Running head: THEORY OF MIND AND SPECIFIC LANGUAGE IMPAIRMENT

# Theory of Mind and Behaviour Disorder in Children with Specific Language Impairment

Alana M. Holmes ©

A dissertation submitted in conformity with the requirements for the degree of

Doctor of Philosophy, Clinical Psychology

Graduate Department of Psychology

Lakehead University

0-612-64711-0

## Acknowledgments

Thank you to the parents who allowed their children to participate in this study. Without such interested parties and open minds it is hard to move forward in our understanding of children and their development. Special thanks to all the children who came and "played games" with me - I really enjoyed it and hope you did too! The quiet and steady support of my dissertation advisory committee was quite agreeable as it enabled me to complete this process in my own way. Appreciation is extended to my friends, family, and especially one Lennie Markwick, who freely gave me the time to complete this work and shared the belief that I could complete this work. Finally, I am indebted to those professionals who took a personal interest in this study and as a result were the ones to "make it all happen": the Speech-Language Pathologists, School Principals, Superintendents, and Program Managers who helped me maneuver through each institution's "red tape" and took time out of their busy days to orient me to each facility.

# Table of Contents

Tables

Abstract

The assertion that children with Specific Language Impairment (SLI) do not experience delays in the development of a theory of mind (ToM) has been made by numerous researchers (Eisenmajer & Prior, 1991; Leslie & Frith, 1988; Perner, Frith, Leslie, & Leekam, 1989; Peterson & Siegal, 1997). Such claims are premised solely upon the results of two studies (Leslie & Frith, 1988; Perner et al., 1989), both of which suffer from design weaknesses. The present study redressed the weaknesses of past research by administering a broad battery of ToM tasks to appropriately aged children and a measure of language ability beyond that of simple vocabulary. The study extended past research and examined the role of working memory and siblings on the ToM performance of SLI children. The relationship between ToM ability and language ability as a means to understanding the elevated incidence rate of behavioural disorder in SLI children was also explored. Forty nine normally developing and 43 SLI children ranging in age from 4- to 7-years completed first- and second-order ToM tasks, the Information subtest from the Wechsler Intelligence Scale, and the Linguistic Concepts subtest from the Clinical Evaluation of Language Fundamentals test. Subjects' parents completed the Child Behaviour Checklist. Results clearly indicated a delay in SLI children's acquisition of ToM. Their performance deficit extended a the nonverbal ToM task and suggests that language has a role to play both in the *conceptualization* and *expression* of ToM understanding. Working memory predicted ToM performance for SLI children but not that of their peers while number of siblings was not correlated with any ToM measure. Behavioural symptomology as rated by parents clustered around the normative mean and did not differentiate groups. The importance of language ability and processing capacity are discussed in relation to children's ability to understand other minds.

# Chapter 1 - Introduction

## Theory of Mind

### The Concept

Theory of mind (ToM) is often explained as "the ability of children to attribute mental states (such as beliefs, desires, intentions, etc.) to themselves and to other people, as a way of making sense of and predicting behavior" (Baron-Cohen, Tager-Flusberg, & Cohen, 1993, p. 3). Described another way, ToM refers to an individual realizing that knowledge is limited by experience and that others' actions are based on *beliefs*, not *facts*. The usefulness of this ability to help decode and comprehend daily and common social situations is clear. It helps individuals understand and predict the behaviour of others. For example, possessing a theory of mind can generate answers to something as simple as: "Why did the girl look in the basket? Because she believed her toy was in there and she wanted to play with it." (Baron-Cohen, Leslie, & Frith, 1986, p.114). More importantly, a theory of mind helps us to understand these actions even if we know that the girl's toy is elsewhere; the girl is looking in the basket because *she thinks* the toy is there. In other words, we know that the girl's behaviour is guided by her beliefs whether they are correct or "false". Generally, then, theory of mind skills might be considered essential to children's social understanding.

### Normal Developmental Time Lines

Given the above definition and stated utility of a theory of mind, this skill may be construed as a developmental milestone of sorts. Indeed, ToM has been the focus of much research within the realm of developmental psychology in recent years (Astington, Harris, & Olson, 1988; Dunn, 1995; Jenkins & Astington, 1996; Keenan, Olson, & Marini, 1998; Leslie, 1987; Wellman & Inagaki, 1997). The body of gathered evidence demonstrates that 2- and 3-

year-olds appreciate that others have desires and thoughts. They are also able to use correct mental state language, but they do not understand that another person may have a belief about the world which is different from their own belief, and different from reality until approximately the fourth year of life. It is now widely accepted that normally developing children are able to acquire a ToM beginning as early as 4 years of age, and that all normally developing children complete the feat by 6 years of age (see Astington et al., 1988 for a review; see also Wimmer & Perner, 1983; Yirmiya, Erel, Shaked, & Solomonica-Levi, 1998). Achieving an understanding of others' minds is not considered to be an "all or none" process, but rather, an ability that develops gradually during the preschool years (Astington & Jenkins, 1995; Dunn, 1995). The unfolding of this process coupled with the variation inherent in normal development likely explains the two year window spanning from ages 4 to 6. The timing of this accomplishment may also be accounted for by such things as neurological maturity, as well as, increasing verbal ability and working memory capacity.

Measurement

ToM ability has typically been measured by using one or more of a variety of tasks within an experimental setting. At least seven different tasks have been used to this effect according to a recent review of the literature (Yirmiya et al., 1998). All the tasks try to determine a child's ability to understand the contents of another person's mind without giving the child explicit access to, or explicit statements regarding, the other person's perspective. Such an ability can be observed by having the child privy to all stages of a hiding task (including an unexpected change of location), having a confederate experimenter present for only some stages of the hiding, and then having the lead experimenter question the child about *where* the confederate believes the hidden object to reside. Often two dolls are used in the stead of two experimenters. The dolls were originally named Sally and Ann and thus this most common ToM task is referred to as "the Sally-Ann task". The Sally-Ann task is a false belief task of the first-order. It is important to distinguish between the level of attribution being tested with a false belief task. First-order false belief tasks require the child to think about another person's thoughts about an objective event. Second-order false belief tasks go further and require the child to think about another person's thoughts about a third person's thought about an objective event (Baron-Cohen, 1989).

Many variations of the standard first-order false belief task described above exist. Some require predicting behaviour, some require explaining behaviour, and some even use a purely visual format to present and respond to the information thus negating the need for a verbal response. Perhaps the most common variant of this first-order false belief task is to extend it by incorporating the concept of ignorance within the paradigm. This is accomplished by adding a question geared to determine if the child *knows* that the confederate experimenter does not know where the hidden object is located. Thus, two sample questions, representing two separate first-order insights into the contents of another mind are: "Does Sally *know* where the marble is hidden?" (ignorance), and, "Where does Sally *think* the marble is hidden?" (false belief).

The ability to understand the perspectives of other minds has also been investigated by requiring children to successfully lie or deceive a third party in order to attain an enticing reward. This type of task requires the ability to intentionally manipulate another person's knowledge and beliefs (Sodian & Frith, 1992). The experimental scenario might run as follows: a child tells a third party that a box with a sweet in it is locked when in reality it is open, thus preventing the third party from trying to open the box and thus ensuring the sweet will be solely available to the child.

Another commonly used ToM paradigm is the "Smarties task" which uses the "deceptive-appearance" paradigm (Perner, Leekam, & Wimmer, 1987). Briefly, in this task, children are shown a Smarties box and asked about its contents. Invariably, the children reply "Smarties" and invariably they are surprised when the experimenter opens the box and reveals its contents - pencils or some such substitution. From here the children are asked to predict another person's response to the original question ("What is in this box?").

Despite the wide variety of ToM tasks described within the literature, no one particular task has proven best or even better at measuring ToM. The various tasks appear to be roughly equal in terms of their degree of difficulty (with the exception of the pre-stated difference between first- and second-order false belief tasks). A recent meta-analysis revealed that type of ToM task was generally not found to moderate ToM performance (Yirmiya et al., 1998). Similarly, Jenkins and Astington (1996) found no significant difference in the degree of difficulty among two versions of the standard "change of location" paradigm and two versions of the "deceptive-appearance" paradigm.

The wide variety of ToM tasks also reflects the fact that ToM is a multi-faceted construct comprised of at least two core components: belief understanding and desire/emotion understanding (Wellman & Bartsch, 1988). Some researchers see socio-emotional understanding as an additional component of ToM (Dunn & Brown, 1994). ToM can also be divided into understanding of representations, beliefs, motivational states, and emotional states. What the literature is emphasizing then with its variety of ToM tasks is that this is a developmental milestone which is not likely to be acquired all of a piece. The different aspects of ToM may develop at different rates, but, eventually combine to result in the tremendous accomplishment of understanding other people in all of their complexity and behavioural variability (Astington & Jenkins, 1995).

What Are The Benefits of an "On-Time" Theory of Mind?

Relatively little is known about the consequences of differences in children's understanding of either emotions or mental states. Investigation into this important line of research has a short history. Still, evidence has begun to accumulate which demonstrates that some aspects of social interaction are associated with performance on false belief tasks. For example, Frith, Happé, and Siddons (1994) investigated the real life social adaptation of autistic, mentally handicapped, and normally developing children as measured by caregiver report on the Vineland Adaptive Behavior Scales (VABS; Sparrow, Balla, & Cicchetti, 1984). The Vineland Adaptive Behaviour Scales were supplemented with a list of statements thought to reflect social behaviour requiring theory of mind (Interactive items) and a list of statements thought to reflect social behaviour not requiring theory of mind and which could be learned (Active items). The normally developing children in the study had an age range of 2 years 9 months to 7 years 4 months (mean age of 4 years 2 months); 60% of this group passed the first-order false beliefs tasks. Interesting differences were found within their scores on the supplemental Interactive and Active items. Normal subjects who comprehended the existence of other minds ("passers") were significantly more inclined to simple sociability (Active) and non-significantly more likely to demonstrate everyday social insight (Interactive) than their peers who did not comprehend the existence of other minds ("failers"). Here then, we have evidence that children who have a timely understanding of the presence and perspective of others' minds appear to be more socially perceptive and more socially orientated than children who have yet to master this concept.

Dunn (1995), using a sample of 46 children aged 3 years 3 months, explored if, and how, individual differences in children's understanding of emotions and of other minds were related to later differences in their social understanding. Understanding of other minds was assessed using a series of five false belief tasks requiring the child to "explain" a puppet character's behaviour based on the character's false belief. The children who demonstrated good ToM ability at 3 years 3 months of age differed from those showing poor ToM ability in several ways. Children with an understanding of other minds were more likely than those without such an understanding to describe some difficulties with the persons populating their kindergarten environment and the work given them within this environment. The study's results were tentatively interpreted as indicating that children with an early understanding of others' minds "may be particularly sensitive to, and aware of others' judgments and suffer accordingly" (Dunn, 1995, p. 198).

This is a fair interpretation; however, another, equally valid, can be offered. Because the false belief tasks required children to "explain" and not "predict" behaviour, perhaps what is measured, more than an understanding of other minds, is language ability or development. Perhaps "passers" on the false belief tasks are children with a predilection for and sensitivity to language. This would explain why these same children are more likely upon entering kindergarten to take to heart a teacher's verbal criticism (negative judgment) than those children less geared to the nuances of language. Notably, language ability was not measured by Dunn (1995) leaving us to wonder about its role in ToM success and in later social understanding. Regardless of the mechanism underlying this displayed difference, the results do indicate that an early understanding of others' minds (at 3 years 3 months of age) is associated with heightened sensitivity to negative appraisals of interpersonal situations as early as the beginning of kindergarten (at 5 years of age).

Other connections between theory of mind development and children's social interaction have also been reported in the literature. Astington and Jenkins (1995), in a sample of 30 children 3 to 5 years of age, found that composite performance on four false belief tasks was associated with displays of joint proposals and explicit role assignments during a 10 minute session of unstructured pretend play. Specifically, those children with higher levels of false belief understanding showed significantly more joint proposals in their pretend play and made more explicit pretend role assignments to both themselves and other children. In other words, performance on false belief tasks was linked to real world behaviours. That is to say, children's social interactions reflect to some degree differences in theory of mind development. These results were produced in a sample of children matched for age and linguistic competence, but, differing in theory of mind development. Were we to extrapolate, we might infer that early attainment of a theory of mind can be associated with leadership or extroverted qualities in young children.

Finally, a relation between children's developing theories of mind and aspects of their social-emotional maturity has been identified by Lalonde and Chandler (1995). The achievement of forty 3-year old children on six measures of false belief understanding was compared to their social-emotional skills and behaviours as rated by their pre-school teachers. Social-emotional maturity was assessed with a 40-item questionnaire half of whose items were deemed "Intentional" (thought to require some insight into the mental lives of others), and half "Conventional" (thought to reflect a simple grasp of social conventions or the exercise of self-control). The measure of false belief understanding proved to be positively correlated with the "Intentional" behaviours. These results led Lalonde and Chandler to conclude that "the consequence of early insights into other people's mental lives is to selectively influence just those aspects of interpersonal functioning that were judged to turn on the achievement and maintenance of an intentional stance" (pp. 180-181). Success on ToM tasks was correlated with such social behaviours as: engages in simple make-believe activities with others, converses with others on topics of mutual interest, able to comment on differences between his or her wishes and those of another, explains rules of game or activity to others, and able to comment on differences between his or her own feelings and those of another.

What Influences the Development of an "On-Time" Theory of Mind?

Research into factors influencing the development of a theory of mind is scant. Several authors have mused that perhaps this line of pursuit has been disregarded in favor of the challenge to establish the timing and window of ToM development (Flavell & Miller, 1998; Jenkins & Astington, 1996; Taylor, 1996). The very recent and very limited literature exploring determinants of theory of mind development has unearthed several factors associated with individual variation in ToM attainment. They are: siblings, general language ability, and working memory span.

Siblings. Perner, Ruffman, and Leekam (1994) found that number of siblings was related to theory of mind performance in a sample of 76 children 3 to 4 years old. Children with two siblings were shown to be approximately twice as likely to pass a ToM task (false belief) than were children without siblings. The linguistic and cognitive abilities of the children under study were not measured and so the role of language and general learning ability could not be teased apart from that of number of siblings.

Jenkins and Astington (1996) examined the relationship of language, memory, and family size relative to false belief understanding in normally developing 3- to 5- year-olds, using four different false belief tasks. The results showed that children with a large number of siblings outperformed children with a small number of siblings even after the effects of age and language were partialled out. The number of siblings remained important regardless of whether they were older or younger than the examinees or how far distant in age they were from the examinees. More interesting still, number of siblings appears to be especially important for the development of ToM in children with poor language skills. The effect of family size was found to be more strongly associated with false belief understanding in the less linguistically adept children than in the more verbally competent children.

A study conducted by Dunn, Brown, Slomkowski, Telsa, and Youngblade (1991) tapped into the effects of both family (though not number of siblings per se) and language. They found an association between some aspects of *family discourse* when children were 2 years 9 months of age and their level of false belief understanding seven months later. The children who performed better on false belief tasks wherein they were required to explain action based on false beliefs at 3 years 3 months of age shared the following characteristics at 2 years 9 months of age. They talked more about feeling states, spoke more with their mother about causal relations, cooperated more with older siblings, and observed more controlling talk between mother and siblings. Simple exposure to other children was not considered to have an important effect as most of the children in the study attended nursery or day care settings for a significant portion of their day. Thus, it seems that there is something unique in the intimate interactions that occur between siblings and other family members which serves to enhance the development of a theory of mind.

Indeed, children have been shown to differ in their interactions with familiar older children and older siblings under unstructured and structured situations. By the end of an unstructured building task, 7-year old children showed increased consultation and imitation as well as improved performance when paired with their older sibling than when working with a familiar older child (Azmitia & Hesser, 1993). Also, the older siblings were noted to be more likely than the older friends to give spontaneous guidance and explanations to the 7-year old child.

The relationship between siblings and ToM performance is not entirely clear however. For example, studies by Anderson (1998) and Cutting and Dunn (1999) failed to find a relationship between number of siblings and metarepresentational ability in two samples of young children.

Language. Ruffman, Slade, Clements, and Import (1999) recently stated, "although it is widely accepted that language relates to ToM, the nature of this link remains vague" (p.3). Indeed, most work on the influence of language ability on the ToM ability of normally developing children has been fuzzy in nature. For instance, Brown, Donelan-McCall, and Dunn (1996) found "conversational mental state language" in 4-year-olds was correlated with their level of false belief understanding. Dunn et al. (1991) found a positive correlation between children's "conversational language" about feelings at 2 years 9 months of age and their level of emotional understanding at 3 years 3 months of age. The importance of language in its conversational form was then suggested. However, the correlation between talk about feelings and ToM performance was independent of the child's general verbal ability and the quantity of talk within the family. Thus, the precise role of language remains unclear.

Such indirect findings make it appear as if the research community has assumed a relationship between language and ToM prior to garnering evidence of such a relationship. This state of affairs was only recently addressed. The findings of Jenkins and Astington (1996) demonstrated a relationship between false belief understanding and achievement on the Test of Early Language Development (TELD; Hresko, Reid, & Hammill, 1981). A threshold effect for language was revealed in that a certain level of linguistic ability (raw score of 14 on the TELD) was required before children could pass the false belief tasks. The TELD is a general measure of language ability which assesses syntactic and semantic skills as well as receptive and expressive ability of children 3 to 7 years of age. The demonstrated link then is between overall language skills, not just vocabulary, and ToM. The important finding is that a certain level of *language competence* appears necessary in order to succeed on ToM tasks.

The notion of language competence being more important than mere vocabulary is supported by a study investigating ToM in autistic persons ranging in age from 6 years 11 months to 22 years 2 months. A strong relationship was found between ToM performance and skills of syntactic comprehension. Moreover, subjects' scores on the syntactic comprehension subtest were better predictors of ToM performance than were scores on a receptive vocabulary test (Peabody Picture Vocabulary Test (PPVT); Dunn & Dunn, 1981) (Tager-Flusberg & Sullivan, 1994a).

