (p. 537). Incorporating PT tasks will possibly expand this set of conceptual acquisitions and will help to inform theories of social understanding in young children.

### **Hypotheses**

Following theories by Moll and Meltzoff (2011a; 2011b), Flavell (1992), and Perner et al. (2002), it is likely that children will perform the best on the perspective-taking task and that they will perform the worst on the explicit false belief task. Flavell and Perner et al. suggest that level 2 PT tasks measure the same ability as false belief tasks; however, false belief tasks require children to understand that a belief can be incorrect, which makes it more difficult than the turtle task and the perspective-confronting task. The turtle task and the perspective-confronting task are very similar, although the perspective-confronting task may be more difficult since the procedure is slightly more complicated and requires children to understand that a tinted screen can alter the apparent colour of an object. It is less clear where diverse desire and diverse belief will fall within the developmental sequence, although following Wellman and Liu's (2004) results, diverse desire may be the easiest task and diverse belief may be slightly more challenging.

Following Flavell's (1992) belief that PT and ToM abilities are conceptually related, and other factors that suggest a strong relationship between the two constructs (Flavell, 2004;

Harwood & Farrar, 2006), it is hypothesized that both PT and ToM tasks will develop along a single dimension. In addition, ToM and PT tasks developing along a single continuum would be consistent with the idea that these constructs are reciprocally related, with the development of earlier abilities facilitating the achievement of later abilities.

1. It is hypothesized that age will be a significant predictor of performance for all of the PT and ToM tasks, as well as the measures of inhibition and verbal ability, with older age predicting better performance.

2. It is hypothesized that task performance, from best to worst performance, will be perspective-taking, turtle task, perspective-confronting, and finally explicit false belief.

3. There are no specific predictions about diverse desire and diverse belief as it is less clear where they will fit among the other tasks; this study will explore whether they fit the scale, and if they do, where they fall in relation to the other tasks.

4. It is hypothesized that ToM and PT tasks will develop along a single dimension, allowing for the possibility of reciprocal relationships among the developing skills.

5. This study will also explore the extent to which relations among PT and ToM tasks were influenced by inhibition and verbal ability.

## Methods

### Participants

Eighty-seven 3- and 4-year old children participated in this study. Participants were a convenience sample collected from the Child Development Lab at the University of Victoria and from daycares in Victoria. Children with developmental disabilities were excluded from the study. Thirty-six participants were collected through the lab and 51 participants were collected through daycares. Three children were excluded from the analysis. One child was not able to identify colours, which is required for completing the perspective-taking and perspective-confronting tasks. A second child from the lab was excluded because the demographic form was not completed and therefore the child's exact age was unknown. A third child was excluded because he/she was a twin. Only one twin's data was kept in the analysis to maintain independence of the data. There were 56 three-year olds ( $M = 3.39$  years,  $SD = 0.29$  years, 34 females, 22 males) and 28 four-year olds in the final sample ( $M = 4.45$  years,  $SD = 0.27$ , 14 females, 14 males). The sample was reflective of the Western Canadian population, with majority of the sample being Caucasian. The sample performed above average (82<sup>nd</sup> percentile) on a measure of expressive vocabulary compared to a sample of preschoolers from the United States.

### Measures

#### Inhibition Task

*Grass/Snow task (Carlson & Moses, 2001)*. This task is a measure of conflict inhibition. The materials for this task included two red, felt hand prints, one laminated 15 x 15 cm green square, and one laminated 15 x 15 cm white square. The hand prints were placed on the table directly in front of the child, and the green and white squares were placed side by side above the

hand prints. At the beginning of the task, the experimenter asked the children what the colour of snow was, and what the colour of grass was. Then the children were told that this is a silly opposite game, and when the experimenter said snow, they should point to the green paper, and when the experimenter said grass, they should point to the white card. Children were instructed to place their hands on top of the red felt hand prints and to move their hand back onto the felt after they pointed to a card. Two practice trials were administered, and if children correctly pointed to the opposite colour for both practice trials, 16 more consecutive trials were run. If children failed one of the practice trials, they were corrected and asked to try again. If children could not complete the practice trials correctly after three attempts, they received a score of zero and the task was finished. Children received two points for responding correctly, one point for initially moving towards the incorrect response but then self-correcting, and zero points for responding incorrectly. The highest possible score was 32.

### **Verbal Ability Task**

*Expressive Vocabulary subtest from the Kaufman Assessment Battery for Children, 2nd edition (K-ABC; Kaufman & Kaufman, 2004).* In this measure of expressive vocabulary, the child was asked to name the object in a series of standardized images. Children received a score of 0 or 1 on each trial according to standardized criteria, and the task was discontinued when 4 consecutive scores of 0 were received. The final score was calculated by subtracting the number of errors from the ceiling item.

### **Theory of Mind Tasks**

*Diverse desire task (Wellman & Liu, 2004).* In the diverse desire task, children were introduced to a character named Mr. Jones and were told that it is snack time. Children were asked to choose which snack they would like best from two options (own-desire). Once they had

chosen, they were told that Mr. Jones actually likes the opposite snack better. Children were then asked what they thought Mr. Jones would choose to have for a snack (target question). Finally, children were asked what snack Mr. Jones likes the best (control question). For the task to be scored as correct, children had to answer the target question and control question opposite to their own desire answer.

***Diverse belief task (Wellman & Liu, 2004).*** In the diverse belief task, children were introduced to a girl named Linda who was looking for her cat. Children were asked whether they thought the cat was in the bushes or in the garage (own-belief). Once they had answered, they were told that Linda thought her cat was in the opposite location to whatever they just answered. Children were then asked where they think Linda would look for her cat (target question), and where Linda thought her cat was (control question). For the task to be scored as correct, children had to answer the target question and the control question opposite from their own-belief answer.

***Explicit false belief task (Wellman & Liu, 2004).*** In the explicit false belief task, or more specifically, the unexpected transfer task, children were introduced to a character named Scott who was looking for his mittens. Children were told that Scott's mittens might be in his backpack or they might be in his closet, but that they were actually in his backpack. Children were then told that Scott thought that his mittens were in his closet. Children were then asked where Scott would look for his mittens (target question) and where his mittens really were (control question). For the task to be scored as correct, children had to answer the target question with "closet" and the control question with "backpack".

### **Perspective-taking Tasks**

***Turtle task (Masangkay et al., 1978).*** In the turtle task, children were sitting across the table from the experimenter and there was an image of a turtle on the table between them.

Children were asked whether the turtle was upside down or right side up. Children then switched seats with the experimenter and were again asked whether the turtle was upside down or right side up. Finally, children were asked how the experimenter saw the turtle from where they were sitting (where the child was originally). For the task to be scored as correct, children had to state that the experimenter saw the turtle the opposite way that they saw the turtle.

***Perspective-taking task (Moll et al., 2013).*** In the perspective-taking task, children were sitting at a table with a half-clear, half-yellow tinted screen in front of them. There were two animals placed on either side of the screen (yellow and clear), and from the children's perspective they both appeared in their true blue colour. The experimenter then sat on the opposite side of the table and remarked that one of the animals looked like a certain colour to her/him (blue/green). The experimenter next asked the children to put the blue/green animal they just mentioned in a bag. For the task to be scored as correct, children had to recognize which animal the experimenter saw as that certain colour from where they sat.

***Perspective-confronting task (Moll et al., 2013).*** In the perspective-confronting task, children were sitting in front of a yellow tinted screen. A blue animal was placed on the children's side of the screen so the animal appeared blue to the children. The experimenter sat on the other side of the yellow screen so the animal appeared green to her. The children were asked both how they saw the animal from their side (blue or green) and also how the experimenter saw the animal from her side (blue or green). For the task to be scored as correct, children had to recognize that the experimenter saw the animal in a different colour than they did.

## **Procedures**

Children were tested at daycares in Victoria and also at the Child Development Lab at the University of Victoria. Parents were asked to fill out consent forms to allow for their child's

participation in the study. Upon receiving a signed consent form from parents, verbal assent from the child was obtained before testing commenced. Once the child had agreed to participate, he/she was individually assessed in a quiet area at the daycare he/she attended or in the Child Development Lab. Children were randomly assigned to one of two fixed orders, with forty-two children completing the tasks in order 1, and forty-two children completing the tasks in order 2.

**Daycares.** Testing sessions at daycares always began with the grass/snow task and ended with the K-ABC. Both orders of the ToM and PT tasks started with diverse desire as it is hypothesized to be easier to understand (Wellman & Liu, 2004) and helped the children to warm up. The perspective-taking task and perspective-confronting task were presented consecutively in both orders to maintain consistency with Moll et al.'s (2013) procedure. Within both orders, PT and ToM tasks alternated as much as they could given the constraints noted above. Aside from diverse desire and the perspective-taking and perspective-confronting tasks, the tasks were reversed between the two orders. Order 1 was the diverse desire task, the turtle task, the diverse belief task, the perspective-taking task, the perspective-confronting task, and finally the false belief task. Order 2 was the diverse desire task, the false belief task, the perspective-taking task, the perspective-confronting task, the diverse belief task, and finally the turtle task. The total time to complete all 8 measures was approximately 20 minutes. At the end of the session, children were given stickers to thank them for participating.

**Child Development Lab.** Children tested on these 8 measures at the Child Development Lab were part of a larger longitudinal study. At the start of the testing session, children completed one brief measure of task-switching before being presented with the grass/snow task. After the grass/snow task, children were randomly assigned to either order 1 or order 2 of the ToM and PT tasks, which were the same orders as presented in the daycares. The major

difference was that the K-ABC subtest was not presented until halfway through the testing session at the Child Development Lab, with 3 additional tasks and a 25 minute break placed in between the end of the ToM and PT tasks and prior to the K-ABC subtest.

## Results

### Data Preparation

Data was analyzed using RStudio version 1.0.143 and IBM SPSS Statistics 23. Correlations and descriptive statistics were calculated using SPSS, logistic and linear regressions were computed using both SPSS and the R ‘stats’ package, and the Rasch analysis was computed using the R ‘eRm’ package. Figures were created using the R ‘ggplot2’ package. The Guttman model was calculated by hand. All variables were screened for data entry errors and none were found.

Scoring for the perspective-confronting task and the turtle task required children to answer two questions: one about their own perspective and one about the experimenter’s perspective. To receive a score of 1 on the trial, children had to answer both questions correctly. Two different types of scores were created for the ToM tasks. One version included both the target question (e.g. where will Linda look for her cat?) and the control question (e.g. where does Linda think her cat is hiding?). The child had to pass both the target question and control question to be assigned a score of 1. The other scoring version for the ToM tasks included only the target question. The child had to pass the target question to be assigned a score of 1. When control questions were not included, the child had a 50% probability of being correct simply by chance.

Although Wellman and Liu (2004) included control questions in their analyses, I am interested in examining children’s performance without the control questions for two reasons. First, when the control questions were included with the target questions, the child’s probability of being correct by chance went down to 25%. Since the PT tasks differ in their probability of being correct by chance (with the turtle task and the perspective-confronting task having a 25%

probability and the perspective-taking task having a 50% probability), I am interested in how children's performance on the ToM tasks may be impacted by the difference in being correct by chance. The turtle task and perspective-confronting task cannot be transformed to have 50% probability of being correct by chance because that would alter the concepts behind the tasks (e.g. that children can understand both their own and someone else's perspective). Second, as will be discussed in more detail later on, when administering the ToM tasks I noticed that children sometimes appeared to be confused by the control question. Therefore, I am interested in looking at children's performance on the target questions alone to eliminate any confounds that may have impacted their response to the control question. To maintain consistency with Wellman and Liu's study, I ran all of the analyses with the control questions included; but to explore my own queries about children's performance on only the target questions, additional analyses excluding the control questions were included for the descriptive statistics and the Rasch analyses.

### **Descriptive Statistics and Correlations**

The range, means, and standard deviations among all eight tasks and age in months are presented in Tables 1 and 2.

Table 1

Mean, standard deviation and range of 8 tasks and age in months, with control questions included

Variable	<i>N</i>	Range	<i>M</i>	<i>SD</i>
Age	84	35.80 – 58.92	44.91	6.91
Expressive Vocabulary	84	6 – 27	19.05	3.97
Grass/snow	84	0 – 32	16.67	11.27
Diverse desire	84	0 – 1	0.68	0.47
Diverse belief	84	0 – 1	0.40	0.49
False Belief	84	0 – 1	0.30	0.46
Turtle Task	84	0 – 1	0.69	0.47
Perspective-taking	84	1 – 4	3.54	0.80
Perspective-confronting	84	0 – 4	2.40	1.64

*Note.* *M* and *SD* are used to represent mean and standard deviation, respectively. Age is reported in months.

Table 2

Mean, standard deviation and range of 3 ToM tasks, with control questions excluded

Variable	<i>N</i>	Range	<i>M</i>	<i>SD</i>
Diverse desire	84	0 – 1	0.83	0.38
Diverse belief	84	0 – 1	0.50	0.50
False Belief	84	0 – 1	0.50	0.50

*Note.* *M* and *SD* are used to represent mean and standard deviation, respectively.

Zero-order correlations and partial correlations are presented in Tables 3, 4 and 5. Tabachnick and Fidell (2007) note that correlations between continuous variables and dichotomous variables, and correlations between two dichotomous variables will always be well below 1.0. With that being said, there are several significant correlations among the dichotomous tasks. To detect a medium effect size ( $r = .30$ ) with an alpha of 0.05 and a power of 0.80, a power analysis demonstrated a desired sample size of 85 (Cohen, 1988).

For the zero-order correlations, the diverse desire task was significantly positively correlated with the diverse belief task and the perspective-confronting task. The turtle task was significantly correlated with the false belief task, the perspective-taking task and the perspective-confronting task. And the perspective-taking task and the perspective-confronting task were also significantly correlated. In addition, the expressive vocabulary subtest was significantly positively correlated with all of the PT tasks and the false belief task. Finally, the grass/snow task was significantly correlated with all of the PT and ToM tasks except for the false belief task and was also significantly positively correlated with the expressive vocabulary subtest.

**Correlations with expressive vocabulary subtest partialled out.** Controlling for the effects of verbal ability (Table 3), partial correlations continued to show significant relationships between most of the PT and ToM tasks. The tasks that remained significantly positively correlated were the diverse desire task and the diverse belief task, the turtle task and the perspective-taking task, the turtle task and the perspective-confronting task, and the perspective-taking task and the perspective-confronting task. This suggests that children's verbal ability accounted for the relationship between the turtle task and the false belief task.

**Correlations with grass/snow task partialled out.** Controlling for the effects of inhibition (Table 4), partial correlations continued to show significant associations between the

diverse desire task and the diverse belief task, the turtle task and the perspective-confronting task, and the perspective-taking task and the perspective-confronting task. This suggests that inhibition accounted for the relationship between the turtle task and the false belief task, and the relationship between the turtle task and the perspective-taking task. In addition, after controlling for inhibition, the relationship between the diverse desire task and the turtle task became significantly negatively correlated.

**Correlations with expressive vocabulary subtest and grass/snow task partialled out.**

Controlling for the effects of both inhibition and verbal ability (Table 5), the partial correlation between the perspective-taking task and the perspective-confronting task remained significant. This demonstrates that inhibition and verbal ability together accounted for all of the other significant positive correlations between the tasks. Additionally, the negative correlation between the diverse desire task and the turtle task remained significant after controlling for both verbal ability and inhibition.

Table 3

Zero-order and partial correlations among PT tasks, ToM tasks, and K-ABC, control questions included

Variable	1	2	3	4	5	6
1. Diverse Desire		.29**	<i>-.11</i>	<i>-.11</i>	.18	.21
2. Diverse Belief	.31**		.01	.01	-.01	.03
3. False Belief	-.05	.05		.17	.06	.07
4. Turtle Task	-.02	.08	.27*		.25*	.23*
5. Perspective-taking	.21	.02	.12	.32**		.32**
6. Perspective-confronting	.27*	.09	.19	.39**	.39**	
7. Expressive Vocabulary	.17	.15	.27*	.45**	.24*	.46**

*Note.* \* indicates  $p < .05$ ; \*\* indicates  $p < .01$ . Italicized correlations represent partial correlations with K-ABC expressive vocabulary partialled out.

Table 4

Zero-order and partial correlations among PT, ToM tasks, and grass/snow, control questions included

Variable	1	2	3	4	5	6
1. Diverse Desire		.22*	<i>-.15</i>	<i>-.22*</i>	.07	.06
2. Diverse Belief	.31**		<i>-.01</i>	<i>-.04</i>	<i>-.10</i>	<i>-.07</i>
3. False Belief	-.05	.05		.21	.05	.10
4. Turtle Task	-.02	.08	.27*		.20	.23*
5. Perspective-taking	.21	.02	.12	.32**		.24*
6. Perspective-confronting	.27*	.09	.19	.39**	.39**	
7. Grass/snow	.41**	.29**	.20	.40**	.37**	.53**

*Note.* \* indicates  $p < .05$ ; \*\* indicates  $p < .01$ . Italicized correlations represent partial correlations with grass/snow partialled out.

Table 5

Zero-order and partial correlations among PT, ToM tasks, K-ABC and grass/snow task, control questions included

Variable	1	2	3	4	5	6	7
1. Diverse Desire		<i>.21</i>	<i>-.16</i>	<i>-.24*</i>	<i>.07</i>	<i>.07</i>	
2. Diverse Belief	<i>.31**</i>		<i>-.02</i>	<i>-.06</i>	<i>-.10</i>	<i>-.09</i>	
3. False Belief	<i>-.05</i>	<i>.05</i>		<i>.15</i>	<i>.03</i>	<i>.04</i>	
4. Turtle Task	<i>-.02</i>	<i>.08</i>	<i>.27*</i>		<i>.18</i>	<i>.13</i>	
5. Perspective-taking	<i>.21</i>	<i>.02</i>	<i>.12</i>	<i>.32**</i>		<i>.22*</i>	
6. Perspective-confronting	<i>.27*</i>	<i>.09</i>	<i>.19</i>	<i>.39**</i>	<i>.39**</i>		
7. Expressive Vocabulary	<i>.17</i>	<i>.15</i>	<i>.27*</i>	<i>.45**</i>	<i>.24*</i>	<i>.46**</i>	
8. Grass/snow	<i>.41**</i>	<i>.29**</i>	<i>.20</i>	<i>.41**</i>	<i>.37**</i>	<i>.53**</i>	<i>.41**</i>

*Note.* \* indicates  $p < .05$ ; \*\* indicates  $p < .01$ . Italicized correlations represent partial correlations with grass/snow and K-ABC expressive vocabulary partialled out.

### Linear and Logistic Regressions

Linear regressions and binary logistic regressions were calculated to predict task success based on age in months. A power analysis for linear and logistic regressions was conducted using an alpha of 0.05, a power of 0.80, and a medium effect size ( $f^2 = .10$ ; Cohen, 1988). Based on these assumptions, the analysis demonstrated a desired sample size of 79. Linear regressions were computed for the multinomial measures (K-ABC subtest, the grass/snow task, the perspective-taking task and the perspective-confronting task; see Table 6), and binary logistic

regressions were computed for the binomial measures (the turtle task and all 3 ToM tasks; see Table 7; Tabachnick & Fidell, 2007). Age was a significant predictor of performance on the grass/snow task, the K-ABC expressive vocabulary subtest, the perspective-taking task, the perspective-confronting task, and the turtle task, with older age predicting better performance. Age was not found to be a significant predictor of performance on the diverse desire, diverse belief, or false belief task. Harrell (2015) recommends reporting the Wald statistics when performing binary logistic regressions. For the turtle task, an increase in age was associated with an increase in the odds of passing the turtle task, with an odds ratio of 1.17 (95% CI, [1.06, 1.28]), Wald  $X^2(1) = 10.47, p < .001$ .