Ruffman et al. (1999) looked at ToM in relation to different language tasks. Their findings also support a link between ToM and general language ability, but, not between ToM and the specific language components of receptive or expressive vocabulary. As well, the results showed that early ToM ability was as good a predictor of subsequent language ability as early language ability was a good predictor of subsequent ToM. The authors' concluded that "the most well replicated finding in the ToM literature - ToM improves with age - seems largely a product of language ability" (Ruffman et al., 1999, p.21). This conclusion is backed up by the findings of a study conducted by Astington and Jenkins (1999) which showed larger and more consistent correlations going from language to ToM than from ToM to language.

Summing up the scant results then, we may conclude that language development seems to precede and assist ToM development.

Short -term memory and working memory. A role for short-term memory (STM) or working memory (WM) in ToM would appear to be obvious. However, Jenkins and Astington (1996) found that short-term memory did not contribute to the variance in ToM performance any more than did a general measure of language. This result was reached after assessing both a verbal and nonverbal measure of short-term memory within a sample of 3- to 5-year-olds. Nonverbal short-term memory made no significant contribution, while the verbal short-term memory appeared to be confounded with or to be sharing the same variance as the more general measure of language. The possible confound lay in the use of the Memory for Sentences subtest of the Stanford-Binet (SB:IV; Thorndike, Hagen, & Sattler, 1986) as a measure of short-term verbal memory. This subtest is highly correlated with overall verbal reasoning ability ($r = .64$ for 3-year-olds; Thorndike et al., 1986). It may also be the case that what is being demonstrated here is a replication of a well established finding in the information-processing/memory literature: STM *capacity* is necessary but not sufficient to ensure correct reasoning in cognitive tasks (Halford, Maybery, & Bain, 1986; Wimmer & Perner, 1983).

The ability to *actively process* information is more likely to be an indicator of ToM performance. Any given ToM task requires holding several pieces of information in mind and then actively processing and reconsidering their temporal location in response to standard questions about perspectives at various points in time. This line of reasoning, if accepted, may explain why working memory tasks that tap into a "central executive" (i.e., require active processing) have been demonstrated as being positively correlated with successful ToM performance. According to the Baddeley (1981) model of working memory, tasks such as digit span forward or memory for sentences would directly measure articulatory loop capacity yet would only indirectly measure central executive capacity, since correct responses require efficient encoding and retrieval, however they require no manipulation of information. By contrast, tasks such as backwards digit span provide a more direct measure of central executive capacity due to the required manipulation of an increasing number of pieces of information.

Davis and Pratt (1995) cite evidence for the role of working memory in children's success with false belief understanding tasks. Within their sample of 54 children 3 to 5 years of age, scores on a backwards digit span task predicted performance on a false belief task. Six percent of the variance was uniquely accounted for by this measure of working memory span above and beyond the variance accounted for by age and language ability.

Keenan (1998), provides evidence to suggest that when a more sensitive measure of working memory is employed, such as a modified version of the Counting Span task (see Case, Kurland, & Goldberg, 1982), working memory assumes a substantial role in accounting for the variance in performance on ToM tasks. Keenan (1998) found that when the effects of age and language were controlled, working memory span accounted for 21% of the variance in the false belief understanding in a sample of normally developing children 4 to 5 years old.

## Theory of Mind - Delayed Developmental Time Lines

A delayed theory of mind has been found in two special populations of children: deaf children and children with Autistic Disorder. The research for each population will be discussed in its turn.

### Theory of Mind and the Deaf - The Literature to Date

Language. This is a very new area in the ToM research literature.

Peterson and Siegal (1995) were the first to suspect and then demonstrate a delayed ToM in deaf children. A slightly modified version (signed, acted out, and requiring only a nonverbal pointing response) of a standard false belief task (Sally-Ann) was administered to a sample of 26 signing, prelingually-deaf children with normal intellectual ability, all between the ages of 8 and 13. The Sally-Ann task is routinely passed by normally developing children between the ages of 4 and 5. However, the majority of this sample of deaf children failed and only 35% of the deaf children with normal intellect and a chronological age above 8 were able to pass this basic task. Russell, Hosie, Gray, Scott, and Hunter (1998) reported a similar delay of several years in deaf children's ability to deduce the thoughts and motives of others. Their study found that only 14% of deaf 4- to 12-year-olds were capable of passing a modified false belief task (signed and acted out and requiring only a nonverbal pointing response).

Beyond this replication, the study by Russell et al. (1998) also revealed a burgeoning in deaf children's understanding others' minds after the age of 13. Sixty percent of their sample of deaf children aged 13 to16 passed the theory of mind task. The authors suggest that their data might reflect the fact that not only is a certain amount of linguistic competence necessary to pass ToM tasks, but that ample opportunity to work with this level of language within social interactions is also required. Due to the restricted opportunities deaf children have for learning about mental states in their largely silent and highly concrete worlds, they may require more time, in fact, years, to gather sufficient experience in this realm to be able to apply and demonstrate a ToM.

Family. Peterson and Siegal's original study (1995) also found evidence that was "consistent with the hypothesis that conversational exposure influences performance on tasks devised to test for theory of mind understanding" (p.469). A significant difference was noted between the ToM performance of deaf children from signing homes as compared to that of deaf children from non-signing homes. Superior performance was shown by deaf children from signing homes relative to their counterparts from non-signing homes: 100% versus 29%.

The importance of a signing home environment was confirmed in a later study which investigated the performance of deaf children from signing and non-signing homes on tasks designed to tap their naïve or folk theories in three distinct causal-exploratory reasoning systems: psychological (theory of mind), physical, and biological (Peterson & Siegal, 1997). As a whole, the deaf children from hearing families performed no better than autistic children on the ToM task. However, the results showed that a subgroup of deaf children from fluently signing households (at least one signing deaf conversational partner) exceeded the performance of all other signing deaf classmates. Eighty-nine percent of those with a signing relative passed the false belief task, while only 46% of those without a signing relative passed the false belief task. "This finding points strongly to early conversational experience at home as a determining factor in the acquisition of a theory of mind" (Peterson & Siegal, 1997, p. 66). Thus, the research on ToM development in deaf children supports the role of language and of family in acquiring a timely understanding of others' minds.

Short-term memory and working memory. To date there exists no research exploring the role of short-term memory or working memory in the ToM performance of deaf children. This might be a viable area of research given the mixed results documenting the importance of short-term memory and working memory in normally developing children's understanding of other minds.

Behaviour. The association between ToM delay and behaviour in deaf children has not been specifically explored. However, unrelated studies document behavioural and social abnormalities in deaf children. For instance, personality studies involving the deaf indicate that they have a general lack of emotional control and a social immaturity (Kusche, Garfield, & Greenberg, 1983). Deaf children and adolescents have also demonstrated consistent deficits in social interaction (Garrison, Emerton, & Layne, 1978) and in empathy development (Bachara, Raphael, & Phelan, 1980). While these behaviours could indeed be associated with a lack of ToM, it remains for future investigators to prove this connection.

Theory of Mind and Autistic Disorder- The Literature to Date

Language. It has been a slow process, but the studies investigating ToM ability in autistic children appear to be converging on the fact that verbal mental age or language ability plays a role in determining the achievement of ToM. Early studies such as Baron-Cohen (1989) and Leslie and Frith (1988) revealed a trend toward increased verbal ability in ToM "passers" but the trends usually failed to reach significance. It is likely that the small number of subjects involved in the early studies and their extremely diverse verbal mental ages (VMAs) were factors in this outcome.

A study conducted by Eisenmajer and Prior (1991) was one of the first to demonstrate a significant difference in verbal mental age between autistic "passers" and "failers" of a theory of mind task. These experimenters demonstrated with a sample of relatively able autistic subjects that when a certain level of verbal competence is reached, an autistic child becomes likely to succeed on first-order theory of mind tasks and thus display mentalizing ability. Having unearthed the influences of verbal mental age on ToM ability, Eisenmajer and Prior (1991) then make the point that verbal mental age was not the only factor regulating the demonstration of

ToM in autistic children. They also found in their study that some autistic children, despite having relatively high language skills, were unable to pass a standard first-order false belief task. Another pertinent finding revealed by Eisenmajer and Prior (1991) was that no autistic child below the chronological age of 8 years 7 months was able to pass the false belief task (in their study) and that this appeared similar to trends found in earlier studies. This suggests that not only is a certain level of language competence required in order to pass a ToM task, but, that perhaps ample opportunity to acquire skilled performance in these language skills in social environments is also needed to hone the skill.

Responding to the observation that high verbal ability in autistic subjects was associated with passing ToM tasks, Frith, Morton, and Leslie (1991) proposed that these children may be using verbally mediated routes (not used by other children) to answer these meta-representational questions. In other words, autistic subjects might be "hacking out" strategies to solve the ToM tasks and so require much more verbal skill in order to do so. It is indeed conceivable that "hacking out" could be successful in structured ToM tasks where elements of visual access and information are spelled out. However, it seems less likely that this approach would work in real life situations which are necessarily tinged with ambiguity. This hypothesis was tested using a variety of stories about everyday situations where people say things they dc not mean literally (Happé , 1994). The high-functioning autistic children had difficulty with the 12 naturalistic stories. The authors interpreted this as lending support to the idea that autistic children use something other than the usual approach to succeed on ToM tasks. But, perhaps it also lends support to the idea that children with late developing language need extra time and social opportunity to develop a true theory of mind. Or perhaps it demonstrates that sufficient language skills initially permit interpretation of theory of mind in highly cued experimental conditions and that the generalization of these skills to real life requires extra time and opportunity.

Happé (1995), after performing a meta-analysis on the results of several studies on theory of mind tasks employing small numbers of autistic, mentally handicapped, and normal young subjects, reached four conclusions regarding this literature. Of interest is the assertion that success on ToM tasks is indeed related to verbal level as determined by a picture card receptive vocabulary task. More interesting still is that the pooled data for normal children and for children with Autistic Disorder closely fit a two-threshold model in explanation of a pass or fail on ToM tasks. A two-threshold model based on verbal mental age implies that all children below a certain VMA fail while all children above a certain VMA pass standard first-order ToM tasks. Different VMA thresholds were found for autistic versus normal children. Young normals begin to pass the tasks with a VMA of 2 year 10 months, while autistic subjects require a VMA of at least 5 years 6 months before having a chance of passing the tasks. Autistic subjects thus require at least twice as much receptive vocabulary relative to normal children in order to comprehend the same concept. Again, we are tempted by the idea that when language is late to develop, extra time is required to maximize its potential - to be able to use it to interpret real-life interpersonal situations.

Recently, three large scale meta-analyses which compared the theory of mind abilities of individuals with Autistic Disorder, individuals with mental retardation, and normally developing individuals were completed (Yirmiya et al., 1998). This series of statistical studies reached several important conclusions. First, ToM deficits can no longer be considered unique to autism as they are also evidenced in persons with mental retardation to a significant degree. Second, the selection of comparison groups must be chosen carefully with regards to specifying their diagnoses as this factor was shown to moderate ToM performance. Third, care should also be taken in considering matching criteria as chronological age was identified as an important moderator variable when comparing all three groups of children. And, finally, ToM ability should be studied in different clinical groups in order that its varying components and their origins may be more completely understood (Yirmiya et al., 1998).

Short-term memory and working memory and siblings. To date there exists no research exploring the role of short-term memory or working memory or family size in the ToM performance of autistic children. This could be a viable area of research given the mixed results documenting the importance of short-term memory and working memory and siblings in normally developing children's understanding of other minds.

Behaviour. The relation between ToM ability (or lack thereof) and behaviour has not been well explored within the autistic population. The discovery of a ToM deficit in persons with Autistic Disorder supplied the field with a very tidy theoretical explanation for the core symptoms of the disorder. That is, autistic children's limited language, unusual behaviours, and social isolation could all be attributed to their more primary inability to recognize and understand the presence of other persons' minds and mental states. Logically this makes good sense. It is also a parsimonious explanation - the most coveted criterion of good research theories! Nevertheless, the fact remains that few attempts have been made to demonstrate empirically the association between impaired ToM and individual differences in daily life social behaviour in persons with autism.

Frith et al. (1994) examined the real life competence of 24 autistic children, one third of whom were found capable of passing two first-order false belief tasks. Real life social adaptation was assessed by caregiver report using the VABS and two supplemental scales comprised of

"Interactive" items (social behaviour thought to require a theory of mind) and "Active" items (social behaviour thought to be learned by rote). Significant differences were present between the Interactive scores of the ToM "passers" and the ToM "failers" but not between their Active scores. So, ToM ability as measured in experimental situations does have applied consequences in the social interaction behaviour of autistic children. It is linked with more instances of everyday social insight.

Prior et al. (1998) also found the variable of language ability to be important for success on ToM tasks in a large sample of children diagnosed with disorders from the autism spectrum. Their results suggest that when verbal ability is not too removed from "average", theory of mind deficits are less in evidence. Within this sample, performance on the ToM tasks was able to divide the autistic children into three clinically recognizable groups. The authors, then proposed that these children's levels of cognitive and language competence moderate the nature and severity of their behavioural symptoms. "Hence, we argue that the results of this research support the concept of a spectrum of autistic disorders in which severity of social and communicative impairments underlie individual differences in the cognitive, behavioural, and adaptive functioning deficits observed" (Prior et al., 1998, p.900)

Summary of Findings on ToM Development

Theory of mind development within normally developing children has been shown to occur between the ages of 4 and 6. Possession of an understanding of others' minds at this age, has also been shown to manifest itself in social behaviours that demonstrate leadership, extroversion, and sensitivity to negative judgment. Of the many factors believed to influence the development of ToM in normal children, the best documented to date are language, working memory, and number of siblings. As would be expected when exploring developmental skills and behaviours, that which is found to be influential within a normally developing population is also found to hold sway within an abnormally developing population. Thus, the literature on ToM development in deaf and autistic children contains similar findings regarding the factors thought to influence ToM development in normal children. Given the language and socialization limitations of deaf and autistic children, a ToM delay is hardly surprising. Exploring ToM development within another clinical group would serve to further document the role of language, working memory, and siblings, and to possibly strengthen the hypothesized link between ToM development and behaviour.

In identifying a potentially relevant clinical population in which to study ToM it makes inherent sense to consider shared characteristics as well as shared etiology. For example, although both deaf and autistic populations demonstrate language and socialization delays, the centrality of these traits to the respective disorder and the etiology behind them vary. That is to say, the language and socialization delays of autistic individuals are thought to be core symptoms of the disorder and to reflect an underlying neurological abnormality. The language and socialization delays of deaf children are however, considered to be secondary symptoms caused by the deafness, not by underlying neurological difficulties. Given this, it becomes most interesting to consider researching ToM development in a clinical population where the etiology behind the language impairment is believed to be neurological and the delays in socialization deemed secondary symptoms resulting from the poor language skills. The prime candidate for extending the research on ToM then is the population of Specific Language Impaired (SLI) children.

Specific Language Impairment - A Similar Special Population?

For a variety of reasons, children with Specific Language Impairment (SLI) present as prime candidates for furthering research on ToM development. Among the most important are the several key traits they share with children with deafness or autism: difficulty with language acquisition, impaired access to social opportunities, and patterns of abnormal behaviour.

Language impairment. The language impairment shared among these three clinical groups is of course not identical, but it is similar. Specific Language Impairment by definition means that the language impairment is primary in nature and not secondary to some other condition (Craig, 1993). Thus, the criteria for diagnosis are basically exclusionary in nature: no hearing loss or history of recurrent otitis media effusion, no significant emotional or behavioural problems, no mental retardation, no evidence of frank neurological problems, and no sensory or oral defects (Craig, 1993; Leonard, 1998; Stark & Tallal, 1981). A receptive or expressive or overall language score 1.25 standard deviations below the mean on a standardized language test completes the rather stringent definition of SLI. Articulation problems are usually defined as a speech disorder and so are distinguished from the more severely disabling language disorders. Deaf children have receptive and expressive language problems which are responsive to treatment and likely to improve to within near normal levels given time and exposure to sign language. The language delays of autistic children are well documented with such things as absence of speech, weak receptive language skills, and unusual use of language such as echolalia.

Diminished social opportunity. Children with SLI, much like autistic and deaf children, have difficulty integrating themselves into the social fabric that surrounds them. Their language limitations as well as their under-developed social behaviour (relative to same-aged peers) stand in the way of normal and appropriate social interaction. This can lead to limited excursion into the regular social channels and interactions which can in turn suppress the development of language and social skills. Indeed, the quantity and quality of social interaction experienced by SLI children has been documented to be markedly different from that of their linguistically able peers. For example, the comments of SLI children are responded to approximately half as much as those of normally developing children (Craig & Gallagher, 1986). Children with SLI are also ignored by their peers twice as much as other children and SLI children participate in fewer interactions with other children (Hadley & Rice, 1991). As well, SLI children seem to have fewer positive social interactions with their peers as they are more often interrupted (Wellen & Broen, 1982), and are less successful in entering into an activity already in progress (Rice, Sell, & Hadley, 1991). In essence a downward spiral is created which removes these children from everyday social exchanges.

It is likely that when deaf children are raised in an environment composed primarily of hearing persons, they will acquire most of their knowledge through language that is directed to *them*. As such they will generally be unable to benefit from incidentally overhearing communications between other persons in their environments. The same may be true of autistic children and SLI children though to a lesser extent. They are often unable to communicate well with their family members and so they may have difficulty learning to recognize attitudes which are mainly learned through language's more subtle attributes (e.g., innuendo, tone, and intonation). In essence, the case can be made that social, emotional, and perspective understanding is not a direct result of chronological maturation and increased number of life experiences, but, it is also influenced through socialization and language (Luria, 1976).

Behaviour. The concurrence of language problems and behaviour problems has been documented with increasing frequency within the psychological literature. What is accepted to date is that the two often occur together, they may begin early in life, and their presence is disturbing to normal development. Richman, Stevenson, and Graham (1982) found that 14% of the general population of 3-year-olds showed behaviour problems. When this same definition was applied to children with language delay, 59% of them were found to have a behaviour problem. Overall, the findings on behaviour problems in children with language delay are strong and significant: approximately half of children with language delay have been observed to demonstrate behaviour problems (Stevenson, 1996).