Table 6

## Linear regressions

Dependent Variable	R <sup>2</sup>	F <sup>2</sup>	Mean Square	F	Significance
Grass/Snow	.32	.47	3413.01	39.22	.000
K-ABC	.26	.35	339.85	28.79	.000
Perspective-taking task	.10	.11	5.502	9.52	.003
Perspective-confronting task	.30	.43	66.45	34.98	.000

*Note.* Predictor is age in months.

Table 7  
Logistic binary regressions

Dependent Variable	<i>B</i>	<i>SE B</i>	R <sup>2</sup>	F <sup>2</sup>	Exp( <i>B</i> )	CI Exp( <i>B</i> )	Wald <i>X</i> <sup>2</sup>	Significance
Turtle task	.15	.05	.15	.18	1.17	1.06 – 1.28	10.47	.001
Diverse desire task	.07	.04	.05	.05	1.07	1.00 – 1.16	3.62	.057
Diverse belief task	.00	.03	.00	.00	1.00	.94 – 1.07	.006	.939
False belief task	.05	.04	.03	.03	1.05	.99 – 1.13	2.31	.129

*Note.* Predictor is age in months. R-squared is Cox & Snell R-squared. *SE* and *CI* are used to represent standard error and 95% confidence interval, respectively.

## Item Difficulty

Another aim of the study was to assess the difficulty of each task. Tables 8 and 9 present the proportion of children correct on each task. With control questions included, the order of task difficulty from easiest to hardest was the perspective-taking task, the turtle task, the diverse desire task, the perspective-confronting task, the diverse belief task, and finally the false belief task. With control questions excluded, the order of task difficulty from easiest to hardest was the perspective-taking task, the diverse desire task, the turtle task, the perspective-confronting task, and finally the diverse belief task and the false belief task with the exact same proportion of children correct.

Table 8

Proportion of children correct on each task, control questions included

Task	Description
Perspective-taking (87%)	Child understands that someone else can see the same object in a different colour than they see it.
Turtle Task (69%)	Child understands that someone can see a picture from a different perspective than they do.
Diverse desire (68%)	Child understands that someone else can have a different desire than himself/herself about the same objects.
Perspective-confronting (60%)	Child understands both how they see an object and also how someone else sees that object differently.
Diverse belief (40%)	Child understands that someone else can have a different belief than himself/herself about the same object, when the child does not know which belief is true or false.
False Belief (30%)	Child understands how someone will search for an object, given that person's incorrect belief.

Table 9

Proportion of children correct on each task, with control questions excluded

Task	Description
Perspective-taking (87%)	Child understands that someone else can see the same object in a different colour than they see it.
Diverse desire (83%)	Child understands that someone else can have a different desire than himself/herself about the same objects.
Turtle Task (69%)	Child understands that someone can see a picture from a different perspective than they do.
Perspective-confronting (60%)	Child understands both how they see an object and also how someone else sees that object differently.
Diverse belief (50%)	Child understands that someone else can have a different belief than himself/herself about the same object, when the child does not know which belief is true or false.
False Belief (50%)	Child understands how someone will search for an object, given that person's incorrect belief.

In addition to investigating the relative difficulties of the tasks, the scalability of the tasks was assessed using Rasch and Guttman methods. Both Rasch and Guttman models order dichotomous tasks on a continuum. To prepare the data for the Rasch and Guttman analysis, the perspective-confronting task and the perspective-taking task scores were converted into dichotomous scores. To maintain consistency, since there was only a single trial of the turtle task, only the first trials from both tasks were used. If children passed the first trial, they were assigned a score of 1 and if they failed the first trial, they were assigned a score of score of 0. For the turtle task and the perspective-confronting task, the child had to get both their own perspective and the experimenter's perspective correct from the first trial to receive a score of 1. For the perspective-taking task, the child was only required to answer a question about the experimenter's perspective from the first trial to receive a score of 1. Descriptive statistics for the

perspective-confronting task and the perspective-taking task after being converted into dichotomous scores are presented in Table 10.

Table 10

Mean, standard deviation and range of tasks converted to dichotomous scores

Variable	<i>N</i>	Range	<i>M</i>	<i>SD</i>
Perspective-taking	84	0 – 1	0.87	0.34
Perspective-confronting	84	0 – 1	0.60	0.49

*Note.* *M* and *SD* are used to represent mean and standard deviation, respectively.

**Rasch analysis.** The Rasch model is a type of one-parameter item response theory model. In contrast to a three-parameter item response theory model, the Rasch model does not include a parameter that accounts for guessing, and the Rasch model assumes that all items that fit the model have equivalent discriminations. The Rasch model creates the most probable response pattern for a participant when items are ordered from least difficult to most difficult, and each participant and item deviation from that order can be assessed (Bond & Fox, 2001). The Rasch model is probabilistic. If a participant passed a task that was in the middle of the scale, it is probable that they also passed the tasks ranked as easier on the scale. Table 11 shows the item parameters, person parameters, error, and standardized infit and outfit values for the six item Rasch model with control questions included. Table 12 presents the same information but with the control questions excluded. Means and standard deviations for parameter estimates, error, infit and outfit values are also reported. The Rasch model provides a statistical test for measuring unidimensionality. If the items fit the Rasch model, then they measure the same dimension (Müller, Sokol, & Overton, 1999).

Item parameter values are estimated as easiness parameters using conditional maximum likelihood (Mair, Hatzinger, & Maier, 2009). Infit and outfit values are calculated for both tasks and participants and are used to identify the tasks and participants whose pattern of responses deviate more than expected from the single underlying continuum (Bond & Fox, 2011). Outfit values are more sensitive to participants who have an unexpected response on an item that would theoretically be either easy or hard for them (e.g., an item far from a person's or item's parameter estimate). Infit values are more sensitive to participants with unexpected responses on an item that is roughly at their difficulty level (e.g., an item close to a person's or item's parameter estimate). Infit and outfit  $t$  values greater than 2.0 indicate misfit (Bond & Fox). Negative values suggest that there is overfit of the model, indicating that the model may be more deterministic (like the Guttman scale) than probabilistic. In the computation of item and person infit and outfit, participants who got all tasks correct (6/6) or got all tasks incorrect (0/6) were excluded as the analysis cannot provide accurate estimates for these individuals because they are either too capable or not capable enough, and it is unclear where their true ability level lies (e.g. very low or very high; Müller et al., 1999).

As shown in Table 11, all item fit statistics are less than 2.0, and the mean fit statistics for item fit are reasonable (standardized infit,  $M = -0.53$ ,  $SD = 0.69$ ; standardized outfit,  $M = -0.41$ ,  $SD = 0.91$ ). Mean fit statistics for person fit are good, with the means being close to the expected value of 0 (standardized infit,  $M = -0.10$ ,  $SD = 0.87$ ; standardized outfit,  $M = -0.02$ ,  $SD = 0.82$ ). As shown in Table 12, all the item fit statistics for the model without control questions are less than 2.0 except the turtle task, which has an infit value of -2.14 and an outfit value of -2.23. For this model, the mean fit statistics are also reasonable (standardized infit,  $M = -0.41$ ,  $SD = 1.12$ ; standardized outfit,  $M = -0.33$ ,  $SD = 1.15$ ), and the mean fit statistics for person fit are also good

(standardized infit,  $M = -0.07$ ,  $SD = 0.91$ ; standardized outfit,  $M = -0.02$ ,  $SD = 0.93$ ). These fit statistics demonstrate that these 6 items fit the Rasch model well, and that the items fit slightly better when control questions are included in the model.

Person parameter estimates were correlated with age in months. With control questions included, person parameters and age in months were significantly correlated,  $r(82) = .42$ ,  $p < .001$ . Without control questions included, person parameters and age in months were also significantly correlated,  $r(82) = .34$ ,  $p < .01$ .

Table 11

Item and Person Summary and Fit Statistics for the Rasch Model with control questions included

	Parameter Estimate	Error	Infit <i>t</i>	Outfit <i>t</i>
Diverse Belief	-0.96	0.23	-0.11	0.57
False Belief	-1.49	0.24	0.04	-0.48
Turtle Task	0.43	0.24	-1.27	-1.69
Perspective-taking	1.72	0.32	-0.27	0.39
Perspective-confronting	-0.06	0.23	-1.55	-1.27
Diverse Desire	0.37	0.24	-0.03	0.03
Mean	0.00	0.25	-0.53	-0.41
<i>SD</i>	1.13	0.04	0.69	0.91
Person ability and fit statistics				
Mean	0.47	1.01	-0.10	-0.02
<i>SD</i>	1.15	0.11	0.87	0.82

*Note.* Expected values for infit *t* and outfit *t* is a mean of 0 and standard deviation of 1.0; fit statistics > 2.0 indicate misfit. Item parameters estimated in terms of easiness, with higher numbers indicating an easier task. Infit and outfit values calculated using 77 participants.

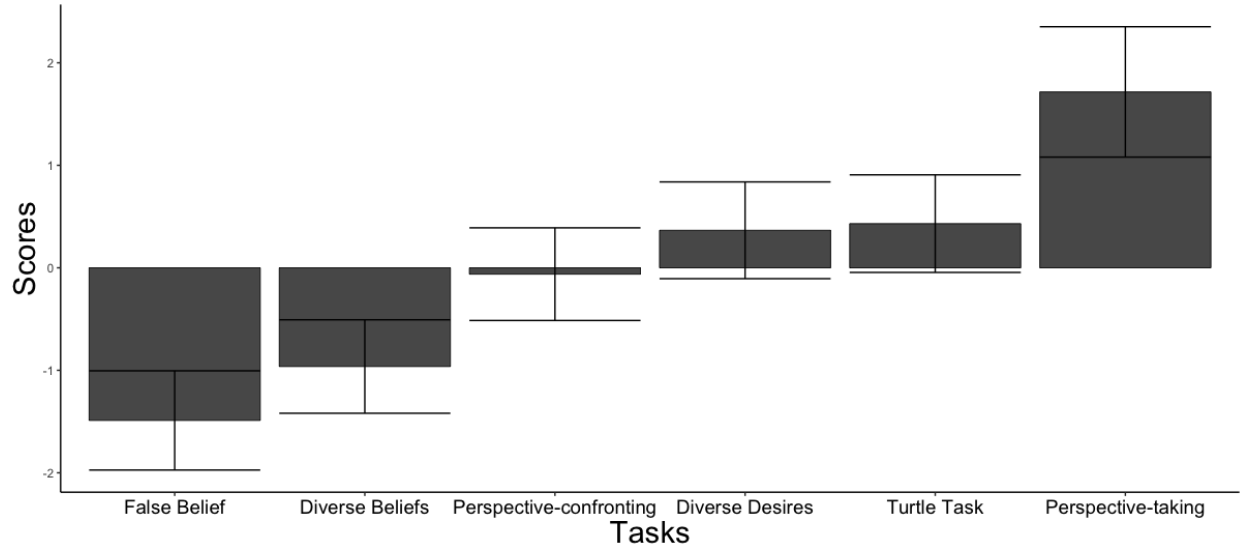
Table 12

Item and Person Summary and Fit Statistics for the Rasch Model with control questions excluded

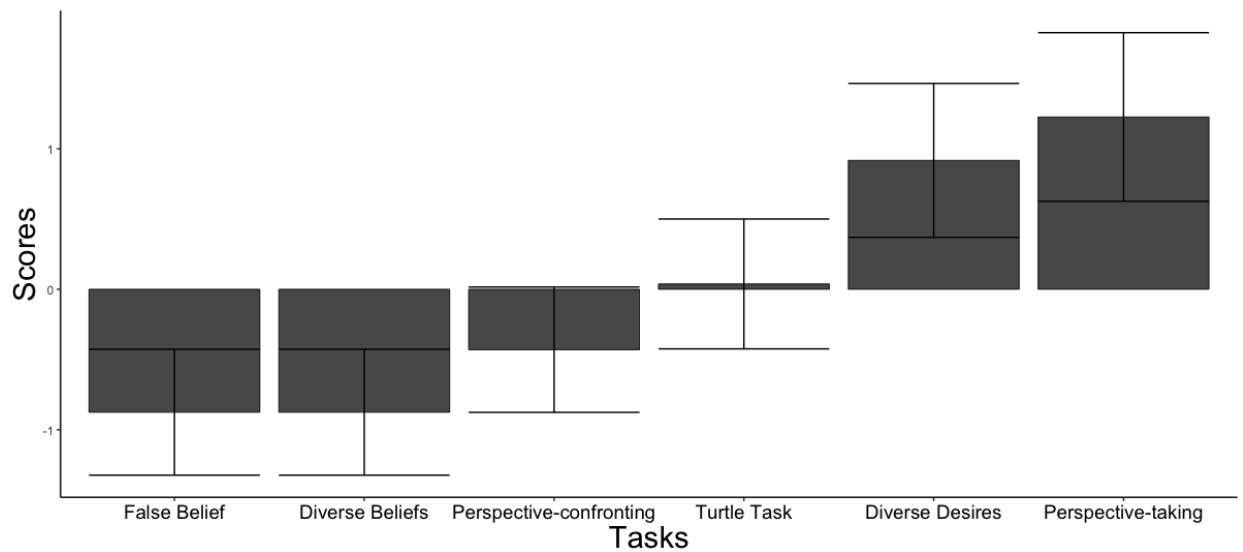
	Parameter Estimate	Error	Infit $t$	Outfit $t$
Diverse Belief	-0.88	0.22	0.17	-0.04
False Belief	-0.88	0.22	0.69	0.57
Turtle Task	0.04	0.23	-2.14	-2.23
Perspective-taking	1.23	0.30	-0.65	-0.15
Perspective-confronting	-0.43	0.22	-1.17	-1.26
Diverse Desire	0.92	0.27	0.63	0.80
Mean	0.00	0.25	-0.41	-0.33
<i>SD</i>	0.90	0.03	1.12	1.15
Person ability and fit statistics				
Mean	0.60	0.97	-0.07	-0.02
<i>SD</i>	0.93	0.10	0.91	0.93

*Note.* Expected values for infit  $t$  and outfit  $t$  is a mean of 0 and standard deviation of 1.0; fit statistics  $> 2.0$  indicate misfit. Item parameters estimated in terms of easiness, with higher numbers indicating an easier task. Infit and outfit values calculated using 70 participants.

Figure 1 shows the parameter estimates and error values for the Rasch model for all six tasks with the control questions included. Figure 2 shows the parameter estimates and error values for the Rasch model for all six tasks with the control questions excluded.



*Figure 1.* Parameter estimates and error values for the Rasch model with control questions included.



*Figure 2.* Parameter estimates and error values for the Rasch model with control questions excluded.

**Guttman scaling.** In contrast to the Rasch model, Guttman scaling is a stringent way of ranking items. Guttman scaling requires that when a child gets one scale item incorrect, they must have passed all previous items on the scale. Therefore, if we know a child's score on the

scale, then there is only one possible response pattern they could have exhibited and we can conclude exactly how the child performed on each task. Guttman scalogram analysis was completed for the model with the control questions included because it showed a better fit with the Rasch model. In addition, since diverse belief and false belief were comparable in difficulty level when control questions were excluded, criteria would likely not be met for a Guttman scale (Hiller, Weber, & Young, 2014). Table 13 shows the Guttman scalogram results for the 6 tasks with control questions included. As shown in Table 13, 42% of children (35 of 84) fit this 6-item Guttman scale precisely. The coefficient of reproducibility, a standard statistic for a Guttman scale, is .87 (Green, 1956). Values greater than .90 suggest that the items are scalable.

Table 13

## Guttman scalogram for six tasks

Pattern	1	2	3	4	5	6	7	Other patterns	N
PT	–	+	+	+	+	+	+		
TT	–	–	+	+	+	+	+		
DD	–	–	–	+	+	+	+		
PC	–	–	–	–	+	+	+		
DB	–	–	–	–	–	+	+		
FB	–	–	–	–	–	–	+		
Participants									
Total	2	4	3	3	8	10	5	49	84

## Discussion

The aim of the present study was to gather information about the relationship between PT and ToM. The primary objectives were to investigate the relative difficulties of several PT and ToM tasks, and to evaluate whether PT and ToM develop along a single continuum. Additional aims of the present study were to investigate the relationships between the PT and ToM tasks with inhibition and verbal ability, and to examine whether the distinctions made within PT (e.g. perspective-taking and perspective-confronting) are valid.

### Task Difficulty

It was hypothesized that task difficulty, from easiest to most challenging, would be perspective-taking, turtle task, perspective-confronting, and finally explicit false belief. No specific hypotheses were made regarding diverse desire and diverse belief. Results of task performance support this hypothesized order. The difficulty of diverse desire and diverse belief changed depending on whether control questions were included or not. With control questions included, diverse desire and diverse belief were much more challenging for children to pass. On the one hand, this makes sense due to the change in probability of passing the task by chance (must get both target and control questions correct to receive a score of 1). On the other hand, it appeared that children would sometimes become confused by the control question after answering the target question correctly. The repetitive nature of asking “Where does Linda think her cat is?” right after asking “Where will Linda look for her cat?” appeared to confuse some children even when it seemed they had a good grasp of the task. Of the 42 children who passed the diverse belief target question, 8 failed the related control question. Of the 70 children who passed the diverse desire target question, 13 failed the related control question. Control questions were asked to assess the child’s understanding of reality, or of someone else’s state of mind. If a

child answered the target question incorrectly, we could check the control question to evaluate whether they understood what the other person desired or believed, yet still held an egocentric perspective that the individual would choose/believe the same as themselves. If the child did not correctly answer the target and control questions, then one can assume they did not understand the task format and questions (Wellman & Liu, 2004). The children who answered the target question correctly, but not the control question, are somewhat of an anomaly. They were either arbitrarily guessing for the target question and getting it correct, without really understanding the task, or they were able to answer the target question correctly because they understood the task but then became confused by the repetitive questioning, and changed their answer for the control question (to the incorrect response). Children may infer that we repeated the question because their first response was incorrect, consequently changing their answer for the control question.

When control questions were excluded, children performed equally well on the diverse belief task and the false belief task. This demonstrates that children correctly answered that Scott would look in his closet for his mittens even though they were actually in his backpack just as often as children correctly answered that Linda would look for her cat in the opposite location compared to where the children thought the cat was hiding. When children were asked where the mittens really were, and where Linda thought her cat was, children were more successful at answering the diverse belief control question (about the cat's location) correctly than the false belief control question (about the mittens' location). This suggests that the control questions are an important part of the task, because past research has consistently found that the diverse belief task is easier for children than the false belief task (Hiller, Weber, & Young, 2014; Wellman & Liu, 2004; Zhang, Shao, & Zhang, 2016).

Although children performed as we hypothesized on the turtle task relative to the other tasks in the scale, children in this study were able to pass the turtle task at a younger age than past research has suggested. Perner et al. (2002) and Moll et al. (2013) argue that children cannot pass perspective-confronting tasks until 4- to 4.5-years old, and more specifically, Masangkay et al. (1978) found that it was not until 4.5-years old when children were able to pass the turtle task. The findings of the present study suggest that children as young as 3-years old are able to pass the turtle task. Out of the 56 three-year olds in the study, 59% passed the turtle task. In addition, results of the present study suggest that Flavell (1992) and Perner et al.'s (2002) claim that level 2 PT tasks and false belief tasks measure the same ability is incorrect. Results suggest that false belief understanding is more challenging for preschoolers than all three of the level 2 PT tasks administered. False belief understanding requires children to accept that someone's belief can be false, which is not a requirement of the level 2 PT tasks.

It was also hypothesized that age would be a significant predictor of task success on all eight tasks, with older age predicting better performance. Linear and logistic regressions provided support for this hypothesis for the grass/snow task, the expressive vocabulary subtest, the turtle task, the perspective-taking task, and the perspective-confronting task. Age did not predict success on the three ToM tasks. One potential reason for this finding may be the limited age range of the children tested. In addition, there was an uneven distribution of 3-year olds and 4-year olds, with more 3-year olds included in the sample.