Considering the reverse relationship, language delays discovered in children identified with behaviour problems, reveals just how strong and how common the overlap is between these two childhood disorders. When a definition of language delay that identified 3% of the general population, was applied to the population of children with behaviour problems, language delay was found in 13% of the behaviour disordered children (Stevenson, 1996). Although the relationship does not appear to be as strong when viewed in the opposite direction, the overlap between the two domains is remarkable.

Other estimates provide a slightly higher incidence rate. For example, unsuspected language delay was discovered in 34% of a sample of 4- to 12-year old children referred for behavioural or emotional problems to a mental health centre in a large metropolitan city (Cohen, Davine, Horodezky, Lipsett, & Isaacson, 1993). And unsuspected language delay was revealed in 40% of a sample of 7- to 14-year-olds referred for psychiatric services in a large metropolitan city (Cohen, Barwick, Horodezky, Vallance, & Im, 1998). Language impairment was defined as one language test score two standard deviations below the normative mean or two language test scores one standard deviation below the mean.

Spectrum disorder. The investigation of ToM development in SLI children is further legitimized by recent changes in the theoretical conceptualization of autism. This childhood disorder is currently being conceptualized as existing on a continuum and as being expressed with varying degrees of severity (American Psychiatric Association [APA], 1994). Autism at its most extreme represents an individual with significantly depressed cognitive skills, no verbal skills, a repertoire of repetitive behaviours (often self-injurious in nature), and an apparent unawareness of the existence of other persons. The least severe form of autism is often categorized and dubbed Asperger's syndrome. Individuals with Asperger's syndrome have intact cognitive faculties and language skills, a restricted range of interests and behaviours, and fairly gross social difficulties. In theory, it is possible to extend this continuum further still to encompass the more severe cases of language impairment.

This theoretical extension has been understood by practicing clinicians: "It has long been recognized that high functioning individuals with autistic disorder and those with specific developmental language disorder or developmental dysphasia share a number of characteristics in common" (Konstanareas & Beitchman, 1996, p.178). Children with severe receptive and mixed receptive-expressive language disorders, in fact appear to represent a very mild presentation of autistic symptomology: a history of delayed language development, notable problems in understanding abstract concepts (better suited to understanding highly literal concepts), along with poorly developed imaginative play and social cognition (Konstanareas & Beitchman, 1996).

The conception of a broader phenotype in autism has also found support in family and genetic studies. Tanguay, Robertson, and Derrick (1998) found that the pragmatic language scores of the parents of autistic children were significantly more abnormal than those of control adults. Similarly, Piven, Palmer, Jacobi, Childress, and Arndt (1997) found higher rates of social and communication weaknesses, and stereotypic behaviours demonstrated in 25 families with multiple-incidence autism compared to families of Down syndrome children. As well, a twin study by Le Couteur et al. (1996) wherein one or both twins per pair had autism, determined that in monozygotic and dizygotic, same-sex twins discordant for autism, many of the non-autistic twins demonstrated language impairments and social deficits beginning in childhood and continuing into adulthood. The degree of demonstrated language impairments and social deficits has been found to be much less in the case of dizygotic twins discordant for autism (Folstein & Rutter, 1978; Le Couteur et al., 1996, Rutter, Bailey, Bolton, & Le Couteur, 1993).

Prior et al. (1998) gathered extensive data through parent interviews on the developmental history and current behaviour of 110 high functioning children and adolescents with diagnoses of autism, Asperger's, or related disorders, such as Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS). Cluster analysis resulted in three subgroups which differed on theory of mind performance and verbal ability. These results confirm the importance of ability and age variables in succeeding on ToM tasks. The three groups were roughly teased apart into the original diagnoses of autism, Asperger's, and "other" (such as PDD-NOS). One important finding was that the displayed behaviours per se did not differentiate the children, but, the *severity* of the behaviours did serve to separate the sample into three distinct groups. Overall, evidence is provided for the taxonomic validity of a "spectrum of autistic disorder on which children differ primarily in term of *degrees* of social and cognitive impairments" (Prior et al., 1998, p.893). Thus, some of the traits that SLI and autistic children do share can reasonably be viewed as existing on a continuum with the traits within the SLI population being displayed in a muted form.

Summary of Similarities Between SLI, Deaf and Autism

Difficulty with language acquisition, impaired access to social opportunities, and patterns of abnormal behaviour are traits shared among the special populations of deaf, autistic, and SLI children. More importantly, these shared traits are those which have been identified as factors associated with ToM development. For these reasons then, a delayed ToM is to be suspected in children with SLI. The necessary research questions become: How does theory of mind develop in children with SLI? What are the issues surrounding timing and development of this ability to mentalize? Does ToM mediate the link between language and behaviour in SLI children? Do the demonstrated behaviours change with age and degree of ToM delay? What is the influence of family size and working memory on ToM development in SLI children?

Theory of Mind and Language Impairment - The Literature to Date

There exists at this point in time limited reference to SLI within the ToM literature. Entering a variety of terms meant to capture language impairment and pairing these with the term "theory of mind" led to zero hits in the Psychology Abstracts database spanning from 1981 to 1998. The terms "language disorder" and "theory of mind" produced three reference articles. Perusal of their abstracts indicated that only two of the studies actually employed children with language disorder, the other simply made reference to the population. Despite this paucity of empirical research, numerous articles issue the bold conclusion that SLI children do not experience any delay or deviance in the development of a theory of mind (Eisenmajer & Prior, 1991; Leslie & Frith, 1988; Perner et al., 1989; Peterson & Siegal, 1997). Invariably such authors cite the same two studies - both of which only incidentally explored the issue. Two independent studies can hardly be considered a thorough investigation of the issue, especially when each has several features which could conceivably shroud any ToM deficits existing in SLI children. The limitations of the select studies on SLI and ToM will now be discussed.

(1.) Leslie and Frith (1988) conducted a study investigating the ToM ability of autistic subjects and employed SLI children to serve as a control group and a means of determining the role of language in mastering ToM tasks. The subjects were matched on verbal mental age for comparison purposes. The SLI subjects in this study had a mean chronological age of 8 years 8 months and a mean receptive verbal mental age of 6 years 9 months. When given first-order ToM tasks, they completed them successfully. The researchers took this to mean that language impaired children are without deficit or delay in achieving ToM and that language development is not related to ToM. However, given that normal children are able to master this same level of ToM task by 3 or 4 years of age, this cannot be deemed a major accomplishment, nor normal development, on the part of the 8 year 8 month old SLI subjects. It also does not rule out the possibility that their ToM was delayed. It is necessary to test SLI subjects of a younger age on this first-order ToM task to determine if indeed they are on track with developmental expectations. It also necessary to test SLI subjects with higher order ToM tasks - again - to determine if they are on track with developmental norms.

Another limitation of the Leslie and Frith (1988) study is that the ToM task used is fairly artificial in design. It is created in a manner designed to point out all of the visual information necessary to the participating subjects (i.e., the examiner pointed out all the relevant information with questions and teaching). Happé (1994) has found that higher-functioning autistic subjects are able to perform successfully on experimental second-order ToM tasks with such highlighted information. The same high-functioning autistic subjects are not able to completely transfer this achievement to ToM tasks that are more naturalistic in design. This too may be the case with SLI subjects. They may be able to perform in the formal and obvious circumstances of an experimental ToM task, but, not be able to transfer this ability to a more naturalistic type of ToM task. Less experimental ways of assessing ToM exist and it would be both interesting and useful to employ these methods with SLI children to determine if their ToM is functional within more life-like settings.

Furthermore, the measure of language ability used in this study is weak and uni-dimensional at best. VMA was determined by the British Picture Vocabulary Scale (BPVS; Dunn, Dunn, Whetton, & Pintillie, 1982), a nonverbal measure of receptive vocabulary. It is doubtful how the skill of associating a single concept with a pictorial equivalent would seriously contribute to the strategy and multi-tasking required to perform a ToM task. ToM tasks typically employ lengthier narratives, and generally require several concepts to be held in mind in serial order and potentially rearranged before arriving at a correct interpretation of the situation. In support of this argument, recent research has begun to query the usefulness of the BPVS (or any nonverbal receptive measure of vocabulary) for measuring the language skills related to ToM performance. Other measures of language ability (such as pragmatic language skills, syntactic abilities, verbal IQ, and various verbal subtests from the Wechsler scales) have proven to be better predictors of ToM ability than the BPVS (Astington & Jenkins, 1996; Eisenmajer & Prior, 1991; Ruffman et al., 1999; Tager-Flusberg & Sullivan, 1994a). So, if the BPVS does not correlate as highly with ToM as other measures of language then it cannot be used to reliably assess for, or partial out, the effects of language on ToM tasks. "It is not sufficient to partial out BPVS performance when examining whether two variables correlate. The relation between the two variables in question could very well stem from a common linguistic core" (Ruffman et al., 1999).

Finally, the Leslie and Frith (1988) study reported an age trend with regards to success on ToM tasks, but argued that these trends were invalid and could be ignored. A relationship was not found for autistic children between failing or passing and verbal mental age, however a trend for older autistic children to perform better than younger children almost reached significance. Design weaknesses also included a small sample size (N=12) and a very large span of verbal ability (4 years 5 months to 12 years 8 months).

(2.) Perner, Frith, Leslie, and Leekam (1989) conducted the second study to explore ToM ability in autistic children and used SLI children matched for verbal mental age as a control group. This study suffers from shortcomings similar to the above mentioned study (inappropriate chronological age of subjects, weak assessment of language skill, and inappropriate choice of ToM tasks) and therefore reached similarly unjustified conclusions which have been cited in the literature. Again, first-order ToM tasks were utilized thus indicating that the level of challenge was apt to be met by normal 3- or 4-year old children. Again, VMAs were calculated via the BPVS and so provide a poor estimate of language ability especially in relation to ToM performance. And once again, design weaknesses include a small sample size (N=12).

Also, the mental ages of these subjects (assessed by the British Picture Vocabulary Test - an equivalent to the North American PPVT) ranged from 5 years 5 months to 8 years 7 months, with a mean of 6 years 9 months in age. These ages are still well above those usually required to succeed on such a simple ToM task. Given the advanced age of the SLI subjects and the minimal challenge of the ToM task, it is not surprising that 11 out of 12 SLI subjects passed the task with flying colours. What is surprising is that this performance then led the authors to the following conclusion: "We can rule out the possibility that general impairment in language comprehension is responsible for failure because of the near-perfect performance of non-autistic children with specific language impairment" (Perner et al., 1989, p.697).

Does the Literature Predict What Behaviours Will be Affected by Delayed ToM in Children with SLI?

The link between behaviour and language delay appears to be quite specific in language delayed children: specific, likely even constant, but possibly changing over time. Behavioural immaturity and over-activity are common in young language delayed children. However, internalizing or neurotic problems are common in older language delayed children and adults (Beitchman et al., 2001; Stevenson, 1996).

A review of the link between language delay and later psychopathology in children with early presentation of language delay concluded, after judging all available data that the main increase in psychopathology appeared to be anchored in the domain of anxiety, social relationships, and attention-deficit problems rather than in conduct disturbance or antisocial behaviour (Rutter & Mawhood, 1991). Stevenson (1996) also found that behaviour problems occur and persist or develop in as many as 60% of children with early language delay. A high rate of internalizing problems was found. This is clearly unexpected in that the majority of children with language delay are boys, and this gender is more prone to demonstrating externalizing behaviours (APA, 1994).

Beitchman, Wilson, Brownlie, Walters, Inglis, and Lancee (1996) documented a finding that "children with receptive and pervasive speech/language problems at age 5 demonstrated greater behavioural disturbance than children without such impairment"(p. 815). When the initial behavioural status was controlled for statistically, early childhood language competency was still linked with behavioural and social competence ratings, seven years later. Also, children with low overall or poor comprehension scores (i.e., mixed or pure receptive difficulties) show the greatest impairment on these measures.

Haynes and Naidoo (1991) conducted a longitudinal study investigating the behaviour of school-aged children suffering from speech and language delays. The data accumulated from teacher ratings revealed a trend for behaviour to quickly move from high rates of frustration and aggression in younger children (lower age limit 6 years) to high rates of low expressed self-confidence, low self-esteem, and increased social withdrawal (upper age limit 11 years of age).

Baker and Cantwell (1987) found that the psychiatric disturbance found in SLI children varied with age. Much of the earlier disorders are of an externalizing nature such as Attention Deficit/Hyperactivity Disorder (ADHD) and may be related to neurological immaturity. But, the later psychiatric disorders in this population of children are predominately anxiety based and cannot easily be explained by neurological immaturity. This suggests that later psychiatric disturbance in SLI children may be a response to a life of language impairment and difficulties in communication and social skills.

Tallal, Dukette, and Curtiss (1989) investigated the relationship between developmental language and psychiatric disorders in preschool-age language impaired children using the parent version of the Achenbach Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1991a). One hundred and one 4-year-olds with specific developmental language impairment were compared to a matched group of control children. Significant between group differences were found for boys, but not girls, on broad-band syndromes; increased Total Behaviour Scores were found for the language-impaired boys, but not the language-impaired girls relative to their matched peers in the control group. While few between group differences were found for narrow-band syndromes, the Immaturity scale significantly differentiated language-impaired boys from control boys and the Social Withdrawal significantly differentiated language-impaired girls from control girls.

Stevenson, Richman, and Graham (1985) conducted a study in an attempt to establish whether early language development was related to later behavioural deviance. A variety of language and behaviour measures were administered to 535 children on their third and eighth birthdays. Three-year-olds with poor language skills but no reported behaviour problems were found at age 8 to show a high rate of neurotic deviance on the Rutter Teacher's Scale (Rutter, 1967), when behaviour at age 3 was controlled for. Thus, a specific association between early language and later behaviour was demonstrated. The association is not with degrees of language disability, but, of children with poor language structure (i.e., not using certain features of their expressive speech as determined by scores on the Reynell Development Language Scales). An additional finding of this study which also employed the English Picture Vocabulary Test (Brimer & Dunn, 1962) as a measure of language development (receptive and expressive) was that language structure presented as a better predictor of later behaviour than did pure expressive or receptive language skill.

Cohen, Menna, Vallance, Barwick, Im, and Horodezky (1998) examined the social cognitive skills, behavioural ratings, and psychiatric diagnoses of 380 children 7 to 14 years of age with identified and unsuspected language impairment who had been referred for psychiatric services to two mental health centres in Toronto. The results indicated that children with language impairment showed greater deficits in social cognitive processing relative to children with normally developing language. Also, children with previously identified language impairment showed different psychiatric diagnoses and behaviour problems only in relation to children with normally developing language. The top diagnoses for children with previously identified language impairment derived from parent interview were ADHD (42.9%), Oppositional Defiant Disorder (39.5%), and Dysthymia (28%). These were closely followed by several anxiety based disorders: Over-anxious Disorder (23%), Separation Anxiety Disorder (15.3%), and Phobias (11.3%). Findings derived from ratings on the Teacher Report Form (TRF; Achenbach & Edelbrock, 1991b) also found the children with previously identified language impairment to demonstrate significantly more of the following problem behaviours: Attention, Withdrawal, Anxious, and Depressed. Parent ratings on the CBCL (Achenbach & Edelbrock, 1991a) indicated that language impaired children received more severe total scores than children without language problems. Here then, we have mixed evidence: children with language problems do entertain significantly more behaviour problems than children with normally developing language, however, the form it takes can be either internalizing or externalizing in nature.

<u>What Behaviour Would a Delayed ToM Predict: Externalizing then Internalizing?</u>

A lack of theory of mind means that an individual is necessarily reading the script of the social world in a very present-based and self-centered manner. Thus, there will be little motivation for, or insight into, interaction on anything other than a concrete and self-referenced level. Given this, there will be very little in the way of connection with other persons; the means of gratification is not in place for either party. In essence, there will be an inability to share points of mutual interest.

What might the effect of this be on the behaviour of the language delayed individual? The absence of studies addressing this question creates the need to pose the speculative answers contained in the following paragraphs. Initially delays in ToM might lead to acting out behaviour as the individual experiences frustration and tries "more of the same" in order to get their point across or to capture the attention of another child. The child may turn to an excess of physical behaviour in an effort to draw attention to themselves, or may move from one activity to another in the hopes of keeping social engagement on a surface level. A short attention span may be displayed because the child cannot follow the true intent of others. Play in a self-centered manner might lead to conflicts of interest with others or disputes over such things as roles and turn-taking. Unwanted or unsatisfied, the child may act out or move on. In essence, the child may use behavior to construct his or her social needs because she or he is unable to follow those constructed by others.

A lack of ToM may also inhibit a child's acquisition of skills in conflict resolution. Stevens and Bliss (1995) found that SLI children tended to consider threats and physical action to be their best bet for solving hypothetical disputes, whereas normally developing children enlisted persuasion, explanations, and questioning as the best means to resolve a hypothetical dispute. The reliance of the latter strategies on an understanding of others' minds is quite clear.

Children without a ToM will continue to respond to the concrete stimuli of the moment and not the hypothesized thoughts, feelings, or desires of their counterparts in social interactions. Eventually, as peers continue to advance in their understanding of this "secret" or "invisible" code of knowledge, and the child with delayed ToM remains on the outside of this knowledge, so too, will he or she move to the outside of social interactions with their same-aged peers. Unable to fathom what motivates others or to reliably predict their behaviour, social interaction becomes perhaps too threatening and incomprehensible to bother with. Thus a withdrawal from others is predicted for the later years in children with delayed ToM. Conceivably the fear and worry over interacting with others is due to their incomprehensible points of reference and choice of responses. Repeated failure to learn proper interaction or an involuntary, but, self-imposed isolation from others could naturally lead to acquisition of internalized behaviours as anxiety, withdrawal, and shyness. If the child with delayed ToM comes to believe that the behaviour of others is arbitrary, then withdrawal from the bumpy and awkward realm of peer play seems a reasonable choice.

Stevenson (1996) states, "the putative effect of delayed language on internalizing problems through an influence on an impairment in social cognition needs to be investigated more fully" (p.94). Theory of mind ability is obviously an element of social cognition and so its role in influencing the behavioural development of language impaired children needs to be investigated. Stevenson (1996) hypothesizes in the same spirit as the above paragraph that the reduced social cognitive ability of language delayed children (due to poor early language experience and limited opportunity for effective language use) will reduce a child's tendency to enter into social interaction. The language delayed child, aware that he or she cannot read or interpret the nuances of social interaction will shy away from such circumstances. The child who for example cannot perceive order and structure in social interaction will likely avoid such unpredictable circumstances.