### **Subjectivity Scale**

Another main goal of this study was to evaluate whether PT and ToM tasks would develop along a single dimension. Although the criterion was not met for the stringent Guttman scale, results of the probabilistic Rasch analysis demonstrated good internal consistency between

the six tasks, with all PT and ToM tasks fitting the Rasch model. This finding supports past hypotheses that PT and ToM may be measuring the same underlying concept (Flavell, 1992). Fit statistics were better for the model with the control questions included. This could indicate that the ToM tasks with the control questions are capturing a child's understanding of subjectivity more accurately compared to using the target question alone. This finding is somewhat contradictory to the above discussion regarding children being confused by the control questions. Perhaps the better model fit with control questions is because including both target and control questions more accurately represents children's understanding of the task. If children were not confused by the control question, including the control question as part of the dichotomous score more accurately captures the level of understanding of children who passed the target question by guessing (since they should not pass the task), which in turn would result in a better fit for the Rasch model.

The results of the Rasch analysis suggest that there are multiple steps preschoolers go through to develop an increasing understanding of subjectivity. I think conceptualizing these achievements in PT and ToM as the increasing understanding of subjectivity incorporates both constructs nicely. Although being aware of subjectivity requires the ability to see different perspectives, it is a broader term that encompasses a wide range of understanding including both visual and mentalistic perspectives. The Rasch scale represents a series of steps children take as they advance and learn more about varying perspectives and mental states. The good fit statistics for the Rasch model indicate that this six-item model could be used in the future to measure the degree of subjectivity understanding in preschoolers. Using their ToM scale, Wellman and Liu (2004) found progressive conceptual development underlying ToM understanding with some aspects of ToM being reliably more challenging than others. This study expands those findings

and provides evidence for a wider degree of conceptual development in preschoolers that encompasses both PT and ToM, while still showing that some aspects of social understanding are reliably easier (e.g. perspective-taking) than others (e.g. false belief understanding).

### **Inhibition and Verbal ability**

Before controlling for verbal abilities or inhibitory skills, many of the PT and ToM tasks were significantly correlated. Interestingly, after controlling for both verbal ability and inhibition, the only tasks that remained significantly positively correlated were the perspective-taking task and the perspective-confronting task. Verbal ability on its own accounted for the relationship between the turtle task and the false belief task, and inhibition accounted for the relationship between the turtle task and the false belief task, and the relationship between the turtle task and the perspective-taking task. Meta-analyses by Devine and Hughes (2014) and Milligan et al. (2007) investigated the relationship between false belief understanding with executive function and language, respectively. These studies found that language and executive function play a large role in the development of false belief understanding, with longitudinal research demonstrating that language and executive function predict false belief understanding. This study adds to this past research by showing that verbal ability and inhibition are not just related to false belief understanding, but also to other ToM abilities and visual PT tasks as well.

Thinking about these relationships in regards to the Rasch model, it raises the question of whether the Rasch scale is really measuring subjectivity if executive function and verbal abilities account for the relationships among the PT and ToM tasks. On the one hand, it is possible that the Rasch model is actually measuring children's executive function and verbal abilities, with the more challenging tasks requiring higher inhibitory and/or verbal abilities. On the other hand, although the grass/snow task and the expressive vocabulary subtest were significantly correlated

with almost all of the PT and ToM tasks, there was no clear finding that suggests the tasks that were more challenging (e.g. false belief) required higher inhibitory or language skills than the easier tasks (e.g. the turtle task). More research needs to be conducted to determine whether the relationship between ToM and PT is due to the involvement of executive function and verbal ability, or whether there is a unique ability that holds ToM and PT together.

### **Relationship between PT and ToM**

An additional goal of this study was to investigate the relationship between PT and ToM. Past research has discussed PT as a specific case of ToM (Barnes-Holmes et al., 2004), and although on the basis of findings from this study I cannot conclude whether that is the case or not, the results do confirm that the two constructs are related. The results demonstrate that PT and ToM skills develop together, with easier PT and ToM skills developing before more challenging PT and ToM skills. The scaling of both PT and ToM tasks along one dimension suggests a possible reciprocal relationship between the constructs, with the understanding of simpler PT and ToM abilities supporting the understanding of more challenging PT and ToM skills. If the understanding of subjectivity is what unites both PT and ToM, perhaps easier levels of subjectivity understanding (e.g., that someone can have a different perspective from oneself) can scaffold the understanding of more complex levels of subjectivity (e.g., understanding that someone can have a belief that is not true). Referring back to Flavell's (1972) mediation sequence, earlier developing subjectivity understanding may help in the achievement of later developing subjectivity understanding.

### **Distinctions Among Constructs**

When verbal ability and inhibition were controlled for separately, the relation between the turtle task and the perspective-confronting task remained significant. This supports the

hypothesis by Moll et al. (2013) that their perspective-confronting task is measuring the same ability as the turtle task. However, when both verbal ability and inhibition were controlled for, the correlation between the turtle task and the perspective-confronting task was no longer significant. Although Moll et al. is correct that the two tasks are related, they may be related due to similar cognitive demands (verbal ability and inhibition) as opposed to their relation in confronting perspectives.

Even though the turtle task and the perspective-confronting task were positively correlated, the turtle task was less challenging for children than the perspective-confronting task. This could be due to more difficult task demands in the perspective-confronting task, in which children must understand that a yellow-tinted screen changes the perception of a blue animal to appear green. The turtle task does not require that extra level of understanding and only requires a child to understand that depending on where an individual is sitting at a table, a stationary picture will look different (e.g., upside down) from different positions. Furthermore, the turtle task includes an extra level of interaction when the child is asked to switch seats with the experimenter. The meta-analysis by Wellman et al. (2001) demonstrated that when children are actively engaged in ToM tasks (e.g. helping with the change of location during a false belief task), it increases the odds of them being correct by 1.96 over task versions with a passive participant who only observes the task. In the turtle task, switching seats is similar to the active engagement demonstrated in the ToM tasks, in which children play an active role in changing either their own visual perspective or that of the experimenter, or the belief that a character holds. This extra engagement with the children may increase their motivation and attention, and in turn help to improve their performance on the task.

The difference in difficulty level between the perspective-taking task and the perspective-confronting task provides evidence for Moll et al.'s (2013) and Perner et al.'s (2002) distinction between the two abilities. The turtle task distributes between these two tasks, but is closer in difficulty level to the perspective-confronting task. Furthermore, the perspective-taking task matches the perspective-confronting task in terms of more challenging task demands; therefore, children may actually perform better on a different type of perspective-taking task that does not require the understanding of the yellow screen.

The results of this study add to the literature on social understanding by providing information about ToM, visual PT, inhibition, verbal ability, and their relationships. As discussed earlier, aspects of social understanding such as ToM and PT are predictive of children's social skills. The results of this study suggest that young children could benefit from interventions targeting their inhibition and verbal ability to indirectly aid in their social understanding. In addition, directly training ToM and PT skills could also benefit children's social understanding and improve social skills. A review by Kloo and Perner (2008) found several ToM and executive function training programs that have been shown to improve ToM and executive functions skills in young children. One study by Fisher and Happé (2005) that examined the effectiveness of training children with autism spectrum disorders on ToM found that not only did training children on ToM improve ToM, training children on an executive function task that required set-shifting also improved ToM. There is less research that has investigated training visual PT, but one study found that training visual PT using a drawing intervention resulted in a positive change in PT abilities (Cigala, Mori, & Fangareggi, 2015).

### **Limitations of the Current Study**

Although the results of the Rasch analysis do suggest that the tasks fit along one continuum, meaning they could be potentially measuring the same underlying trait, it is important to be aware of the possibility for confirmation bias. We can reject the hypothesis that these six items do not fit the Rasch model, but we cannot say for certain that all six tasks are measuring the same latent trait. For example, it is possible to include a task in a Rasch analysis that is completely unrelated (e.g. one math task along with the ToM and PT tasks) and still find that all tasks fit the model. Consequently, although it is important and meaningful that the six ToM and PT tasks do not *not* fit the Rasch model, it is of note that we cannot guarantee they are measuring the same trait.

An additional limitation of this study is the potential working memory demands during the ToM/PT tasks. This is a limitation because I hypothesize that some tasks require more working memory than others, impacting task difficulty differentially based on working memory abilities. While many of the tasks administered were quite short in length, there was still a component of working memory required to be successful. In regards to the ToM tasks, once the premise is explained to the child, they are immediately asked the target question, and then asked the control question. Although the premise is relatively short (see Appendix B, D and G), there is still enough information that it could be challenging for children to hold it in their mind as they answer both the target and control questions. I believe that ToM tasks require more working memory than the visual PT tasks do, as the ToM tasks require you to hold the story in mind and manipulate the information to answer correctly.

Although we did not include a specific measure of working memory, Carlson and Moses (2001) argue that the grass/snow task is a measure of what they call conflict inhibition, and that it

actually requires both inhibitory abilities and working memory. For the grass/snow task, there are 16 trials for children to complete and there is no repetition of the task rules partway through the trials. Children must hold the rules of the game in their mind as they play the game and try to inhibit their pre-potent response. It is clear that there is a relationship between the grass/snow task and almost all of the PT and ToM tasks administered in this study. However, it is unclear whether the relationship is in regards to specific task demands in the ToM and PT tasks that require working memory, or whether the ability to take someone else's perspective and understand others' mental states requires working memory. The grass/snow task was significantly correlated with all PT and ToM tasks except for the false belief task. Although I postulated that working memory would play a larger role in the ToM tasks than the PT tasks, since grass/snow measures both working memory and inhibition, we do not have a clear picture of which tasks involved more working memory independently of involving inhibition.

Another limitation is the cross-sectional nature of the study. Although past research suggests that executive function and language predict false belief understanding, this study cannot comment on the directionality of the relationship between conflict inhibition, verbal ability, PT and ToM.