Happé and Frith (1996) looked at ToM ability in children with the diagnosis of Conduct Disorder because, as they reasoned, this is a group of children that display problems in social interaction that are somewhat similar to those displayed by children with autism. Because a lack of ToM has been relatively "successful in explaining many of the social difficulties in autistic children (both in laboratory tests and in everyday life)" (Happé & Frith, 1996, p.385), it was hypothesized that the social impairment seen in Conduct Disorder may have a similar foundation. The study showed that simple tests of understanding false beliefs (first-order) did not discriminate children with Conduct Disorder from normal controls. Perhaps this is not surprising given that the average verbal mental age of the children was 8 years 0 months and given that the task is usually passed by 4-year-olds. This finding implies that if Conduct Disordered children have a delayed theory of mind, it is not as significantly delayed as say that of deaf or autistic children. However, of interest in this particular study is the fact that the Vineland subdomains of Communication and Socialization, as well as additionally created items (by the authors), showed marked and specific real-life differences. The Conduct Disordered children displayed widespread social dysfunction and the identified atypical behaviours were very much those that presuppose a well-functioning theory of mind.

## The Present Study

The primary aims of the present study were to add to the meager body of literature that investigates ToM development in children with significant language impairment and to improve upon the investigative procedures used to study the issue. A thorough investigation of the development of ToM in SLI children was thus proposed. As well, known correlates and predictors of ToM for normally developing children were investigated to determine their generalizability to a special population of children (SLI). Finally, this study sought to determine if ToM is related to the development of particular behaviour disorders within SLI populations.

### Design Improvements

Age. The most significant weakness in the studies to date is the chosen subject pool. Subjects have tended to be of a very broad range of ages, to be older than the ages at which the given ToM tasks are believed to be normally mastered, and to be few in number. The subjects in Leslie and Frith's study (1988) and Perner et al.'s study (1989) ranged in age from 6 years 11 months to 9 years 11 months. These three years of development were represented by a sample of only 12 children; at best this means that four children represented each year of development. The mean chronological age of the sample was 8 years 8 months which is well above the 4-year old level at which first-order ToM tasks are usually successfully completed.

The present study addressed the issue of subject age by selecting a sample of SLI and normally developing children whose ages correspond to the timing of ToM achievement in a normal population. The subject sample was divided into two groups based on the ages at which first- and second-order ToM tasks are passed by normally developing children. Thus, 4- and 5-year old children formed one group while 6- and 7-year old children formed the second group.

It was assumed that the performance of the normally developing children would match that reported in the developmental literature: 60% - 80% of the 4- and 5-year-olds should pass the first-order tasks, and, 60% -80% of the 6- and 7-year-olds should pass the second-order tasks. It was predicted that the performance of the language disordered children would fall below that of the normally developing children with less than 60% -80% of the 4- and 5-year-olds passing the first-order tasks, and, less than 60%-80% of the 6- and 7-year-olds passing the second-order tasks.

ToM. The second crucial shortcoming of the literature investigating ToM development in language impaired children is the inappropriateness of ToM measurement. Leslie and Frith (1988) ran two different tasks which involved one limited knowledge question and two prediction questions as their measure of ToM development. Perner et al. (1989) tested SLI children on two false-belief tasks which were based on only two prediction questions. The tasks from both experiments were rated as being of first-order difficulty. Given the complexity of ToM as a construct, a fuller assessment approach is required before making comment on the ability of SLI children to understand other minds.

The present study utilized a full battery of ToM tasks to tap the degree of ToM development. First- and second-order ToM tasks were given so that the level of ToM achievement could be contrasted with both age and diagnostic group (normally developing vs. SLI). First-order ToM was assessed using the total number of correct answers to five false belief questions. The questions were posed within the paradigm of an "unexpected change of location" task and a "deceptive appearance" task. Second-order ToM was assessed using the total number of correct answers to four questions assessing either ignorance or false belief. The questions were generated by two second-order stories involving the "unexpected change of location" paradigm. Additional estimates of first- and second-order ToM were gathered with the administration of three sarcasm vignettes. One first-order and second-order question accompanied each story. Utilizing such a broad base of measures ensured that subjects' scores were not the result of "chance" or guessing on one or two questions. Also, this number of questions allowed ToM development to be categorized as either intact, transitional, or beginning. ToM was therefore not reduced to an all-or-none state as warned against by various researchers (Astington & Jenkins, 1995; Keenan et al., 1998). Rather, ToM was conceptualized as an unfolding and variably developing, newly learned skill.

It was predicted that both populations would follow the trends prominent in developmental literature wherein children will pass first-order ToM tasks before they pass second-order ones. In other words, ToM will increase with age and be reflected in higher scores on first- and second-order tasks. Also, normally developing children were predicted to have a more secure ToM than SLI children as indicated by higher total ToM scores.

It was expected that with the above noted design improvements in place, a difference would indeed surface between the ToM development of SLI children and that of their normally developing peers. The delay indicated by this difference was not expected to be of a magnitude equal to that found in autistic children. However, it was expected that the difference between SLI children's ToM scores and normally developing children's ToM scores would be statistically significant.

Extension of Previous Research

A second goal of the present study was to extend findings from the body of literature investigating factors influencing ToM development in normally developing children to incorporate SLI children. Accordingly, this study assessed the contribution of language, working memory, and siblings to the ToM development of SLI children.

Language. The function of language in relation to ToM development is currently a topic of great interest. As a result, the contribution of language was analyzed in several ways. It was assessed in a general manner by comparing SLI children's performance on ToM tasks to that of their normally developing peers. This provided a gross estimate of language's contribution to understanding other minds.

Language's relation to theory of mind was also considered by comparing subjects' performance on two brief measures of language to their performance on the ToM battery. Given that measures of receptive vocabulary have demonstrated weak and unstable relationships with ToM ability (see Ruffman et al., 1999; Tager-Flushberg & Sullivan, 1994a) more specific aspects of language functioning were measured in the present study. The Information subtest from the appropriate Wechsler test (either the Wechsler Preschool and Primary Scale of Intelligence - Revised [WPPSI-R; Wechsler, 1989] or the Wechsler Intelligence Scale for Children - Third Edition [WISC-III; Wechsler, 1991], depending on the subject's age) and the Linguistic Concepts subtest from the appropriate Clinical Evaluation of Language Fundamentals (CELF) test (either the CELF-Primary [Wiig, Secord, & Semel, 1992] or the CELF-3 [Semel, Wiig, & Secord, 1995], depending on the subject's age) provided the necessary language measures. It was predicted that these scores would be highly correlated with ToM performance for both SLI and normally developing children. In accordance with Ruffman et al.'s findings (1999) it was predicted that the Information subtest would have a slight edge in predictive value.

Finally, a nonverbal measure of ToM was also administered to the subjects. This provided a finer analysis of the role of language in ToM achievement. The picture sequencing task of Baron-Cohen et al. (1986) was employed. Three trials of the "intentional" picture cards were administered. Thus, nonverbal ToM was assessed using the total number of correct answers to three first-order intention questions. Eliminating the verbal presentation of a ToM scenario, ToM question(s), and the need for a verbal response (be it an open-ended or forced choice answer) allowed for a purer test for the presence of ToM understanding. In this manner, a nonverbal measure of ToM helps to determine if impaired language skills impair the *communication* of an understanding of ToM or if impaired language skills impair the *conceptualization* of a theory of mind. In adherence to the notion that impaired language development retards conceptualization of the working of other minds, it was predicted that the SLI children would also be less successful on this measure than their normally developing peers.

Siblings. The importance of family size in aiding ToM development through sibling interaction has been found in several studies of normally developing children (Jenkins & Astington, 1996; Perner et al., 1994). The increased significance of sibling interaction for children with weak language skills has been noted in only one study to date (Jenkins & Astington, 1996). The present study extended previous results by looking at the role of family size in the achievement of a theory of mind in children diagnosed with SLI. Family size was represented by the number of siblings currently living within the household. The same questions posed by Jenkins and Astington (1996) were given to the parents of all participating subjects. In line with the hypothesis that social interaction supports the development of theory of mind ability, it was predicted that SLI children from larger families would show a ToM advantage relative to SLI children from smaller families.

Working memory. A growing number of studies document the importance of working memory in successful performance on ToM tasks (Davis & Pratt, 1995; Keenan, 1998, 1999). Working memory span has been measured in several ways throughout the ToM literature. Difficulty has been noted with a floor effect on the backwards digit span task, so this approach was not used. Memory for sentences tasks could easily contain a confound between memory and language skills, also ruling out the usefulness of this task. The counting span task originally created by Case et al. (1982) has merit as does the modified version employed by Keenan (1999). The modified version has proven successful with children as young as 3 years of age and so it was utilized in the present study. Three trials of this task were given at each level to provide working memory span scores ranging from 0 to 9.

## A New Piece to the Puzzle

The third objective of the present study was to try and unearth an element which may help explain the often noted overlap of language disorder and behaviour problems. It was proposed that possession of an underdeveloped ToM would be linked with problems in the areas of language and behaviour. This study looked for correlations between language disorder and ToM and between ToM and behaviour problems. It was predicted that as the level of ToM increased, the level of language disorder and behaviour disorder would decrease.

Behaviour. Research exploring the concomitance of language and behaviour problems provides tentative evidence to suggest that the nature of behaviour problems in children with language disorders changes over time. A move from externalizing behaviours to internalizing behaviours is proposed as occurring around the age of 6 (Haynes & Naidoo, 1991). The current version of the Child Behaviour Checklist (CBCL; Achenbach & Edelbrock, 1991a) was completed by the parents of all study participants. The CBCL is a common measure in studies investigating the overlap between language and behaviour disorders (Stevenson, 1996). It provides a standardized description of children's problems as reported by parents.

This behavioural measure provides a list of behavioural problems and competencies which are rated by parents. The behavioural problem scale consists of items grouped into a variety of narrow-band scales which are then grouped into two broad-band factors: Internalizing and Externalizing. It was predicted that data gathered from the present study would support a metamorphosis of behavioural expression with the 4- and 5-year old SLI children demonstrating more externalizing behaviours than internalizing behaviours and the 6- and 7-year old SLI children demonstrating the opposite relationship (more internalizing behaviours and less externalizing behaviours).

Hypotheses

Therefore, the principal hypotheses addressed in this study are: (1) between group differences will be observed for the SLI and normally developing children on measures of language and ToM. SLI children were expected to have lower levels of language development and to be delayed in their ToM development. This delay was expected to hold even on the nonverbal measure of ToM as it was predicted that disordered language retards conceptualization of ToM above and beyond communication of ToM; (2) a developmental course of ToM ability would be demonstrated with older children out-performing younger children within both the SLI and normal groups. To this end, it was expected that children will pass first-order ToM tasks before they pass second-order ones; (3) the variables of working memory, language, and siblings found to predict ToM performance in normal children would also predict ToM performance in SLI children. Specifically, as working memory capacity increases, language skill increases, and number of siblings increases, so too will the level of ToM development, in other words, a positive correlation is expected for all three predictor variables; (4) it was predicted that a relationship would emerge between degree of language impairment and ToM delay relative to ratings of behavioural disturbance on the CBCL. It was expected that the greater the lag in language and ToM ability, the greater would be the overall display of inappropriate behaviours. As well, the type of noted behavioural problems were expected to be rated as internalizing for children below the age of 6 and to be rated as externalizing for children above the age of 6.

# Chapter 2 - Method

## Participants

Forty-nine normally developing children ranging from 4 to 7 years of age were recruited through three local Day Care Centres and two local public schools. There were 26 boys and 23 girls. Subjects in this group were considered to be developing normally based on the opinion of school personnel, and an absence of formal identification and diagnostic labels.

Forty-three SLI children ranging from 4 to 7 years of age and identified as SLI by Speech-Language Pathologists were recruited. There were 28 boys and 15 girls. The exclusionary criteria for SLI were given to the Speech-Language Pathologists: no hearing loss or history of recurrent otitis media, no significant emotional or behavioural problems (in particular such states as autism, Asperger's or PDD-NOS), no mental retardation or neurological problems, and no sensory or oral defects. The Speech-Language Pathologists considered these criteria when choosing among children on their case lists who had achieved scores (receptive or expressive or both) 1 ¼ standard deviations below the norm on a standardized test of language. Locating local children who met this standard was difficult. Subjects were therefore referred from a children's centre and several public schools in Southern Ontario as well as from two local agencies.

The normally developing and SLI children were then divided into two groups: "younger" (4- and 5-year-olds) and "older" (6- and 7-year-olds). This age division reflects the timing of first-order and second-order ToM achievement in normally developing children as noted in the literature. The number of children in each group along with their gender, mean age, age range and standard deviation is given in the table below.

Table 1

Subject Characteristics

| Group | N | Mean Age | Standard Deviation | Range | Gender (Male: Female) |
|---|---|---|---|---|---|
| Younger | 47 | | | | 29 : 18 |
| SLI | 23 | 5 yr. 4 mo. | 5.30 mo. | 4 yr. 4 mo. - 5 yr. 11 mo. | 15 : 8 |
| NSLI | 24 | 5 yr. | 5.68 mo. | 4 yr. 11 mo. - 5 yr. 10 mo. | 14 : 10 |
| Older | 45 | | | | 25 : 20 |
| SLI | 20 | 7 yr. 1 mo. | 7.80 mo. | 6 yr. 0 mo. - 8 yr. 1 mo. | 13 : 7 |
| NSLI | 25 | 7 yr. | 6.69 mo. | 6 yr. 0 mo. - 7 yr. 10 mo. | 12 : 13 |

Overall, 93 children were tested. The data from one subject was excluded from the analysis due to an incomplete protocol. The mean ages of the older groups of SLI and NSLI children were not significantly different ($t$ = .58, 43df, $p$=.56). However, the mean age of the group of young SLI children was greater than that of their NSLI counterparts ($t$ = 2.47, 45df, $p$<.02). This provides the young SLI children with a slight advantage concerning any of the benefits accompanying age. Therefore, should ToM deficits be found in the young SLI group relative to the young NSLI group, they must be considered that much more striking.

Materials

First-order ToM tasks. Three trials of the standard "unexpected change of location" paradigm were given. These standard false belief tasks are based on the original false belief task designed by Wimmer and Perner (1983). The "Smarties" tasks was also given along with its two false belief questions. This standard unexpected contents tasks is based on the version used by Perner et al. (1987). Three sarcasm stories containing three first-order sarcasm questions as described by Keenan and Quigley (1999) were also used.

Second-order ToM tasks. Two second-order stories containing six second-order ToM questions as described by Sullivan, Zaitchik, and Tager-Flusberg (1994) were given. Three sarcasm stories with three second-order sarcasm questions as described by Keenan and Quigley (1999) were also used.

Nonverbal tasks. The picture sequencing tasks of Baron-Cohen et al. (1986) were used. A preliminary teaching trial was given using a "mechanical" story to demonstrate the task requirements. Three trials of the "intentional" picture cards were then administered.

Language tasks. The Information subtest from the appropriate Wechsler Intelligence Scale and the Linguistic Concepts subtest from the appropriate version of the Clinical Evaluation of Language Fundamentals test were used. The Information subtest of the WPPSI-R was used with children aged 3 years 11 months to 6 years 11 months. The Information subtest of the WISC-III was used with children aged 7 years 0 months to 7 years 11 months. These two subtests are considered to be downward or upward extensions of one another (Wechsler, 1991). There is overlap in the age range for which these two versions are applicable: 6 years 0 months to 7 years 3 months. Given the lowered language ability of the SLI children, the younger version of this test was chosen for all the 6-year old children. This is in accordance with the manual guidelines which state, "in most cases for children of below-average ability, the WPPSI-R should be used because it has a lower floor than the WISC-III" (WISC-III Manual, 1991, p.33). The CELF-P and the CELF-3 possess an age range overlap that is similar to that of the WPPSI-R and the WISC-III. Continuity of logic and test materials then dictated that the younger version of the

CELF would be used for children in the age range of 6 years 0 months to 6 years 11 months. The Linguistic Concepts subtest of the CELF-Preschool was therefore used with children aged 3 years 11 months to 6 years 11 months. The Linguistics Concepts and Oral Directions subtest of the CELF-3 was used with children aged 7 years 0 months to 7 years 11 months. These two subtests are considered to be downward or upward extensions of one another (Wiig et al., 1995).

Behaviour. The parent version of the Child Behaviour Checklist (CBCL; Achenbach & Edelbrock, 1991a) was completed by subjects' parents.

Sibling measure. Parents of subjects were asked how many siblings resided in the subject's home and the ages of these siblings.

Working memory span. The houses spatial working memory task as described by Keenan (1999) was administered to subjects.

Design

Three sets of tasks were completed with all subjects: working memory span, theory of mind, and language sample. A behaviour measure and a sibling measure were collected by having parents complete a CBCL and several questions about household occupants.

Given the number of tasks and levels within tasks, the battery was administered in a standard order: working memory span task, first-order ToM (three stories then Smarties task), second-order ToM (three sarcasm stories then two standard version stories), nonverbal ToM (three trials), language (WISC Information subtest, CELF Concepts subtest). This particular order was chosen as it is presumed to represent an increase in difficulty with the harder tasks thus receiving the benefit of some familiarity. Also, it has been noted by at least one researcher that children given a language task first are often tired and restless throughout the remainder of the testing session (Keenan, 1998). The presentation of first-order, second-order, and nonverbal

ToM stories was randomized by having the children draw numbered pieces cardboard out of an empty film canister - the stories were given in order of number drawn.

Procedure

Informed consent for participation was gathered in written format from the subjects' parents. A letter was disseminated through the various institutions explaining that the study posed no risk to participants and that it had been declared as conforming to ethical standards by the University's Ethics Committee and deemed acceptable by the governing body of each institution (school boards, principals, day care supervisors, health unit directors). The purpose of the study, a brief description of the tasks, time commitment, the voluntary nature of participation and the right to withdraw at any point in the process were also delineated in the letter. It was explained that all results would be kept confidential and that individual results could not be made available to anyone, parents or otherwise. Finally, the letter provided contact information for the parents to use should they have questions or concerns related to the study. Consenting parents were asked to complete and return a CBCL and a form gathering the following information: parents' name, child's name, child's date of birth, number of siblings, and number of older siblings residing in the home. A full version of the covering letter to parents and consent form are respectively contained in Appendix A and B.