### **Future Directions**

To expand on the findings from this study, future research should conduct a longitudinal study to look at ToM, PT, executive function and language. Examining the directionality between executive function, language, and different measures of ToM and PT would help add to the literature that has found an asymmetric pattern of relations between executive function and language with false belief understanding (Devine & Hughes, 2014; Milligan et al., 2007). Furthermore, incorporating more varied measures of executive function (e.g. working memory,

task switching) would help elucidate which specific aspects of executive function are involved in which tasks. Additionally, longitudinal research would help to clarify the potential reciprocal relations between PT and ToM suggested earlier.

In addition to the directionality of these relationships, trying to tease apart extraneous cognitive demands from the PT and ToM tasks is another area for future research. Although there is not much that can be done to reduce the inhibitory abilities required for these tasks, there are task manipulations that could be incorporated to reduce the involvement of language and working memory. Having a more clear visual display of the events in each ToM story would help to lessen the load on children's working memory. There are also modifications that could be made to reduce the amount of language required in task performance. Call and Tomasello (1999) tested children on both a traditional false belief task and a nonverbal false belief task, and found that the two tasks were highly correlated. In this study, the turtle task required more expressive language than the other tasks since all the other tasks offered a pointing/grabbing option. Choosing tasks that require an equal amount of language would be helpful when investigating task difficulties relative to one another. Using nonverbal tasks would help to tease apart language from the ToM and PT tasks and provide a more pure measure of ToM and PT.

One interesting finding of this study was that after controlling for verbal ability and inhibition, there was a significant negative correlation between the diverse desire task and the turtle task. This unexpected and unusual negative relationship between a PT and ToM task lends itself to further investigation to see if this association would be replicable in a future study.

Another relevant follow-up to this study would be to add more ToM and PT tasks and evaluate whether they still create a unitary scale. If a wider range of PT and ToM tasks (e.g. adding emotion-related ToM tasks and level 1 PT tasks) were found to scale onto one dimension,

this could provide a broader measure of subjectivity understanding for children. In addition, it would provide more information about the relative difficulties of different PT and ToM skills.

Finally, administering measures of perspective-taking and perspective-confronting that do not require the understanding of the yellow-tinted screen would provide more clear evidence for the distinction between these abilities. Furthermore, comparing Perner et al.'s (2002) understanding of confronting perspectives with Moll et al.'s (2013) description of perspective-confronting is another area for future research. Looking at within-person confronting perspectives (Perner et al.) compared to between-person perspective-confronting (Moll et al.) and whether they are equally as challenging for children would be an interesting study to help confirm the different levels of PT and the varying difficulty of different tasks.

## **Conclusion**

The aim of the present study was to add to the literature on social understanding in children by examining the relationship between PT and ToM and how they develop in preschoolers. In line with past research that has suggested an association between PT and ToM, the results indicated that PT and ToM develop along a single continuum. I propose that these six tasks are measuring an increasing understanding of subjectivity in children, and that this six-item scale can be used to assess preschoolers' understanding of subjectivity. Results also show that understanding that others' beliefs can be false is the most challenging of all six of the tasks administered. Findings of this study support Moll et al.'s (2013) suggestion that level 2 PT can be further divided into perspective-taking and perspective-confronting, with perspective-confronting being more challenging. Furthermore, results highlighted the significant role inhibition and verbal ability play in the completion of these tasks. Future longitudinal research

should be conducted to understand the directionality of the relationships between language, executive function, and a variety of PT and ToM tasks.

## References

- Apperly, I.A. (2008). Beyond Simulation–Theory and Theory–Theory: Why social cognitive neuroscience should use its own concepts to study “ToM”. *Cognition*, *107*, 266-283.
- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953-970.
- Ball, L. H., Besozzi, M., Ball, S. E., Anderson, L., & Geye, T. (2013). Strategies for deploying theory of mind in adults: Theory vs simulation is not either/or. *The Irish Journal of Psychology*, *34*(2), 81-92. doi:10.1080/03033910.2013.808160
- Barnes-Holmes, Y., McHugh, L., & Barnes-Holmes, D. (2004). Perspective-taking and theory of mind: A relational frame account. *The Behavior Analyst Today*, *5*(1), 15-25.  
doi:10.1037/h0100133
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, *21*, 37– 46. doi:10.1016/0010-0277(85)90022-8
- Bigelow, A. E., & Dugas, K. (2008). Relations among preschool children’s understanding of visual perspective taking, false belief, and lying. *Journal Of Cognition And Development*, *9*(4), 411-433. doi:10.1080/15248370802678299
- Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, *77*(1), B25-B31. doi:10.1016/S0010-0277(00)00096-2
- Bond, T.G., & Fox, C.M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.
- Borke, H. (1971). Interpersonal perception of young children: Egocentrism or empathy? *Developmental Psychology*, *5*(2), 263-269. doi:10.1037/h0031267

- Borke, H. (1975). Piaget's mountains revisited: Changes in the egocentric landscape. *Developmental Psychology, 11*(2), 240-243. doi:10.1037/h0076459
- Bradmetz, J., & Schneider, R. (1999). Is Little Red Riding Hood afraid of her grandmother? Cognitive vs emotional response to a false belief. *British Journal of Developmental Psychology, 17*, 501–514.
- Buttelmann, D., Carpenter, M. & Tomasello, M. (2009) Eighteen-month-old infants show false belief understanding in an active helping paradigm, *Cognition, 112*, 337–342.
- Call, J., & Tomasello, M. (1999). A nonverbal false belief task: The performance of children and great apes. *Child Development, 70*(2), 381-395. doi:10.1111/1467-8624.00028
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's ToM. *Child Development, 72*(4), 1032-1053. doi:10.1111/1467-8624.00333
- Carlson, S. M., Claxton, L. J., & Moses, L. J. (2015). The relation between executive function and theory of mind is more than skin deep. *Journal of Cognition and Development, 16*(1), 186-197. doi:10.1080/15248372.2013.824883
- Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant And Child Development, 11*(2), 73-92. doi:10.1002/icd.298
- Carpendale, J. M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences, 27*(1), 79-151. doi:10.1017/S0140525X04000032
- Carpendale, J. M., & Lewis, C. (2015). The development of social understanding. In L. S. Liben, U. Müller, R. M. Lerner (Eds.), *Handbook of child psychology and developmental*

- Science, Vol. 2: Cognitive processes (7th ed.)* (pp. 381-424). Hoboken, NJ, US: John Wiley & Sons Inc.
- Carruthers, P. (1996). Simulation and self-knowledge: a defence of theory-theory. In P. Carruthers & P.K. Smith (Eds.), *Theories of theories of mind* (pp. 22-38). Cambridge, England: Cambridge University Press.
- Chandler, M. J., & Greenspan, S. (1972). Ersatz egocentrism: A reply to H. Borke. *Developmental Psychology*, 7(2), 104-106. doi:10.1037/h0033145
- Cigala, A., Mori, A., & Fangareggi, F. (2015). Learning others' point of view: Perspective taking and prosocial behaviour in preschoolers. *Early Child Development And Care*, 185(8), 1199-1215. doi:10.1080/03004430.2014.987272
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cook, C. R., Frye, M., Slemrod, T., Lyon, A. R., Renshaw, T. L., & Zhang, Y. (2015). An integrated approach to universal prevention: Independent and combined effects of PBIS and SEL on youths' mental health. *School Psychology Quarterly*, 30(2), 166-183. doi:10.1037/spq0000102
- Dennett, D.C. (1978). Beliefs about beliefs [P&W, SR&B]. *Behavioral and Brain Sciences*, 1, 568-570 doi:10.1017/S0140525X00076664
- Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development*, 85(5), 1777-1794.
- Ebert, S. (2015). Longitudinal relations between ToM and metacognition and the impact of language. *Journal of Cognition and Development*, 16(4), 559-586. doi:10.1080/15248372.2014.926272

- Fisher, N., & Happé, F. (2005). A training study of theory of mind and executive function in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 35*, 757–771.
- Fizke, E., Barthel, D., Peters, T., & Rakoczy, H. (2014). Executive function plays a role in coordinating different perspectives, particularly when one's own perspective is involved. *Cognition, 130*(3), 315-334. doi:10.1016/j.cognition.2013.11.017
- Flavell, J. H. (1972). An analysis of cognitive-developmental sequences. *Genetic Psychology Monographs, 86*(2), 279-350.
- Flavell, J. H. (1992). Perspectives on perspective taking. In H. Beilin, P. B. Pufall (Eds.), *Piaget's theory: Prospects and possibilities* (pp. 107-139). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly, 50*(3), 274-290. doi:10.1353/mpq.2004.0018
- Flavell, J. H., Shipstead, S. G., & Croft, K. (1978). Young children's knowledge about visual perception: Hiding objects from others. *Child Development, 49*(4), 1208-1211. doi:10.2307/1128761
- Frick, H., Strobl, C., Leisch, F., & Zeileis, A. (2011). *Flexible Rasch mixture models with package psychomix* (No. 2011-21). Working Papers in Economics and Statistics.
- Gallagher, S. (2006). The narrative alternative to theory of mind. In R. Menary (Ed.), *Radical enactivism: Intentionality, phenomenology and narrative: Focus on the philosophy of Daniel D. Hutto* (pp. 223-229). Amsterdam, Netherlands: John Benjamins Publishing Company. doi:10.1075/ceb.2.15gal

- Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition: An International Journal*, 17(2), 535-543. doi:10.1016/j.concog.2008.03.003
- Gallagher, S. (2015a). The new hybrids: Continuing debates on social perception. *Consciousness and Cognition: An International Journal*, 36, 452-465. doi:10.1016/j.concog.2015.04.002
- Gallagher, S. (2015b). The problem with 3-year-olds. *Journal of Consciousness Studies*, 22(1-2), 160-182.
- Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8(9), 396-403. doi:10.1016/j.tics.2004.07.002
- Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, 138(6), 1085-1108. doi:10.1037/a0028044
- Gordon, R.M. (1996). 'Radical' simulationism. In P. Carruthers & P.K. Smith (Eds.), *Theories of theories of mind* (pp. 11-21). Cambridge, England: Cambridge University Press.
- Green, B. F. (1956). A method of scalogram analysis using summary statistics. *Psychometrika*, 2179-88. doi:10.1007/BF02289088
- Harrell, F. (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Switzerland: Springer.
- Harwood, M. D., & Farrar, M. J. (2006). Conflicting emotions: The connection between affective perspective taking and theory of mind. *British Journal of Developmental Psychology*, 24(2), 401-418. doi:10.1348/026151005X50302
- Herschbach, M. (2007) The phenomenological critics of folk psychology: The case of false belief, paper presented at the *Berkeley-Stanford-Davis Graduate Student Conference*, 14 April 2007.