Children were tested individually in a quiet area of their school, home, or agency office. The children were asked to name one or two of their friends as part of initial rapport building. One of these names was then recorded for use within the Smarties task. All subjects were seen individually by the researcher over a period of 9 months. The test battery was completed in one session that varied in length from 30 minutes to 60 minutes depending on the age and cooperation level of the child. Subjects were introduced to the researcher by their teacher or parent, then escorted to the testing area.

All answers were recorded for scoring after the sessions. All protocols were scored first by the researcher. The scoring of the sarcasm task contains an element of subjectivity and so it was also scored by a fellow graduate student who applied the same scoring criteria. Inter-rater agreement of 95% was initially achieved; after discussion this rose to 100%.

Working memory spatial span task. This task consisted of a series of index cards (3" x 5"). Centered in the middle of each card was a nine squared grid topped with an inverted, extended "V". The overall effect thus resembled a house. The houses were identical apart from their variation in colour (red, yellow or green) and the location of a single red dot. House colour varied for each item within a level to help children organize their responses. An adhesive red dot had been randomly placed in one of the nine boxes contained within the grid. The child was given 3 seconds to note and remember the location of the dot before the index card was flipped over and a blank index card was presented for 2 seconds. Then a house identical (colour) to the original stimulus but empty (no red dot) was shown and the child was required to point to the location of the missing red dot. Three levels of difficulty exist and each level consists of three sets of items. Level 1 involves one stimulus card and one blank card. Level 2 involves two stimulus cards each separated by a blank card; the child must work through all four cards before pointing out the two correct dot locations in the correct order on the two blank houses. Level 3 involves three stimulus cards each separated by a blank card; the child must work through all six cards before pointing out the three correct dot locations in the correct order on the three blank houses. To begin, children were given two practice trials at level one and one practice trial at level two. Three trials were then given at each level. The task ended when a child failed two consecutive trials on any given level. A score of 1 was given for each trial correctly answered and so scores ranged from 0 to 9 on this task. The script for this task is located in Appendix C.

First-order ToM, unexpected change of location. The children were given three variants of the false belief task originally devised by Wimmer and Perner (1983). The scenarios were presented in story book form using the mother, father, and Calvin characters from the Calvin & Hobbes cartoon strip. The main events of the three stories were each represented in three colourful, full 8½" x 11"page drawings and laminated for protection. The pictures were placed in a three-ring binder and indexed for ease of location and display. The standard scenario ran as follows: two characters are in a room and a toy or object is obviously placed in one of two locations. Character #1 then leaves the scene and is out of sight and unable to see what Character #2 is doing or saying. Character #2 moves the toy or object to a second location. At this point, the child is asked three control questions: memory question (where did Character #1 put the toy or object in the beginning?), reality question (where is the toy or object now?), knowledge question (did Character #1 see Character #2 move the toy or object?). A correct answer on a control question resulted in a score of 1. Control question scores were summed across the three trials resulting in a control question score ranging from 0 to 9. Finally, Character #1 is brought back into view and the child is asked the false belief test question, "where will Character #1 look for his or her toy or object?" Children were given a score of 1 for each correct answer, and the score was totaled across trials to create a first-order score ranging from 0 to 3. The script for this task is in Appendix D.

First-order ToM, unexpected contents. In this task, children were shown a Smarties box and asked to state what they believed to be inside. The contents of the box were then revealed as a number of small pencils. The box was closed up again and children were asked two control questions to gauge their understanding of the task: "what do you think is in here?", and "what is really inside the box?" A score of 1 was given for each correct answer and the scores tallied to form another first-order control score that ranged from 0 to 2. The children were asked a representational change question ("When you first saw the box, before we opened it up, what did you think was inside it?") and a standard false belief question (" [Name of one of the child's friends] hasn't seen the inside of the box yet. When she or he first sees the box, all closed up like this, what will she or he think is inside it?"). Non responders were given forced choice versions of the questions, ("Did you think there were Smarties inside or did you think there were pencils inside?" and "Will [name of friend] think there are Smarties or pencils inside the box?) Children were given a score of 1 for each of the false belief questions they correctly answered. The range of possible scores was 0 to 2. The script for this task can be found in Appendix E.

Children's scores were tallied across the two first-order tasks to form a first-order ToM composite score ranging from 0 to 5. Children's control scores were also tallied across the two first-order tasks and formed a first-order control composite ranging from 0 to 11.

First- and second-order sarcasm. Three sarcasm stories as originally devised by Keenan and Quigley (1999) were given to all participants. Each story is approximately 7 lines long and contains approximately 100 words. The stories were tape recorded and then played to the accompaniment of a series of colour pictures depicting the stated action. The pictures were drawn on 8.5" x 11" sheets of white paper and laminated for protection. Four pictures accompanied each story. The pictures were placed in a three-ring binder and indexed to allow for easy location and display of the correct set of pictures. Two memory for fact questions, a first-order sarcasm question and a second-order sarcasm question were posed to the subjects after they heard each story. A correct answer on a memory for fact question was given a score of 1. A sarcasm composite control score ranging from 0 to 6 was formed by summing fact question scores from across stories. Each correct answer to each ToM question earned a score of 1. Scores were summed across stories to form a first-order sarcasm composite score (range 0 to 3) and a second-order sarcasm composite (range 0 to 3). The scripts for the sarcasm stories are located in Appendix F.

Second-order ToM, unexpected change of location and ignorance. Two second-order stories as originally designed by Sullivan et al. (1994) were given to all participants. Six memory for fact questions were posed to the subjects throughout the telling of the story. A score of one was given for each correct answer on all memory for fact questions resulting in a second-order control composite score ranging from 0 to 6. A second-order ignorance and second-order false belief question were posed to the subjects after the reading of the story. Each correct answer to these ToM questions earned a score of 1. Scores were summed across both stories to form a second-order theory of mind question composite score which ranged from 0 to 4. The scripts and materials used in the second-order tasks are given in Appendix G.

Nonverbal ToM. Three "intentional" stories as originally devised by Baron-Cohen et al. (1986) were given to all participants. "Intentional" stories reflected people acting in everyday activities requiring attribution of mental states. The pictures were drawn on blank 5"x5" index cards and laminated for protection. Four pictures depicted each scenario. A child's ability to arrange given pictures into a predetermined sequence was taken to indicate their understanding of the story depicted in the sequence.

A teaching trial was given using a "mechanical" story prior to administering the stories thought to reflect theory of mind. The procedure for the teaching trial is closely based upon that used for the Picture Arrangement subtest on the WISC-III. The Picture Arrangement subtest is similar in design and intent to this nonverbal ToM task. As well, the Picture Arrangement subtest is deemed appropriate to administer to children with limited language skills because it can be both demonstrated and completed without words.

The teaching trial ran as follows: a set of four cards was placed face up on the table, one at a time, at the pace of about one card per second. The initial card was in the correct position, while the remaining three were in a pre-determined and jumbled order. While the cards were being placed face up on the table, the experimenter said, "These pictures tell a story of a boy who pushes a rock and makes it roll down a hill. This card goes first, it shows the boy standing on top of the hill and looking at the rock by his feet. Show me which cards go next." If the child placed the cards in the correct order she or he was congratulated and the story explained out loud while the pictures were picked up by the experimenter. If the child did not place the cards in the correct order, the next correct card in the sequence was arranged for the child and the prompt "Show me which cards go next' reiterated. This continued until the cards had been arranged in the correct order and the story line verbalized by the experimenter.

Then, three "Intentional" stories were given. The pictures were placed on the table in a set order with the beginning picture being the correct one for the sequence. This meant that the children had only to place three pictures within each trial. The verbal directions to accompany this task were: "This is the first picture. Look at the other pictures and see if you can make a story with them." The card arrangement chosen by the child was noted down, after any self-corrections. Each child was allowed to proceed at their own pace, but given only one attempt at each of the three stories.

The three "Intentional" stories run as follows: Intentional Story #1: 1) boy buys sweets, 2) leaves shop, 3) sweets drop out of bag, 4) boy sees sweets gone. Intentional Story #2: 1) girl puts teddy down, 2) turns to pick flower, 3) boy takes teddy, 4) girl sees teddy gone. Intentional Story #3: 1) boy puts chocolate in box, 2) goes out to play, 3) Mum eats chocolate, 4) boy sees chocolate gone. A completely correct sequence earned 2 points. A sequence having the correct end point for the story was given 1 point. Scores across the three trials were summed to form a nonverbal ToM composite ranging from 0 to 6.

Language tasks. Children were administered the Information subtest from the appropriate Wechsler Intelligence Scale following the standardized instructions given in the manual. Children were also administered the Linguistic Concepts subtest from the appropriate CELF test following the standardized instructions given in the manual. The children's scale scores from each subtest were used as a measure of the level of language development of each group.

Behaviour. The CBCL was completed by the parents of the participants. It was then scored according to the procedures described in the manual. A total behaviour score, an internalizing score, and an externalizing score resulted. These were in the form of standard scores and they were assessed for statistical significance according to the norms table provided in the manual.

Siblings. The parent of each participant was asked to provide answers to the following two questions: 1) How many children reside in your household?; 2) What are the ages of the children residing in your household? A siblings score was formed using the number given in answer to the first question. An older siblings score was calculated by counting the number of older aged siblings listed as living with the child participating in the study.

# Chapter 3 - Results

## Introduction

The results of the study are organized into six sections. Section 1 presents a summary of the correlations observed among all variables organized by dependent variables. Section 2 presents an analysis of the differences between age and diagnostic groups on all ToM tasks. These results are then compared and contrasted to the between age and diagnostic group differences on the remaining outcome measures of working memory, language, behaviour, and siblings. This was accomplished by analyzing children's scores for correct answers with a series of two-way ANOVAs. Section 3 provides an analysis of the developmental progression of ToM ability. This was carried out by comparing the number of children in each age and diagnostic group passing first-order and second-order ToM and sarcasm tasks. Chi-square analyses were run and gammas calculated to determine the significance of these differences, and the results of pertinent two-way ANOVAs are reviewed. Section 4 explores the ability of working memory, language, and age to predict children's performance on ToM tasks. A series of multiple regressions were run to this effect. Section 5 provides an analysis of the relationship between parental ratings of children's behavioural problems and language and ToM ability. Chi-square analyses were run to determine the significance of the differences in the number of children in each age and diagnostic group rated as displaying normal or significantly elevated levels of inappropriate behaviour on the Internalizing, Externalizing and Total scales of the CBCL. Finally Section 6 addresses incidental questions which arose as a result of findings in the previous sections.

Section 1 - Correlations

The correlations for all 16 variables are given in Table 2. Strong negative correlations occurred between diagnostic group and the two language measures (WISC $r$ = -.60 and CELF $r$ = -.65), indicating that higher language scores were associated with the NSLI group (group 1) and lower language scores were associated with the SLI group (group 2). Both language measures were positively and significantly related to all ToM tasks and first-order sarcasm with $r_s$ ranging from .24 to .51. Thus an association between language and ToM ability is suggested. Both language measures were significantly and negatively related to the number of siblings in the household ($r$ = -.16 and -.29). This relationship is the inverse of what was predicted from the literature.

Diagnostic group was significantly and negatively associated with all ToM measures ($r$ = -.34 to -.41) and with first-order sarcasm ($r$ = -.33). This reflects the superior performance of the NSLI children (group 1) over the SLI children (group 2). As well, diagnostic group was positively associated with the number of siblings in the household ($r$ = .27), indicating that the SLI children tended to have more siblings than the NSLI children.

All ToM measures were significantly correlated with one another ($r$ = .40 to .68), language ($r_s$ as given above), first-order sarcasm ($r_s$ as given below), and working memory ($r_s$ as given below). Of note, the nonverbal measure of ToM had a slightly lower correlation with the language measures ($r$ = .24 and .29) than did the verbal measures of first-order ToM ($r$ = .33) and second-order ToM ($r$ = .51). Also of note, second-order ToM and nonverbal ToM had stronger positive associations with working memory ($r$ = .34 and .35 respectively) than did first-order ToM ($r$ = .29).

Regarding sarcasm, first-order sarcasm appeared to function as a competent measure of ToM and was positively correlated with: language (CELF $r = .43$ and WISC $r = .29$), first-order ToM ($r = .46$), second-order ToM ($r = .63$), nonverbal ToM ($r = .47$), and working memory ($r = .42$). Second-order order sarcasm however, was generally unrelated to the given measures, apart from a positive association with first-order sarcasm ($r = .27$) and working memory ($r = .23$).

Moderate positive associations were exhibited between working memory and measures of ToM ($r = .29$ to .35), and sarcasm ($r = .23$ to .42). A moderate correlation also occurred between working memory and the CELF language measure ($r = .22$). A link between working memory and ToM ability is thus suggested.

Behaviour was negatively and significantly correlated with both siblings measures. Number of *older* siblings was associated with all three scales of the CBCL: Internalizing ($r = -.23$), Externalizing ($r = -.23$), and Total ($r = -.25$). Number of siblings was associated with both the Externalizing ($r = -.26$) and Total ($r = -.25$) scales. Thus, a trend appears to be present in which the behaviour of children is rated as less and less problematic as the number of siblings increases. A small but significant and positive correlation was observed between the Internalizing scale and second-order ToM ($r = .20$).

Age, as expected, was significantly and positively correlated with most of the measures of interest: first-order ToM ($r = .45$), second-order ToM ($r = .57$), nonverbal ToM ($r = .49$), working memory ($r = .47$), and first-order sarcasm ($r = .54$). These correlations likely reflect the developmental nature of the tasks. A significant correlation was not obtained between age and either of the language measures. This reflects the fact that age-corrected scaled scores were used and that the two age groups achieved almost identical scaled scores.

Gender was not significantly correlated with any of the other 15 variables. This indicates that gender was in no way associated with performance on any of the tasks, or behaviour ratings, or the number of siblings, or diagnostic group membership.

Section 2 - Between Group Differences

Language. Table 3 presents the mean scale scores achieved on the CELF Concepts Subtest and the WISC Information subtest, broken down by age and diagnostic group. From Table 3 it is clear that the NSLI children in each age group performed at a level very close to the subtest's mean scale score. Also evident from Table 3 is that the SLI children in each age group performed at a level approximately one and a half standard deviations below the subtest's mean scale score. Thus it can be concluded that the children within each diagnostic category did indeed possess the language skills indicated by their category level: the NSLI children had language scores falling within the average range while the SLI children had scores falling significantly below the average range. A further conclusion to be drawn from Table 3 is that very little difference occurred between the mean language scores of the two age groups within each category. The mean language scores are age-corrected scale scores. Thus, the virtually identical scores achieved by the older and younger SLI children indicate that both age levels are equally impaired in their language achievement. Similarly, the virtually identical scores achieved by the older and younger NSLI children indicate that both age levels are equally developed in their language achievement.

To test for group differences in language ability, two one-way ANOVAs were run using children's scaled language scores on the WISC and the CELF as the dependent measures. The results of the one-way ANOVA comparing WISC scores revealed significant differences between the two diagnostic groups in the level of language achievement [$F$ (1,90) = 51.58, $p<.001$, $MS_e$ =

6.426]. The NSLI group performed significantly better (mean = 10.73) than the SLI group (mean = 6.93). The results of the ANOVA comparing CELF scores also revealed significant differences in language achievement between the two diagnostic groups [$\underline{F}$ (1,90) = 65.07, $\underline{p}$<.001, $\underline{MS_e}$ = 7.073]. The NSLI group performed significantly better (mean = 10.20) than the SLI group (mean = 5.72). Thus the diagnostic separation created prior to testing, based on the presence or absence of a SLI diagnosis, is validated by language scores achieved in this study.

First-order ToM . Table 4 presents the mean number of first-order ToM questions answered correctly, arranged by age and diagnostic group. From Table 4, it is clear that all children's performance is almost at ceiling level for the first-order control questions. The older children and the younger NSLI children performed at ceiling level while the younger SLI children were not far behind with a mean score of 9.57 out of a possible 11. This finding shows that the children had little difficulty answering the control questions and thus indicates that they comprehended the basic facts of the first-order stories. Also evident from Table 4 is that the older children in each diagnostic group outperformed all the younger children on the first-order ToM questions. Thus, an age effect is suggested for first-order ToM, with older children in both diagnostic groups showing better performance than younger counterparts. Another pertinent conclusion to be drawn from Table 4 is that the NSLI children outperformed the SLI children at both age levels indicating that first-order ToM may be slower to develop in children with SLI. A two-way ANOVA was run entering age and diagnostic group as the between group factors, and the first-order ToM score as the dependent variable. Results of the two-way ANOVA investigating first-order ToM revealed significant main effects for group [$\underline{F}$ (1,88) = 21.68, $\underline{p}$<.001] and age [$\underline{F}$ (1,88) = 29.12, $\underline{p}$<.001]. The NSLI group performed significantly better (mean = 4.37) than the SLI group (mean = 2.98). Similarly, the 6- and 7-year-olds outperformed (mean = 4.51) the 4- and 5-year-olds (mean = 2.96). No significant interaction was found between age and diagnostic group on first-order ToM [$F$ (1,88), = 1.94, $p$ = .17, $MS_e$ = 1.814].

Table 4 also reveals a similar profile of findings (age and group differences) occurring for the nonverbal ToM task. A two-way ANOVA was run entering age and diagnostic group as the between group factors, and the nonverbal ToM score as the dependent variable. The results of this analysis revealed significant main effects for group [$F$ (1,88) = 13.68, $p$<.001] and age [$F$ (1,88) = 30.85, $p$<.001]. The NSLI group outperformed (mean = 4.33) the SLI group (mean = 2.91) and the older children outperformed (mean = 4.71) the younger children (mean = 2.66). No significant interaction was found between age and diagnostic group on nonverbal ToM [$F$ < 1].

Second-order ToM. Table 5 presents the mean number of second-order ToM questions answered correctly by age and diagnostic group. From Table 5 it is clear that all children performed at a level above chance on the second-order control questions. The young SLI children answered approximately 4 out of 6 control questions correctly while all other children correctly answered approximately 5 out of 6. This suggests a good but slightly incomplete understanding of the basic facts of the stories. Also evident from Table 5 is that the older children in both diagnostic groups outperformed all the younger children on the second-order ToM questions. Thus an age effect is suggested for second-order ToM, with older children in both diagnostic groups showing improved performance over their younger counterparts. Another pertinent conclusion to be drawn from Table 5 is that the NSLI outperformed the SLI children at both age levels indicating that second-order ToM, like first-order ToM may be slower to develop in children with SLI. A two-way ANOVA was run entering age and diagnostic group as the between group factors and the second-order ToM score as the dependent variable. Significant main effects for group [$F$ (1,88) = 23.94, $p$<.001] and age [$F$ (1,88) = 63.12, $p$<.001] were observed. Once again, the NSLI group performed significantly better (mean = 2.82) than the SLI group (mean = 1.67) and the older children outperformed (mean = 3.20) the younger children (mean = 1.40). No significant interaction was found between age and diagnostic group on second-order ToM [$F < 1$].

Thus, clear age and group differences in favour of the NSLI children and the older children are revealed by their scores on the standard first-order, second-order, and nonverbal measures of ToM. The hypothesis of delayed ToM acquisition in children with SLI is thus supported.

Sarcasm. Table 6 presents the mean number of sarcasm questions answered correctly by age and diagnostic group and ToM composite scores arranged by age and diagnostic group. From Table 6 it is clear that all children performed almost at ceiling level for the control questions. This finding shows that the children had little difficulty answering the control questions and thus indicates that they comprehended the basic facts of the stories. Also evident from Table 6 is that the older children in both diagnostic groups achieved a higher mean score than all the younger children on the first-order and second-order sarcasm questions. Once again an age effect is suggested, with the older children in both diagnostic groups showing higher levels of performance over their younger counterparts on both first-order and second-order sarcasm questions. Table 6 also illustrates that the NSLI children outperformed the SLI children at both age levels on the first-order sarcasm task indicating that comprehension of first-order sarcasm may be slower to develop in children with SLI. This trend did not hold however on the second-order sarcasm questions wherein the older SLI children marginally outperformed the older NSLI children. Two two-way ANOVAs were run entering age and diagnostic group as the between group factors and the first-order and second-order sarcasm scores as dependent variables. Results of the two-way ANOVA investigating first-order sarcasm revealed significant main effects for group [$F$ (1,88) = 12.60, $p$<.001] and age [$F$ (1,88) = 28.02, $p$<.001]. The NSLI children outperformed (mean = 1.98) the SLI children (mean = 1.12) and the older children outperformed (mean = 2.22) the younger children (mean = 0.96). No significant interaction was found between age and diagnostic group on the first-order sarcasm task [$F$ < 1].

Results of the analysis of second-order sarcasm revealed no significant differences between age groups [$F$ (1,88), = 1.88, $p$ = .17] or diagnostic groups [$F$ <1]. No significant interaction was found between age and diagnostic group on the second-order sarcasm task [$F$ (1,88) = 1.04, $p$ = .31, $MS_e$ = .625]. Thus, age and diagnostic group differences occurred on the first-order sarcasm task indicating that children's ability to comprehend sarcasm increases with age and is slower to develop in children with SLI. Age and diagnostic group differences did not occur on the second-order sarcasm task. The overall low scores on this task suggest that the task was beyond the developmental level of this study's sample of children. This interpretation is supported by the previous findings of Ackerman (1981, 1986) and Keenan and Quigley (1999).

ToM composite. A ToM composite score was created by adding up the children's scores on all of the ToM tasks (first-order, second-order, first-order sarcasm, second-order sarcasm, nonverbal). Table 6 shows that the older children in both diagnostic groups achieved higher scores on the ToM composite than all the younger children. This is to be expected given the trend demonstrated on all individual ToM measures. Thus, the age effect is apparent in the ToM composite. Also apparent from Table 6 is the stronger performance of the NSLI children over the SLI children at both age levels on the ToM composite. A two-way ANOVA investigating group and age effects for the ToM composite score was conducted. Significant main effects for diagnostic group [$F$ (1,88) = 36.56, $p$<.001] and age group [$F$ (1,88) = 77.70, $p$<.001] were observed. The NSLI children performed significantly better (mean = 14.06) than the SLI group (mean = 9.16) and the older children outperformed (mean = 15.29) the younger children (mean = 8.40) on the composite of ToM tasks. No significant interaction was found between age and diagnostic group on the composite of ToM tasks [$F < 1$]. Thus, further support is given for age and group differences in favour of the older children and the NSLI children on ToM tasks.

The results thus far, offer strong support for the primary hypothesis of group differences between NSLI and SLI children on the cardinal measures of language and ToM. NSLI children significantly outperformed same-aged SLI children on both measures of language and on every measure of ToM other than second-order sarcasm wherein a floor effect appears to be in evidence.

Working memory and sibling measures. Table 7 presents the mean scores achieved on the spatial working memory task, the mean number of siblings living in the home, and the mean number of older siblings living in the home, arranged by age and diagnostic group. From Table 7 it can be seen that scores on the working memory task appear to improve with age within both diagnostic groups. Also apparent from Table 7 is that little difference appears to exist in the working memory scores between diagnostic groups. Results of a two-way ANOVA investigating working memory scores by age and diagnostic group revealed a significant main effect for age [$F(1,88) = 15.16, p < .001$]. The older children outperformed (mean = 4.96) the younger children (mean = 3.38) on the working memory task. No main effect was found for diagnostic group [$F < 1$], and no interaction effect was observed [$F < 1$].

Table 7 also shows some small differences in the mean number of siblings in the homes of all children. A two-way ANOVA investigating number of siblings by age and diagnostic group confirmed this and showed a significant main effect for diagnostic group [$F(1,88) = 7.38$,

$p < .01$]. The SLI children had significantly more siblings (mean = 1.86) than the NSLI children (mean = 1.25). No main effect was found for age group [$F < 1$], and no interaction effect was observed [$F(1,88) = 1.66$, $p = .02$, $MS_e = 1.179$]. A similar analysis investigating number of older siblings in the home revealed no main effects for age or diagnostic group and no interaction effect. Given that neither number of siblings nor number of older siblings sizeably correlated with any of the measures of ToM, the role of older siblings will not be further explored and the hypothesis of its possible predictive influence on ToM cannot be confirmed or supported through the data gathered in this study.

Behaviour. Table 8 presents the mean T scores achieved on the CBCL Internalizing, Externalizing, and Total scales as rated by parents and organized by age and diagnostic group. Given that the mean T score for each scale according to the standardization norms is 50 with a standard deviation of 10, it is apparent that on the whole, the means of both groups at each age level are soundly within the average range. Thus, neither group at either age is exhibiting significantly more behavioural symptoms than most children their age.

This equivalency between age and diagnostic groups on the behaviour variables makes moot any further comparisons involving this variable. Thus, the evidence does not support the hypotheses that behavioural problems occur alongside weak language and ToM ability or that behavioural problems move from favouring internalization to externalization as age increases.

Section 3 - Developmental course for ToM in SLI as in NSLI

The second primary hypothesis of ToM following a similar developmental course in SLI children as in NSLI children was investigated in two ways. The significance of the number of children passing first-order ToM and sarcasm tasks before passing second-order ToM and sarcasm tasks was assessed. Then the significance of the performance levels of younger versus older children on first-order and second-order ToM tasks is reviewed.

Diagnostic group performance on first-order ToM versus second-order ToM. The significance of the number of children passing first-order ToM before passing second-order ToM was assessed through a series of K-W Gammas, chi-squares, and McNemar statistics. The Gammas were calculated as a more stringent measure of the observed concordant and discordant responses in light of criticisms that the chi-square measure of statistical significance is problematic because the marginal relations between categorical variables is confounded with the interaction between these variables (see Nelson, 1984; Upton, 1978). To conduct these analyses subjects were categorized as passing or failing first-order and second-order ToM and sarcasm tasks. A perfect score or giving only one incorrect answer on an individual ToM task was considered a pass and assigned a score of 1. Two or more incorrect answers on any of the individual ToM tasks was considered a fail and assigned a score of 0.

Table 9 shows the number of NSLI children giving each of the 4 possible pass/fail patterns on the two levels of ToM. A chi square analysis of this table showed a high association (K-W Gamma = .85, $p<.01$) between children's performance on the first-order ToM task and their performance on the second-order ToM task [$\chi^2(1) = 7.85$, $p<.01$; Yates correction for continuity]. The NSLI children's performance on first-order ToM usefully predicts their performance on second-order ToM. The observed relationship between performance on first-order and second-order ToM is in accordance with that commonly found in the literature (Baron-Cohen, 1989; Perner et al., 1989; Wimmer & Perner, 1983). As can be seen from Table 9, 31 (63.3%) subjects correctly answered both questions and 7 (14.3%) subjects were incorrect on both questions. Table 9 also shows that 11 (22.4%) of the NSLI subjects passed only one of the

ToM tasks. Of these 11 subjects, 9 passed the first-order ToM but failed the second-order ToM, and only 2 subjects showed the reverse pattern. Thus, four and a half times as many show the pattern of passing first-order ToM only over passing second-order ToM only. A McNemar's chi square test performed on these data show that this pattern is significant [$\chi^2(1) = 3.27$, $p<.05$]. Therefore, when NSLI children are only able to pass one level of the ToM tasks, they are more likely to succeed on the first-order ToM task.

Table 9

Relationship Between Diagnostic Group and Number of Children Passing and Failing First-Order and Second-Order ToM

| | Second-Order ToM | | Second Order ToM | |
|---|---|---|---|---|
| | NSLI | | SLI | |
| | Fail | Pass | Fail | Pass |
| First-Order ToM | | | | |
| Fail | 7 (14.3%) | 2 (4.1%) | 23 (53.5%) | 1 (2.3%) |
| Pass | 9 (18.4%) | 31 (63.3%) | 9 (20.9%) | 10 (23.3%) |

Table 9 also shows the number of SLI children giving each of the 4 possible pass/fail patterns on the two levels of ToM. A chi square analysis of this table showed a high association (K-W Gamma = .92, $p<.001$) between children's performance on the first-order ToM task and their performance on the second-order ToM task [$\chi^2(1) = 10.66$, $p<.001$; Yates correction for continuity]. This indicates that the SLI children's performance on first-order ToM usefully predicts their performance on second-order ToM. The relationship between performance on first-order and second-order ToM is the same as that demonstrated by the NSLI children in this study. As can be seen from Table 9, 10 (23.3%) subjects correctly answered both questions and 23 (53.5%) subjects were incorrect on both questions. Table 9 also shows that 10 (23.3%) of the SLI subjects passed only one of the ToM tasks. Of these 10 subjects, 9 passed the first-order ToM but failed the second-order ToM, and only 1 subject showed the reverse pattern. Thus, nine times as many show the pattern of passing first-order ToM only over passing second-order ToM only. A McNemar's chi square test performed on these data show that this pattern is significant [$\chi^2(1) = 4.9$, $p < .05$]. Therefore, when SLI children are only able to pass one level of the ToM tasks, they are significantly more likely to succeed on the first-order ToM task.

Diagnostic group performance on first-order versus second-order sarcasm. Subjects were again categorized as passing (score = 1) or failing (score = 0) first-order and second-order sarcasm tasks using the scores of 1 and 0 assigned to each child's performance on every ToM task when constructing the ToM composite score.

Table 10 shows the number of NSLI children giving each of the 4 possible pass/fail patterns on the two levels of sarcasm. A chi square analysis of this table showed no association (K-W Gamma = -.13, $p = .76$) between children's performance on the first-order sarcasm task and their performance on the second-order sarcasm task [$\chi^2(1) = .00$, $p < .1.00$; Yates correction for continuity]. This is likely because of the small number of subjects passing the second-order task. Performance on first-order sarcasm did not reliably predict performance on second-order sarcasm for the NSLI children. The finding is in contrast to the literature (Ackerman, 1981, 1986; Demorest, Meyer, Phelps, Gardner, & Winner, 1984; Perner & Wimmer, 1985). As can be seen from Table 10, 4 (8.2%) subjects correctly answered both questions and 15 (30.6%) subjects were incorrect on both questions. Table 10 also shows that 30 (61.2%) of the NSLI subjects passed only one of the sarcasm tasks. Of these 30 subjects, 27 passed the first-order sarcasm but failed the second-order sarcasm, and only 3 subjects showed the reverse pattern. Thus, nine times as many show the pattern of passing first-order sarcasm only over passing second-order sarcasm only. A McNemar's chi square test performed on these data show that this pattern is significant [$\chi^2(1) = 17.63$, $p<.001$]. Therefore, when NSLI pass only one sarcasm task, they are significantly more likely to succeed on the first-order than on the second-order task.

Table 10

Relationship Between Diagnostic Group and Number of Children Passing and Failing First- and Second- Order Sarcasm

| | Second-Order Sarcasm | | Second-Order Sarcasm | |
|---|---|---|---|---|
| | NSLI | | SLI | |
| | Fail | Pass | Fail | Pass |
| First-Order Sarcasm | | | | |
| Fail | 15 (30.6%) | 3 (6.1%) | 25 (58.1%) | 1 (2.3%) |
| Pass | 27 (55.1%) | 4 (8.2%) | 14 (32.6%) | 3 (7.0%) |

Table 10 also shows the number of SLI children giving each of the 4 possible pass/fail patterns on the two levels of sarcasm. A chi square analysis of this table showed no association (K-W Gamma = .69, $p = .17$) between children's performance on the first-order sarcasm question and their performance on the second-order sarcasm question [$\chi^2(1) = .97$, $p<.32$; Yates correction for continuity]. This is likely because of the small number of subjects passing the second-order task, and indicates that performance on first-order sarcasm did not reliably predict performance on second-order sarcasm for the SLI children. The pattern is the same as that demonstrated by the NSLI subjects in this study, but, is in contrast to the pattern generally found in other research studies (Ackerman, 1981, 1986; Demorest, Meyer, Phelps, Gardner & Winner, 1984; Perner & Wimmer, 1985). As can be seen from Table 10, 3 (7.0%) subjects correctly answered both questions and 25 (58.1%) subjects were incorrect on both questions. Table 10 also shows that 15 (34.9%) of the SLI subjects passed only one of the sarcasm tasks. Of these 15 subjects, 14 passed the first-order sarcasm but failed the second-order sarcasm, and only 1 subject showed the reverse pattern. Thus fourteen times as many show the pattern of passing first-order sarcasm only over passing second-order sarcasm only. A McNemar's chi square test performed on these data show that this pattern is significant [$\chi^2(1) = 9.6, p<.01$]. Therefore, when SLI children are only able to pass one level of the sarcasm tasks, they are significantly more likely to succeed on the first-order sarcasm task.

Overall, the results of the series of chi-square analyses offer evidence of a developmental progression of ToM in SLI children which is similar to that found in NSLI children. Although the two diagnostic groups are performing at different levels of accuracy as indicated by the means (see Table 1), the pattern of passing first- before second-order ToM tasks holds constant across diagnostic groups. Consider this finding in conjunction with the results of the two-way ANOVAs investigating age and diagnostic group differences on all ToM tasks. The older NSLI and SLI children both consistently outperformed their younger counterparts on tasks measuring nonverbal ToM, first- and second-order ToM and first-order sarcasm. The similarities between the ToM progression of SLI and NSLI children now presents as even more pronounced.

Section 4 - Predictors of ToM

The ability of working memory, language, and age to predict ToM performance in SLI children and in NSLI children was assessed through multiple regression. The ToM composite score of the children was used as the dependent variable in the regression results to follow. Due to the absence of sizable correlations between the sibling and ToM measures (see Table 2), the ability of family size to predict ToM was not assessed.

A simultaneous multiple regression was performed to determine if the ToM composite scores of the NSLI children could be predicted as a function of their working memory scores, CELF and WISC scores, and age group. The four variables together explained 63% of the variance in the ToM composite score for the NSLI children [$R = .79$, $F(4,44) = 18.39$, $p < .001$]. Only two of the four variables were found to contribute significantly to the prediction of the ToM composite: age group ($\beta = .69$, $t = 6.98$, $p < .001$) and CELF subtest score ($\beta = .25$, $t = 2.41$, $p < .02$). A simultaneous multiple regression was next conducted to determine if the ToM composite scores of the SLI children could be predicted as a function of their working memory scores, CELF and WISC scores, and age group. The four variables together accounted for 62% of the variance in ToM composite scores of the SLI children [$R = .79$, $F(4,38) = 15.61$, $p < .001$]. Three of the four variables were found to contribute significantly to the prediction of the ToM composite: age group ($\beta = .52$, $t = 4.59$, $p < .001$), CELF subtest score ($\beta = .29$, $t = 2.50$, $p < .01$), and working memory ($\beta = .26$, $t = 2.12$, $p < .03$). Of interest then is that the combination of four variables accounts for the same amount of variance in the ToM composite scores in each diagnostic group (62%). It is age and language that make significant contributions to ToM composite scores in the NSLI group, while it is age and language supplemented with working memory that make significant contributions to the ToM composite scores of SLI children. The

WISC subtest did not contribute any unique variance likely because of its strong correlation with the CELF subtest ($r = .64$). These data provide support for the current hypothesis that working memory, language, and age serve as predictors of ToM in SLI children.

The results as given thus far clearly depict connections between age and ToM ability (older children outperform younger children) and between diagnostic group and ToM ability (NSLI children outperform SLI children). In both cases the prevailing group has more language reflecting either their greater age or their diagnostic grouping. The use of age-corrected language scores foreclosed the possibility of determining the role of absolute level of language in ToM performance: younger and older children could not be compared across diagnostic groups (i.e., older SLI could not be compared to younger NSLI as their age-corrected scores were 5 versus 10 respectively, although their absolute level of language development may have been dead even). A substitute estimate of linguistic competence was therefore derived from the cumulative total of each subject's scores on the control questions for the ToM tasks. The composite control question score can be considered an approximate measure of linguistic competence in that it represents the children's ability to communicate their comprehension of the basic language, facts, and content of the ToM tasks. This control question composite is also an absolute estimate of linguistic competence and so scores can be compared across age groups.