- Hiller, R. M., Weber, N., & Young, R. L. (2014). The validity and scalability of the Theory of Mind Scale with toddlers and preschoolers. *Psychological Assessment, 26*(4), 1388-1393. doi:10.1037/a0038320
- Kaufman, A.S., & Kaufman, N.L. (2004). Kaufman Assessment Battery for Children Second Edition. Circle Pines, MN: American Guidance Service.
- Kesselring, T., & Müller, U. (2011). The concept of egocentrism in the context of Piaget's theory. *New Ideas in Psychology, 29*(3), 327-345. doi:10.1016/j.newideapsych.2010.03.008
- Kloo, D., & Perner, J. (2008). Training theory of mind and executive control: A tool for improving school achievement?. *Mind, Brain, And Education, 2*(3), 122-127. doi:10.1111/j.1751-228X.2008.00042.x
- Laranjo, J., Bernier, A., Meins, E., & Carlson, S. M. (2014). The roles of maternal mind-mindedness and infant security of attachment in predicting preschoolers' understanding of visual perspective taking and false belief. *Journal of Experimental Child Psychology, 125*, 48-62. doi:10.1016/j.jecp.2014.02.005
- Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12-and 18-Month-Olds Point to Provide Information for Others. *Journal of Cognition And Development, 7*(2), 173-187. doi:10.1207/s15327647jcd0702\_2
- Mair, P., Hatzinger, R., & Maier, M. J. (2009). Extended Rasch Modeling: The R Package eRm. Retrieved from ftp://cygwin.uib.no/pub/cran/web/packages/eRm/vignettes/eRm.pdf
- Martin, J., Sokol, B. W., & Elfers, T. (2008). Taking and coordinating perspectives: From prereflective interactivity, through reflective intersubjectivity, to metareflective sociality. *Human Development, 51*(5-6), 294-317. doi:10.1159/000170892

- Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development, 45*(2), 357-366. doi:10.2307/1127956
- Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*(2), 622-646. doi:10.1111/j.1467-8624.2007.01018.x
- Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100. doi: 10.1006/cogp.1999.0734
- Moll, H., & Meltzoff, A. N. (2011a). How does it look? Level 2 perspective-taking at 36 months of age. *Child Development, 82*(2), 661-673. doi:10.1111/j.1467-8624.2010.01571.x
- Moll, H., & Meltzoff, A. N. (2011b). Perspective-taking and its foundation in joint attention. In N. Eilan, H. Lerman, & J. Roessler (Eds.), *Perception, causation, and objectivity. Issues in philosophy and psychology* (pp. 286-304). Oxford, England: Oxford University Press.
- Moll, H., Meltzoff, A. N., Merzsch, K., & Tomasello, M. (2013). Taking versus confronting visual perspectives in preschool children. *Developmental Psychology, 49*(4), 646-654. doi:10.1037/a0028633
- Müller, U., & Kerns, K. (2015). The development of executive function. In L. S. Liben, U. Müller, R. M. Lerner (Eds.), *Handbook of child psychology and developmental science, Vol. 2: Cognitive processes (7th ed.)* (pp. 381-424). Hoboken, NJ, US: John Wiley & Sons Inc.
- Müller, U., Liebermann-Finestone, D. P., Carpendale, J. M., Hammond, S. I., & Bibok, M. B.

- (2012). Knowing minds, controlling actions: The developmental relations between theory of mind and executive function from 2 to 4 years of age. *Journal Of Experimental Child Psychology*, *111*(2), 331-348. doi:10.1016/j.jecp.2011.08.014
- Müller, U., Sokol, B., & Overton, W. F. (1999). Developmental sequences in class reasoning and propositional reasoning. *Journal Of Experimental Child Psychology*, *74*(2), 69-106. doi:10.1006/jecp.1999.2510
- Mutter, B., Alcorn, M. B., & Welsh, M. (2006). Theory of Mind and Executive Function: Working-Memory Capacity and Inhibitory Control as Predictors of False-Belief Task Performance. *Perceptual And Motor Skills*, *102*(3), 819-835. doi:10.2466/PMS.102.3.819-835
- Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology*, *58*(2), 220-249. doi:10.1016/j.cogpsych.2008.07.002
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science*, *308*(5719), 255-258. doi:10.1126/science.1107621
- Osterhaus, C., Koerber, S., & Sodian, B. (2016). Scaling of advanced theory-of-mind tasks. *Child Development*, *87*(6), 1971-1991. doi:10.1111/cdev.12566
- Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA, US: The MIT Press.
- Perner, J. (2010). Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In P. A. Frensch & R. Schwarzer (Eds.), *Cognition and neuropsychology: International perspectives on psychological science (Vol. 1, pp. 241–261)*. Hove, UK: Psychology Press.

- Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). ToM finds its Piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development, 17*(3-4), 1451-1472. doi:10.1016/S0885-2014(02)00127-2
- Piaget, J., & Inhelder, B. (1956). *The Child's Conception of Space*. London: Routledge & Regan Paul.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a ToM? *Behavioral and Brain Sciences, 1*, 515-526. doi:10.1017/S0140525X00076512
- Rothbaum, F. (1979). Comprehension of the objectivity–subjectivity distinction in childhood and early adolescence. *Child Development, 50*(4), 1184-1191. doi:10.2307/1129347
- Schurz, M., Kronbichler, M., Weissengruber, S., Surtees, A., Samson, D., & Perner, J. (2015). Clarifying the role of theory of mind areas during visual perspective taking: Issues of spontaneity and domain-specificity. *Neuroimage, 117*, 386-396. doi:10.1016/j.neuroimage.2015.04.031
- Selman, R.L. (1973). A structural analysis of the ability to take another's social perspective: Stages in the development of role-taking ability. Paper presented at the meeting of the Society for Research and Child Development, Philadelphia.
- Tabachnick, B.G., & Fidell, L.S. (2007). *Using Multivariate Statistics* (5<sup>th</sup> ed.). Boston, MA: Pearson Education, Inc.
- Wardlow, L. (2013). Individual differences in speakers' perspective taking: The roles of executive control and working memory. *Psychonomic Bulletin & Review, 20*(4), 766-772. doi:10.3758/s13423-013-0396-1

- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655-684. doi:10.1111/1467-8624.00304
- Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development, 75*(2), 523-541. doi:10.1111/j.1467-8624.2004.00691.x
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128. doi:10.1016/0010-0277(83)90004-5
- Zhang, T., Shao, Z., & Zhang, Y. (2016). Developmental steps in theory of mind of typical Chinese children and Chinese children with autism spectrum disorder. *Research In Autism Spectrum Disorders, 23*, 210-220. doi:10.1016/j.rasd.2015.10.005

## Appendix A

### 1. Grass/Snow (Carlson & Moses, 2001)

*Materials:* Green square, white square, felt hand prints for children to place their hand

*Notes:* 1. Trial is considered self-corrected if child originally moves towards the wrong colour, but then changes directions and chooses the right colour. 2. Timing through the trials should be paced appropriately, leaving enough time for children to return to the red handprints each time. 3. Provide reminders if child is not returning to handprints.

Can you tell me what is the color of snow?

**Can you tell me what is the color of grass? Very good!** (Place green and white squares on table in front of child approximately 12-18 inches apart from one another)

Now, we are going to play a silly opposites game, where every time I say snow I want you to point to the green color (*point to green square*) and every time I say grass, I want you to point to the white color (*point to white square*).

Let's practice, first please put your hands on the red hand shapes on the table (*Tape handprints to table*). Please put both your hands back on the handprints each time. Ready? Grass. Snow. These are teaching trials, and instructions can be repeated if child does not understand.

Trials will be administered in the following fixed order:

Target Card	Response	Score
1. Grass	White	2 1 0
2. Snow	Green	2 1 0
3. Snow	Green	2 1 0
4. Grass	White	2 1 0
5. Snow	Green	2 1 0
6. Grass	White	2 1 0
7. Grass	White	2 1 0
8. Snow	Green	2 1 0
9. Snow	Green	2 1 0
10. Grass	White	2 1 0
11. Snow	Green	2 1 0
12. Grass	White	2 1 0
13. Grass	White	2 1 0
14. Snow	Green	2 1 0
15. Grass	White	2 1 0
16. Snow	Green	2 1 0

Total Score Grass/Snow: \_\_\_\_\_ / 32

Scoring: 0 for pointing incorrectly, 1 for self-correcting, 2 for correct pointing.

## Appendix B

### 2. Diverse desire Task (Wellman & Liu, 2004)

*Materials:* Mr. Jones figurine, and pictures of a cookie and a carrot

*Notes:* If child says they do not know, or want to choose more than one option, you can tell them to just choose one and to make their best guess.

**Here's Mr. Jones** (*hold up Mr. Jones*). **It's snack time, so, Mr. Jones wants a snack to eat. Here are two different snacks: a carrot and a cookie** (*place pictures down on table*). **Which snack would you like best? Would you like a carrot or a cookie best?**

Own desire:                carrot                cookie

Well, that's a good choice, but Mr. Jones really likes [opposite of own desire]. He doesn't like [own desire]. What he likes best are [opposite of own desire].

'So, now it's time to eat. Mr. Jones can only choose one snack, just one. Which snack will Mr. Jones choose? A carrot or a cookie?

Target answer:                carrot                cookie

**What snack does Mr. Jones like best?**

Control question:                carrot                cookie

Score:                0                1

(0 if target is same as own answer, or if target and own are opposite but control is wrong, 1 if target is opposite of own answer and control question correct)

## Appendix C

### 3. Turtle Task (Masangkay et al., 1978)

*Materials:* Picture of a turtle, cup

Use a cup and demonstrate what right side up and upside-down means. See this cup? When I place the cup like this (*right side up*), it is right side up, but when I place the cup like this (*upside down*), now it is upside down. (*Place cup right side up again*) Is the cup upside down or right side up? (*Put cup away*). If child answers incorrectly, explain again and then move on.

Did child understand cup demo?                      Yes                      No

Place a card with a picture of a turtle on the table between the child and the experimenter. The child sits opposite the experimenter facing him/her.

**Look at the picture of this turtle. Is the turtle lying upside down or right side up?** This is a teaching trial. If child answers incorrectly, you can explain the answer and ask the question again.

Circle child's response:                                      upside down                                      right side up

If child isn't able to say "upside down" or "right side up", ask: (and circle response)

Is the turtle lying on its back or standing on its feet?

Ok, let's switch chairs now.

Experimenter and child switch seats so that the child sits where the experimenter used to be sitting.

Now when you see the turtle, is he upside down or right side up?

Circle child's response:                                      upside down                                      right side up

If child isn't able to say "upside down" or "right side up", ask: (and circle response)

Is the turtle lying on its back or standing on its feet?

When I look at the turtle, from my seat, where you were sitting before, do I see the turtle upside down or right side up?

Child's response:                                      upside down                                      right side up

If child isn't able to say "upside down" or "right side up", ask: (and circle response)

Is the turtle lying on its back or standing on its feet?

**Score:**                      0                      1                      2

(0 if none of the previous two questions correct, 1 if one of the previous two questions correct, 2 if both of the previous two questions correct)

## Appendix D

### 4. Diverse belief Task (Wellman & Liu, 2004)

*Materials:* Linda figurine, pictures of a bush and a garage

*Notes:* If child says they do not know, you can tell them to give their best guess.

Here's Linda. Linda wants to find her cat. Her cat might be hiding in the bushes (*place image of bush on table*) or it might be hiding in the garage (*place image of garage on table*). Where do you think the cat is? In the bushes or in the garage?

Own belief:               bushes                               garage

Well, that's a good idea, but Linda thinks her cat is in the [opposite]. She thinks her cat is in the [opposite].

So where will Linda look for her cat? In the bushes or in the garage?

Target answer:           bushes                               garage

Where does Linda think her cat is?

Control question:       bushes                               garage

**Score:**               0                               1

(0 if target is same as own answer or if target and own opposite but control wrong, 1 if target is opposite of own answer and control correct)

## Appendix E

### 5. Perspective-taking Task (Moll, Meltzoff, Merzsch, & Tomasello, 2013)

*Materials:* yellow-tinted screen, clear screen, 4 pairs of blue animal images, stickies for animals, bag

*Notes:* Have a second chair on your side for the child to come sit in.

We are going to play a game where you sometimes sit there (*gesture to child's seat*) and sometimes you sit here (*gesture to own seat*).

Colour Comprehension

\*(Bring out the 4 different coloured flowers and place them on table in the order: white, blue, yellow, and green from the child's left to right)

Here are 4 flowers. Can you show me the green/yellow/blue/white flower? Ask child in the order listed.

Got colours correct? \_\_\_\_\_

Child is encouraged to choose the flower. If child is wrong, point out the correct one and ask again. (**This one is the blue one! Now can you show me the blue one?**) This can be repeated twice if child is incorrect, then move on to next colour.

At the end when all 4 flowers were shown, announce how game goes:

Let's play a game. The game goes that you put the things I like in my bag. This my bag right here. (*Place bag on edge of table in between yourself and child*). Now let's see! I like the white one. Can you put it in the bag for me? I like the yellow one, too. Can you put that one in the bag for me too?

These are teaching trials, children can be corrected if get wrong and asked to try again.

Put white and yellow flowers in bag? \_\_\_\_\_

DEMONSTRATION

Watch, I want to show you something.

\*\*\*(*Get out the yellow and clear screen and place them on the table directly beside one another, with yellow screen on child's RIGHT side*).\*\*\*

**Look, I'll put this here!** (Take out the blue dog from the first pair. Place the dog on the child's side of the screen, on the side close to her own body)

I'll put this here! Now come around the table so you can see what it looks like from here!

Child moves around the table and looks at scene from experimenter's perspective.

Experimenter moves object back and forth quickly three times:

**Look it's like this** (hold dog up behind yellow screen with green flower colour held up as well) **and now look, it's just like this** (move to clear screen and show blue flower color).

**1 more time! Watch, it's like this again** (hold up dog behind yellow screen, holding up green flower) **wow, now it's like that again!** (hold up dog behind clear screen, holding blue flower)

**Now I want to show you something else!** 2 x gradual change: Move dog back and forth very slowly close behind the screens.

**Look, it's like this,** (held behind yellow screen) **and now look** (slowly move to clear screen), **isn't that cool?**

**Can you please come back around here!** (gesture to child's original seat and then move green/blue flowers off of the table, and place dog behind yellow screen)

TEST PHASE

\*\*\* (Rotate screens 180 degrees so that yellow screen is on child's LEFT side)\*\*\*

**I'll put this here!** (Take out second blue dog and places it behind clear screen)

Experimenter exclaims:

Oh, look [child's name]! Look at that DOG, that one looks GREEN!! I really like the GREEN dog. Can you please put the GREEN dog one in the bag for me? Gazing up towards child but not directly at either animal. (*Hold out the bag to the child*). Child makes her choice and places object in the bag.

Child's choice 1:      blue                      green

(Remove the not chosen animal and bring out pair of horses)

Watch, I'll put this here! And I'll put this here!

----- RESPONSE PHASE AGAIN-----

Oh, look [child's name]! Look at that HORSE, that one looks BLUE I really like the BLUE horse. Can you please put the BLUE horse one in the bag for me? Gazing up towards child but not directly at either animal. (*Hold out the bag to the child*). Child makes her choice and places object in the bag.

Child's choice 2:      blue                      green

(Remove the not chosen animal and bring out pair of bears)

Watch, I'll put this here! And I'll put this here!

Oh, look [child's name]! Look at that BEAR, that one looks BLUE!! I really like the BLUE bear. Can you please put the BLUE bear one in the bag for me? Gazing up towards child but not directly at either animal. (*Hold out the bag to the child*). Child makes her choice and places object in the bag.

Child's choice 3:      blue                      green

(Remove the not chosen animal and bring out pair of rabbits)

Watch, I'll put this here! And I'll put this here!

Oh, look [child's name]! Look at that RABBIT that one looks GREEN!! I really like the GREEN rabbit. Can you please put the GREEN rabbit one in the bag for me? Gazing up towards child but not directly at either animal. (*Hold out the bag to the child*). Child makes her choice and places object in the bag.

Child's choice 4:      blue                      green

**Score:** \_\_\_\_\_ (out of 4 for how many they got correct)

## Appendix F

### 6. Perspective-confronting task (Moll et al., 2013)

*Materials:* yellow-tinted screen, 4 blue animals, stickies for animals, blue/green flowers

*Notes:* Place blue and green flowers in front of child, to their left side

Children are sitting in front of a yellow tinted screen. A blue DOG is placed on the child's side of the screen so the dog appears blue to the child. The experimenter sits on the other side of the yellow screen so the dog appears green to them.

**“How do I (*point to self*) see the DOG - like this or like this?”** (pointing at the corresponding coloured blue swatch, and then green swatch).

Child's response (point):      blue                      green

**“How do you (*point to child*) see the DOG - like this or like this?”** (pointing at the corresponding coloured blue swatch, and then green swatch).

Child's response (point):      blue                      green

---3 MORE TRIALS---

A blue HORSE is placed on the child's side of the screen so the horse appears blue to the child. The experimenter sits on the other side of the yellow screen so the horse appears green to them.

**“How do you (*point to child*) see the HORSE - like this or like this?”** (pointing at the corresponding coloured blue swatch, and then green swatch).

Child's response (point):      blue                      green

**“How do I (*point to self*) see the HORSE - like this or like this?”** (pointing at the corresponding coloured blue swatch, and then green swatch).

Child's response (point):      blue                      green

A blue FISH is placed on the child's side of the screen so the fish appears blue to the child. The experimenter sits on the other side of the yellow screen so the fish appears green to them.

**“How do I (*point to self*) see the FISH - like this or like this?”** (pointing at the corresponding coloured blue swatch, and then green swatch).

Child's response (point):      blue                      green

**“How do you (*point to child*) see the FISH - like this or like this?”** (pointing at the corresponding coloured blue swatch, and then green swatch).

Child's response (point):      blue                      green

A blue RABBIT is placed on the child's side of the screen so the rabbit appears blue to the child. The experimenter sits on the other side of the yellow screen so the rabbit appears green to them. **"How do you (*point to child*) see the RABBIT - like this or like this?"** (pointing at the corresponding coloured blue swatch, and then green swatch).

Child's response (point):     blue                             green

**"How do I (*point to self*) see the RABBIT - like this or like this?"** (pointing at the corresponding coloured blue swatch, and then green swatch).

Child's response (point):     blue                             green

**Score:** \_\_\_\_\_(out of 8 for each question asked)

## Appendix G

### 7. False Belief task (Wellman & Liu, 2004)

Children see a toy figure of a boy and a sheet of paper with a backpack and a closet drawn on it. “Here’s Scott. Scott wants to find his mittens. His mittens might be in his backpack (*place image of backpack down*) or they might be in the closet (*place image of closet down*). Really, Scott’s mittens are in his backpack. But Scott thinks his mittens are in the closet.”

“So, where will Scott look for his mittens? In his backpack or in the closet?”

Child’s target answer:            backpack                    closet

“Where are Scott’s mittens really? In his backpack or in the closet?”

Child’s control answer:        backpack                    closet

Score:            0                    1  
(0 if any or all of the two questions wrong, 1 if both questions correct)

## Appendix H

### **8. K-ABC expressive vocabulary subtest**

Materials: K-ABC stimulus book and K-ABC scoring sheet

“Now I’m going to show you some pictures, and I just want you to tell me what those pictures are of.”