A sequential multiple regression was run where the ToM composite score was used as the criterion variable and the composite control question scores of all subjects were entered first, followed by their scores on the memory task and their diagnostic group. When the composite control question score alone was entered, the regression equation explained 46.30% of the variance in ToM understanding, [$R = .68$, $F(1,90) = 77.58$, $p<.001$]. When diagnostic group and memory scores were added, the resulting equation explained an additional 11.70% of the variance, [$\underline{R}$ = .76, $\underline{F}$ (2,88) = 12.25, $\underline{p}$<.001]. All three of the variables contributed significantly to the ToM composite score: control question composite (β =.49, $\underline{t}$ = 6.22, $\underline{p}$<.001), diagnostic group (β =.-.24, $\underline{t}$ = -3.18, $\underline{p}$<.01), and memory (β = .30, $\underline{t}$ = 4.14, $\underline{p}$<.001). Importantly, when the effects of linguistic competence were removed from the equation, diagnostic group continued to influence level of ToM ability. This suggests that differences in absolute level of linguistic competence alone do not account for the differences in ToM understanding that exist between SLI and NSLI children. Linguistic competence contributes to the differences in ToM ability found across ages but so does memory and something else that is inherent in being NSLI or SLI.

<u>Section 5 - Relationship of Behaviour to Language, ToM and Age</u>

Significant differences were not found between age or diagnostic groups on the administered behaviour measure. The means of the children's T scores, at both age levels and in both diagnostic groups, on the three CBCL scales were very much clustered around the test's mean of 50 (see Table 8). As such, direct and solid support was not gathered for any of the hypotheses concerning behaviour.

Overlooking the group means, subjects were assigned to one of two groups depending on whether their three CBCL scale scores were classified as normal (T score ≤ 59) or significantly elevated (T score ≥ 60). The behavioural data were then analyzed for significance using this group membership classification system.

<u>Language and age.</u> Applying a chi-square test to the number of children from each age group in each of the two levels of behaviour ratings, gave non-significant results. Thus no association exits between age and elevated behaviour scores on either the Internalizing, Externalizing, or Total Scale of the CBCL. Applying a chi-square test to the number of children

from each diagnostic group in each of the two levels of behaviour ratings also gave non-significant results for the Internalizing, Externalizing, and Total Scales.

ToM. Using the pass/fail ToM classification system described earlier, level of ToM ability was compared to level of behavioural problems. To test for group differences (passers versus failers) in level of behavioural ratings a series of one-way ANOVAs was run using children's categorical scores as the dependent measures. No significant differences were found for the level of Externalizing or Total scores between those children passing and failing ToM tasks. The level of Internalizing scores was non-significant apart from first-order ToM approaching significance ($p < .09$).

Section 6 - Incidental Questions and Analyses

The observed negative correlations between number of siblings and the CBCL Externalizing and Total scores was something of a surprise finding. Despite the overall homogeneity of the CBCL scores across diagnostic groups and ages a brief analysis of this association was conducted. A simultaneous multiple regression investigating the ability of number of siblings to predict Externalizing and Total scores on the CBCL produced non-significant results. So, although parental perception of a child's behaviour symptoms is significantly associated with the number of siblings in the home, number of siblings does not significantly contribute to the prediction of this rating of behaviour problems.

# Chapter 4 - Discussion

## Support for Hypotheses

Hypothesis 1. A primary hypothesis of this study was that SLI children would exhibit weaker ToM ability than normally developing children of the same age. Strong support is provided for this hypothesis with SLI children in both age groups consistently providing fewer correct responses on ToM tasks than the age-matched NSLI children. This pattern was strong and held with the NSLI children clearly outperforming the SLI children on four of the five given ToM measures: standard first- and second-order ToM, nonverbal first-order ToM, and first-order sarcasm. Because of floor effects, neither age nor diagnostic group differences occurred on the second-order sarcasm task. The majority of children performed poorly on this task - both age and diagnostic groups correctly answered less than 25% of the questions. This poor performance is interpreted as indicating that the task was above the developmental level of the subjects. Such an interpretation is in accord with the findings of Ackerman (1981, 1986) and Keenan and Quigley (1999).

The finding of impaired ToM performance by SLI children is in direct contrast to the two studies (Leslie & Firth, 1988; Perner et al., 1989) widely cited in the literature (Eisenmajer & Prior, 1991; Peterson & Siegal, 1997), which found the performance of SLI children on ToM tasks to not be significantly different from that of normally developing peers. Leslie and Frith (1988) concluded that language delay plays no role in failure to understand mental states since the children with specific language impairment in their study performed at ceiling.

The current results likely differ from those by Leslie and Frith (1988) and Perner et al. (1989) because of design differences. Previous investigations addressing the issue of ToM ability in SLI children failed to employ age appropriate tasks when assessing for ToM ability.

Typically first-order ToM tasks were administered to SLI children aged 7 years or older. Good performance on these tasks was then interpreted as indicating unimpaired development. However, given that first-order ToM is generally mastered by 6 years of age, such results cannot fairly be interpreted as demonstrating flawless ToM development. The previous studies in essence gave the wrong level of task to the wrong aged children. The present study however, administered both first- and second-order ToM tasks to children who ranged in age from 4 years 0 months to 8 years 1 month. This specifically covers the ages at which first-order ToM emerges (4 years) and is mastered (6 years) and the ages at which second-order ToM emerges (6 years) and approaches mastery (7 years). The results of this study therefore demonstrate that when suitable tasks and aged subjects are utilized, a deficit in ToM performance is in evidence for children with SLI.

The suitability of the chosen tasks and subject ages in this study is amply demonstrated by the following results. The older NSLI children achieved 98.40% and 94.00% on first- and second-order ToM tasks respectively while the younger NSLI children achieved 76.00% and 45.75% respectively. These achievement levels coincide with what is known about the milestones of first- and second-order ToM in normally developing children (Perner et al., 1987; Perner & Wimmer, 1985; Sullivan et al., 1995; Wimmer & Perner, 1983). The numbers provided by the SLI subjects tell a different story. The older SLI children achieved 80.00% and 62.50% on first- and second-order ToM tasks respectively while the younger SLI children achieved 41.80% and 24.00% respectively. Thus, SLI children had poorer performances than age-matched NSLI peers on the ToM tasks. Consequently, it can be concluded that ToM ability is affected by language impairment. These findings contradict the previous consensus in the literature and are likely more valid because of the use of appropriately aged children for the tasks.

This finding of delayed ToM development in SLI children is, however, consistent with the body of research examining the relationship between language and ToM ability in normally developing children. Here evidence has been gathered that demonstrates language ability to be consistently associated with ToM performance and to be one of the strongest predictors of ToM ability (Astington & Jenkins, 1999; Jenkins & Astington, 1996; Ruffman et al., 1999). The present results also conform to the findings from the literature on deaf children. Deaf children as a group have demonstrated striking delays in the ability to understand others' mental states as measured by standard ToM tasks (Peterson & Siegal, 1995, 1997, 1999; Russell et al., 1998). The degree of delay has been shown to vary, depending on the deaf child's opportunity to communicate fluently and frequently with other persons in their immediate environment. Deaf children raised in households with at least one fluent signer outperform their deaf peers raised in environments devoid of fluent signers on standard ToM tasks (Peterson & Siegal, 1995, 1999; Russell et al., 1998). As well, native signers, oral deaf children, and normal hearing children have been found to perform similarly on ToM tasks and to outperform signing deaf children from hearing families (Peterson & Siegal, 1999). The finding that ToM development is delayed in children with SLI is also in line with the works investigating ToM ability in autistic children. Language ability and verbal mental age demonstrably differ between those autistic children who can and those who cannot succeed on ToM tasks (Eisenmajer & Prior, 1991; Happé, 1993, 1995; Tager-Flusberg & Sullivan, 1994a).

The results of this study also provide support for the hypothesis that language ability is instrumental in both the *expression* and *conceptualization* of ToM. Age and diagnostic group differences both were observed on the nonverbal measure of first-order ToM. Older children outperformed their younger counterparts within each diagnostic group. The NSLI children once again clearly outperformed their SLI peers. These group differences mirror those discovered on the standard (verbal) measures of first-order ToM. In other words, SLI children continue to exhibit a lag in ToM development even when language is removed from the presentation of the task and when language is unnecessary in formulating a response. The vast majority of ToM investigations to date have employed tasks that are highly verbal in nature (Baron-Cohen, 1989; Jenkins & Astington, 1996; Keenan et al., 1998; Wimmer & Perner, 1983). It is standard for ToM tasks to be presented through words in complement with visual aids. Responses to these standard tasks require verbal utterances or sometimes a pointing response to the visual support.

The present study largely removed language from a first-order ToM task and used two groups of children with significantly different language abilities. This helped to more clearly define the role of language in successfully completing ToM tasks. In other words, is language required simply to communicate an understanding of ToM or does it also play a role in formulating an understanding of ToM? The discrepancy in favour of the NSLI children on the nonverbal ToM task strongly suggests that the formulation of ToM is influenced by level of language development. This finding of poor performance on a nonverbal measure of ToM by children with SLI provides insight into the complex manner in which language affects ToM development. Previous research could only define the role language plays in the *expression* of an understanding of ToM. The current study confirms this and highlights the role language has to play in *formulating* an understanding of ToM. SLI children are less able than normally developing children to succeed on even nonverbal ToM tasks. Thus they are distinctly different from their age-matched peers in their ability to both *express* an understanding of the concept of ToM and their ability to *form* an understanding of the concept of ToM.

Overall, the results indicate that there *is* a delay in ToM development for this clinically distinct group of children. The influence of language ability on ToM ability has been extended beyond the population of normally developing children to a clinically disordered group of children. This is a substantial contribution to the research investigating children's understanding of mental states as it increases the generalizability of the finding that language plays a role in reading others' minds. Furthermore, the demonstration of delayed ToM development in a sample of SLI children provides additional evidence to substantiate the argument that SLI may fit somewhere on the very mild end of a continuum which has the disorder of autism as one of its anchors (Konstanareas & Beitchman, 1996). Difficulty understanding other's mental states might now be added to the list of social and communicative impairments shared, in varying degrees, by individuals with autism, Asperger's syndrome, and specific developmental language disorder.

The poor performance of the SLI children on this battery of ToM tasks also makes a contribution to the literature documenting their social skills. Difficulty gaining access to on-going social interactions (Brinton, Fujiki, Spencer, & Robinson, 1997; Craig & Washington, 1993), difficulty with turn-taking and topic maintenance (Brinton, Fujiki, & Powell, 1997), and poor negotiation strategies (Brinton, Fujiki, & McKee, 1998) have all been noted in children with SLI relative to their age- and language-matched peers. Difficulty understanding the mental states of others may also deserve a place on this list given the findings of the current study.

If not the difference in language ability between the two sampled groups of children, what else might account for this finding of delayed ToM ability in SLI children? Lowered intelligence has certainly been found to be associated with depressed ToM performance as made clear in studies employing Down syndrome subjects and subjects with mental retardation (Baron-Cohen, 1989; Tager-Flusberg & Sullivan, 1994b). The possibility exists that the sample of SLI children was on the whole less cognitively able than the sample of NSLI children.. Because individual measures of cognitive ability were not given to the present sample, this explanation cannot be entirely ruled out. However, several factors combine to make this an unlikely explanation. No SLI subject was consciously selected who had demonstrated marked academic difficulty in areas other than language. Teachers and Speech-Language Pathologists bore this criteria in mind when identifying potential subjects. Further, the recruitment of NSLI children was based on the process of parental interest and consent. Although there may have been a selection bias in that only those parents who felt their children would fare well agreed to participate, this bias would also be in effect for the SLI children who made it past the teacher and Speech-Language Pathologist screening criteria. Finally, the similar performance level of both diagnostic groups on the working memory task provides a rough estimate of cognitive ability and suggests that the two groups were not significantly different in this domain.

Processing capacity presents an another potential explanation for the observed ToM discrepancy in favour of the NSLI subjects. A variety of studies utilizing verbal and nonverbal tasks designed to assess the link between working memory load and language ability have revealed that as the memory and representational demands of a task increase, the performance of SLI children drops further and further below that of their age-matched peers (Bishop & Adams, 1992; Katz, Curtiss, & Tallal, 1992; Records, Tomblin, & Buckwalter, 1995). Compelling as this explanation is, it cannot be substantiated within the current data set. The performance level of the SLI children on the working memory task was not significantly different from that of the NSLI children in either age group.

The diagnostic group differences on the nonverbal ToM task might reflect the fact that the task could not be comprehended without an understanding of the accompanying verbal instructions. If this were the case then the SLI children would be at a disadvantage in performing the task. This is unlikely given that the task is modelled after the Picture Arrangement task of the Wechsler scales and the task is declared as suitable for administration to populations with no or limited language such as deaf children or hard-of-hearing children whose first language is not English (Sattler, 1988). What makes this even more unlikely an explanation is the fact that the training item on the nonverbal task was worked through until each subject had constructed it correctly, either on their own or with assistance.

Hypothesis 2. The current results provide strong support for a development of ToM ability in SLI children that unfolds gradually and in an order similar to that displayed by normally developing children. The developmental literature shows that understanding another person's belief (first-order ToM) occurs prior to understanding another person's belief about a third person's belief (second-order ToM) (Sullivan, Winner, & Hopfield, 1995; Winner & Leekham, 1991). Two significant trends, presented in the performance of the SLI children in this study, demonstrate this same sequence of ToM development. First, the consistently superior performance of the older children in both diagnostic groups on every ToM task supports a gradual and step-wise unfolding of ToM in SLI children that is similar to that demonstrated by NSLI children. This link between age and enhanced ToM performance is commonly illustrated in studies with autistic children (Baron-Cohen, 1989; Eisenmajer & Prior, 1991; Happé, 1995) and deaf children (Peterson & Siegal, 1995; Russell et al., 1998).

Second, the ratio of success on first-order ToM tasks to second-order ToM tasks depicts a progression toward ToM mastery. McNemar chi-square analyses revealed that when subjects were only able to pass one level of standard ToM tasks, they were significantly more likely to succeed on a first-order ToM task. This trend was exhibited whether the child was NSLI or SLI, but was most pronounced in the SLI group. The pattern of passing only first-order ToM over passing only second-order ToM occurred for four and a half times as many NSLI children and nine times as many SLI children. This same trend occurred between the two levels of sarcasm tasks - the unconventional measure of ToM. Nine times as many NSLI subjects and fourteen times as many SLI subjects passed only first-order sarcasm tasks than passed only second-order sarcasm tasks if able to pass only one level. ToM development thus proceeds in the same sequence but with less alacrity for SLI children than it does for NSLI children. This suggests that the mechanism(s) for developing ToM is intact and progressing, but at a slower rate of development.

Interestingly, this pattern is similar to that discovered in the research on ToM development in children with autism. The success/failure ratio is even more extreme in this population with the vast majority of autistic children aged younger than 18 failing first-order ToM tasks and an even greater number failing second-order ToM tasks (Baron-Cohen, 1989; Ozonoff, Pennington, & Rogers, 1991). The autism-ToM literature is moving toward the conclusion that autistic children routinely fail developmentally appropriate ToM tasks because they require “more age” (and therefore more language) than other children to be able to succeed on ToM tasks. The results of the present study suggest a similar though less severe requirement for children with SLI.

The support for the above hypothesis then provides evidence for the idea that SLI children’s ability to understand mental state attributions is not defective but merely delayed. It also adds support to the argument that an understanding of others' mental states is anything but an all-or-none process (Dunn, 1995; Astington & Jenkins, 1995).

Alternative explanations for these findings and their suitability are the same as those discussed for hypothesis 1.

Hypothesis 3. It was hypothesized that the constructs known to predict ToM ability in normally developing children would also serve to predict ToM ability in SLI children. Age, language, family size, and working memory were therefore all anticipated as serving a weighty role in the ToM performance of the SLI subjects. No specific predictions regarding the *order* of strength of the other predictor variables were made due to the early stages of the SLI-ToM literature in this respect.

In fact, family size was not significantly correlated with any ToM variables. This lack of association between number of siblings and ToM achievement contrasts with the literature which has generally found performance on ToM tasks to improve as the number of siblings increases (Jenkins & Astington, 1996; Perner et al., 1994; Ruffman et al., 1998). It is difficult to understand why number of siblings was not correlated with level of ToM performance in either group of children in this study. The lack of association is especially surprising given the study by Jenkins and Astington (1996) which found number of siblings to be important for the ToM performance of children but particularly so for those with poor language skills.

It is easy to generate reasons as to why number siblings might not correlate with the ToM ability of SLI children. It may be that the language skills of the SLI children in the current sample are so poorly developed as to prevent them from benefiting from sibling interactions involving mental state language and situations. A particular level of language competence may be necessary in order for children to benefit from the exposure and experience that comes from living with a larger number of siblings. Just as a certain level of linguistic competence (as measured by the TELD) has been demonstrated as necessary to pass various ToM tasks (Jenkins & Astington, 1996), a given amount of language ability might be necessary to learn about ToM or to begin thinking about it. Alternatively, the poor language and social skills of the SLI children may serve to limit their interaction time with normally developing siblings who are frustrated with their "oddly" behaving brother or sister. Certainly there is evidence to show that SLI children ignore peers' conversational initiations and have their own initiations ignored more often than is the case for children with intact language skills (Brinton et al., 1997; Hadley & Rice, 1991). In a similar vein, studies have shown adults to demonstrate systematic biases toward children with limited communication abilities (Burroughs & Tomblin, 1990; Rice, Hadley, & Alexander, 1993). This has yet to be demonstrated with siblings, so the likelihood remains unknown but possible.

It is much more difficult to generate reasons as to the lack of association between number of siblings and ToM development in the NSLI sample, seeing as this is a much replicated finding in the literature. The lack of observed association between ToM ability and number of siblings may have been created by the varied economic backgrounds of the current subject sample. Subjects were gathered from numerous and dissimilar regions of Ontario. Cole and Mitchell (2000) examined the relationship among the variables of siblings and ToM and executive control abilities in children 3 to 5 years of age. A positive association was not observed between number of sibling and ToM performance. A secondary study and subsequent analyses revealed a possible corrupting influence of socio-economic status on the usual finding of sibling advantage for ToM.

Alternatively, the lack of association between number of siblings and performance on ToM tasks may to some extent reflect design differences. The studies finding a positive association (Jenkins & Astington, 1996; Perner et al., 1994; Ruffman et al. 1998) were conducted with children primarily between 2 and 5 years of age, and employed only first-order ToM tasks. Differences in age and task complexity might therefore be responsible for the discrepant findings of the current study. Number of siblings might be influential in *initiating* a first-order understanding of others' minds at an early age (2, 3 or 4 years), but this effect is muted at later ages (5 to 8 years) by which point most children have acquired first-order ToM as a matter of course and not necessarily with the benefit of siblings. This interpretation receives some support in light of Anderson's (1998) failure to find an association between ToM development and number of siblings in a sample of 6- and 7- year old children. It is interesting to take this explanation a step further to accommodate the observed absence of an association between second-order ToM and siblings in the present study. This finding might represent the fact that number of siblings is influential in *initiating* an understanding of others' minds, but does not effect the *progression* or maturation of a theory of others' minds. Studies are needed which investigate this possibility. The role of siblings in the development of ToM ability in SLI children should therefore not be ruled out. Rather, it should be re-visited in a future study incorporating a larger number of subjects and more complex ToM tasks.

Age made the strongest contribution toward predicting ToM performance (composite) in both the NSLI and SLI group. It also supports the general finding from the literature on normally developing children; older children are more able to pass higher order mental attribution tasks than younger children (Sullivan et al., 1995; Winner & Leekham, 1991). It also goes along well with the findings that both autistic children (Baron-Cohen, 1989; Eisenmajer & Prior, 1991; Happé, 1995) and deaf children (Peterson & Siegal, 1995; Russell et al., 1998) show heightened levels of accuracy on ToM tasks at older ages. The finding that age is a solid predictor of ToM performance is not new or unexpected. Nor does it provide a unique or particularly useful explanation. Still it is useful to find that both sampled populations are "normal" in the sense that their performance on ToM tasks increases as they age. This finding simply falls in line with what is generally known about ToM - ToM is a developmental milestone that is highly associated with age.

Language also made a solid contribution to the prediction of ToM performance (composite) in both the NSLI and SLI group. In fact, language ability provided the second greatest influence for each group. This finding is in agreement with much of the literature (all discussed earlier) investigating ToM ability in normally developing children (Astington & Jenkins, 1999; Jenkins & Astington, 1996; Ruffman et al., 1999), deaf children (Peterson & Siegal, 1995, 1997, 1999; Russell et al., 1998) and autistic children (Eisenmajer & Prior, 1991; Happé, 1993,1995; Tager-Flusberg & Sullivan, 1994a). Interestingly, language made its contribution through the CELF subtest score and not the WISC subtest score. It was the Concepts and Directions subtest of the CELF, not the Information subtest of the WISC, which made the contribution. This was consistent across both diagnostic groups. This finding contrasts mildly with that of Ruffman et al. (1999), who found that while these two language measures shared variance in predicting ToM, the Information subtest had a slight edge in predictive value. The superiority of the Concepts and Directions subtest over the Information subtest however is in line with the thinking and findings of Tager-Flusberg and Sullivan (1994a) and Astington and Jenkins (1999). Both pairs of researchers asserted that the key linguistic ability involved in decoding ToM tasks is syntactic understanding. They argue that an understanding of embedded clauses (or sentential complements) is a requirement to succeeding on ToM questions. Not only is the Concepts and Directions subtest a syntactic measure, its items are also more apt to contain sentential complements than are the items on the Information subtest which are semantic in nature. What this seems to demonstrate is that the particular language skills assessed by the Concepts and Directions subtest of the CELF are more integral to understanding ToM tasks than the particular elements of language assessed by the Information subtest of the WISC. More likely what is being demonstrated is that the two language measures are so highly correlated that once one has been entered into the predictive equation, all the variance that is common to the two measures has been accounted for leaving very little unique variance for the remaining measure to represent.

An obvious design difference between this investigation and those that came before it is the inclusion of a nonverbal measure of ToM. Possibly, the inclusion of a nonverbal measure of ToM influenced the superiority of the CELF subtest over the WISC subtest to predict ToM performance (composite). This makes sense when one considers that the items on the CELF were heavily invested with the requirement to organize components in a specific order and the nonverbal measure of ToM also required organizing and sequencing pieces of information in a specific order. This element of sequencing and active processing is not so apparently contained within the WISC subtest items.

The prediction that working memory would serve as a significant predictor of ToM ability in SLI children was supported. Working memory along with age and language accounted for 62% of the variance in their ToM composite scores. The contribution of working memory ($\beta = .26$) was similar to that of language (CELF subtest, $\beta = .29$) but much less than that of age ($\beta = .52$). Thus, for SLI children, after the benefits inherent in increased age (e.g., neurological maturity, attention span, experience), the ability to handle the syntactic components of language and the ability to hold and manipulate, in mind, multiple pieces of information influences successful understanding of other minds.

This finding represents a novel contribution to the literature. To date no other studies have examined working memory in relation to the ToM ability of SLI children. However, considered alongside the normative literature, this finding is not out of place. Davis and Pratt (1995), using a sample of children 3 to 5 years of age, found that performance on a working memory task (backwards digit span) significantly predicted performance on first-order ToM tasks after controlling for age and verbal skills. Working memory was calculated as providing 6% unique variance. Also, Keenan et al. (1998) determined that the performance of a sample of children 3 to 5 years of age on a working memory measure (counting span task) was a significant predictor of their correct performance on a first-order false belief composite measure after controlling for the effects of age. Working memory contributed 7.4% unique variance. Keenan (1998) then replicated and extended these findings with the addition of a language measure. He found that performance of a sample of children 4 to 5 years of age on a working memory measure (counting span task) predicted their performance on a false belief composite measure after controlling for children's age and individual differences in language ability. Hierarchical regression analysis showed that working memory uniquely accounted for 21% of the variance in false belief understanding. These three studies using two different measures of working memory produced congruent findings - a relationship between children's false belief understanding and developmental increases in working memory.

Less direct evidence was garnered by a study conducted by Sullivan et al. (1994). They found that when the information processing demands (length, format, and story complexity) of a second-order ToM task were sizably reduced, approximately fifty percent of a pre-school aged sample of children was able to correctly answer second-order ToM questions. This dropped the age at which second-order success is generally achieved from 6 or 7 years of age down to 3 or 4 years of age. The present results are also in line with a study investigating information processing capacity and ToM ability in atypical children. Tager-Flusberg and Sullivan (1994b) showed that a portion of autistic subjects were capable of passing a second-order belief task when given a version with the information processing demands reduced.

The finding of the present study is thus in agreement with studies on normally and atypically developing children: working memory capacity is related to successful ToM performance. The finding of the present study is also an extension of research findings to date as it replicates the role of working memory in ToM performance in a different atypical group of children. Given that the same factors (age, language, and working memory) repeatedly emerge as influencing ToM in various populations of children, and that these populations show success on ToM tasks at later points in their development, surely suggests that the mechanism known as ToM is not broken or non-existent, but merely late to emerge.

Although working memory was positively and significantly correlated with all measures of ToM, it did not contribute significantly to the prediction of the ToM performance scores (composite) of the NSLI children. This is interesting given that there were no significant differences between the working memory scores of the two diagnostic groups; only between the two age groups. Yet, age and language accounted for the same amount of variance in the ToM scores of NSLI children as did age, language, and working memory for the SLI children. It seems that where age and language fall short in SLI children, working memory steps in to pick up the slack in predicting ToM performance. It is unclear why working memory did not make a significant contribution to the ToM performance of the NSLI children. Such a finding is very much at odds with the normative literature cited above (Davis & Pratt, 1995; Keenan et al., 1998; Keenan, 1998). The work of these authors all produced findings to show that working memory contributed to ToM performance after the influence of age and language were controlled for. Their subjects were younger than those used in the present study. Perhaps then, working memory is only influential in the initial development of ToM. Its role may fade with age as the child becomes a more competent processor and is able to conduct some tasks and abilities in an automatized manner. This interpretation is consistent with the present findings which show that the SLI children are delayed in their ToM development (demonstrated by their poorer scores on all ToM tasks). Thus, they may still be at the stage where a reliance on working memory is necessary to successfully complete ToM tasks. Good support exists for this in studies by White (2000) and Keenan (2000). These researchers tested 3- to 5-year-olds and found the expected relationship between working memory and ToM performance. However, when the data were analyzed separately for each age group, the working memory-ToM association proved to be strongest in the 3- to 4-year old group. The authors interpreted this as meaning that working memory is not required for successful ToM performance as children become more expert at thinking about people's representational states.

Following this line of reasoning, the present study included second-order ToM tasks as part of the ToM battery. Research has not been conducted that specifically investigates the role of working memory on second-order ToM tasks. Perhaps it does not have a role. Indirect evidence by Sullivan et al. (1994) and Tager-Flusberg and Sullivan (1994b) as reviewed above suggests that this is not the case. However, it cannot be properly ruled out without further investigation. Supposing that the suggested reason is indeed the case - working memory no longer has a significant role to play in successful ToM performance as children reach mastery -

what would this mean is happening as the SLI children struggle to complete the ToM tasks? It may be that SLI children need to activate or draw upon their working memory in order to succeed on ToM tasks because their language is underdeveloped for their age and not able to handle the demands of the task on its own. The ability of SLI children to simultaneously hold and process information thus serves as a predictor of ToM ability. Therefore, the possibility exists that SLI children are using working memory to solve the given ToM tasks and this results in a relationship between their ToM performance and their working memory scores. The NSLI children, on the other hand, although possessing similar working memory skills have no need to draw upon them when solving ToM tasks as their verbal skills are age appropriate and sufficient for reasoning through the given problems.

A further difference between this study and those conducted by Davis and Pratt (1995), Keenan et al. (1998), and Keenan (1998) is the type of working memory measure employed. The present study employed a visual-spatial, non-counting, nonverbal task to measure working memory whereas the other studies employed a visual, verbal counting task to measure working memory. The spatial measure was chosen in light of the severe language problems of the SLI students; an inability to verbalize or count would have unfairly penalized their working memory score - in essence it would have confounded the results. It is unlikely that this differing selection would have caused these differing results given that Keenan (1999) and White (2000) have demonstrated a positive correlation between the two working memory tasks and the ability of both to adequately predict ToM performance in young normally developing children.

Hypothesis 4. The hypothesized relationships between language and behaviour and between ToM ability and behaviour did not receive any clear support in this study. Both age and diagnostic groups had average CBCL scores that clustered closely around the instrument's mean. The limited variance in subject scores did not allow for valid statistical analysis of this variable.

The absence of support for this hypothesis is surprising given the strong body of findings that demonstrate a heightened presence of behaviour problems in language impaired children (Baker & Cantwell, 1987; Beitchman et al., 1996; Cohen et al., 1993, 1998a; Stevenson & Richman, 1978; Stevenson et al., 1985; Tallal et al., 1989). The age range of the subject sample, measures used, and population from which the subjects were drawn all may have contributed to the discrepant finding of the present sample.

For instance, Stevenson's studies found behavioural problems in a random sample of 3-year old children. However, the methodology of these studies does not specifically indicate whether children with mental retardation, hearing loss, early experiences of social deprivation, or neurological delays were screened for exclusion from the category of language delay. Thus, the portion of 3-year-olds identified as language-delayed *with* behaviour problems may include children whose behavioural problems result from any of the listed exclusionary criteria of SLI. In other words, the observed level of behavioural disturbance may include cases where the symptom expression is better accounted for a disorder other than language impairment. The current study adhered to a strict definition of SLI to ensure that the language impairment of the participants was primary in nature and not a secondary consequence of another disorder. Children with significant emotional or behaviour problems as a result of autism, Asperger's, PDD-NOS, or mental retardation were therefore excluded from participation. This tighter screening procedure, designed to eliminate "red herrings", may be responsible for the absence of behavioural problems in the majority of the SLI children. However, while this criterion ruled out inappropriate cases, it did not eliminate all instances of behavioural problems, as approximately 16% of the SLI children achieved CBCL Total scores placing them outside the normal limits of behaviour.

Additionally, Baker and Cantwell (1987) found a heightened expression of behavioural problems in children attending a speech and language clinic. The age range of their sample (2 to 15 years old) was much broader than that of the current study (4 to 7 years old). The difference in ages may be responsible for the absence of behavioural problems in the present study. Baker and Cantwell did not analyse the occurrence of behaviour problems by age group. It is therefore difficult to know if the observed behaviour problems were distributed evenly across the ages or weighted more heavily at any point on the spectrum. If the behavioural problems occurred more frequently in the older subjects (as a response to a life course begun with language difficulties) this would explain why behavioural problems were not observed in the current study - the age range sampled was too restricted and too young.

Furthermore, the ages of the children participating in the investigations by Cohen and colleagues also spanned a broader range. Subjects were 4- to 12-year old psychiatric outpatients (Cohen et al., 1993) and 7-to 14-year old children referred for psychiatric services (Cohen et al., 1998a). The nature and age range of the population being sampled is obviously different than the one sampled in the current study and may account for the difference in results. Subjects in the current study may be too young to be evidencing full clinical disorder. The CBCL scores of a number of the SLI and NSLI subjects placed above the normal range. Perhaps with time these subjects and others might go on to a fuller expression of clinical disorder. Support for an increased rate of behavioural disturbance in language impaired children has been documented by Beitchman et al. (2001). These authors, in sampling the general population, found that normally developing, language impaired, and speech disordered children did not exhibit differing rates of behavioural disturbance until after the age of 8. Moreover, after the age of 12, the disorder rate for the language impaired children rises to a level that also distinguishes it from that of the speech impaired children.

The majority of studies finding increased rates of behavioural problems in language-delayed children employed the CBCL as the measure of behavioural symptomology, however most employed additional measures as well (Beitchman et al., 1996; Cohen et al., 1993, 1998a; Tallal et al., 1989). It is quite likely that the use of more questionnaires completed by more raters (teachers and parents) produced a better estimate of the occurrence of behaviour problems in those samples. Relying on a single measure completed by a single rater to assess behavioural symptomology may be the reason the current study did not find a heightened level of behavioural disorders in SLI children relative to NSLI children.

Perusal of the CBCL protocols revealed a possible response set bias in the parents of the SLI children. These parents often failed to appropriately endorse the item which queried whether or not and to what degree the child in question experienced speech problems. Fifty-four percent indicated that speech problems were "somewhat" or "not true" of their child. If failure to ascribe this behavioural symptom to a child who is currently identified as language-disordered could occur, then it is very possible that failure to acknowledge other behavioural symptoms could come to pass. This interpretation of "faking good" is given tentative support by the fact that more parents of NSLI children than parents of SLI children were willing to endorse sufficient items to tally a score that placed their children's total scores outside the normal range and within the significantly elevated range. Approximately 16% (7 of 43) of the SLI children were perceived by their parents as exhibiting behaviours at a level beyond normal on the Total scale of the CBCL. However, approximately 31% (15 of 49) of the NSLI children were rated as engaging in behaviour that significantly deviated from the norm on the Total scale of the CBCL.

Lack of variation in the behaviour scores generated by parent raters automatically prevented testing of the sub-hypothesis that type of behavioural symptom expressed by children would change with age from externalizing in the younger years to internalizing in the older years. Therefore, no evidence was gathered to support the idea that the nature of behavioural symptom expression changes with age.

Limitations of the Study and Future Research Directions

The present study is the first to demonstrate inferior ToM performance in a sample of SLI children. Replication of this study with its age-appropriate battery of ToM tasks is necessary to lend further credence to these results. Replication is also needed to decrease future reliance on and reference to the contrary findings of Leslie and Frith (1988) and Perner et al. (1989) which most likely resulted from an inappropriate battery of ToM tasks.

The sample size of the current study was relatively small and this may have resulted in some null findings such as the lack of association between number of siblings and performance on ToM tasks. Increasing the number of subjects would allow for a finer analysis of the ToM development of 4-, 5-, 6- and 7-year-olds and negate the necessity of creating "younger" and "older" age groups while still allowing for the construction of such groupings.

Including a brief, nonverbal screen for intelligence would have been useful to rule out any significant differences in general cognitive ability between the two groups of children. In designing the study this measure was not included because the time requirements of the test battery were thought to be approaching a level that would tax the energy and concentration level of children so young as 4 years of age. Conducting two briefer sessions would alleviate this constraint and so could be considered in future investigations.

The nonverbal ToM task used in this study provided a measure of first-order ToM. Development of a series of second-order nonverbal tasks would be useful to extend the current finding of delayed nonverbal ToM performance in SLI children. The first-order nonverbal measure of ToM is not one that is frequently deployed in studies investigating children's ability to "read" others' minds. It would be prudent to determine its psychometric properties to ensure that it is ToM that is truly being measured. Also a study could be designed to determine if this task is readily comprehensible without the accompaniment of verbal instructions. This would lend further credence to the current finding of delayed conceptualization of ToM in SLI children.

Determining the role of absolute level of language in ToM performance was difficult in the current study because of the language measure used. The age-corrected scores necessitated construction of a more informal and approximate measure of absolute language ability (scores on control questions). A language test capable of covering the ages of 4 to 8 and possessing strong psychometric properties would be most useful. This is however a difficult age range for a single test to cover as it incorporates pre-school to primary aged children and many tests are devised based on this distinction.

Administration of a more complete language battery (one that assesses both syntactic and semantic skills) would aid in illustrating which components of language are most integral to the formation and performance of ToM ability. An understanding of this might also be approached from a different angle. Administering ToM tasks to groups of children with different subtypes of language impairment (low expressive, low receptive, low expressive and receptive) might shed further light on the specific language skills that are crucial to success on ToM tasks and understanding the mental states of others. If and when these components are identified, then research can begin to look into whether or not ToM can be taught or improved through language intervention.

The role of language in the expression and conceptualization of ToM might be further explored by administering the nonverbal task to children with different language delay profiles. Children with poor expressive (but adequate receptive) ability might be expected to perform at age level on a nonverbal measure of ToM and any measure of ToM that did not require a verbal response. A different success pattern would be predicted for children with poor receptive (but adequate expressive) ability; poor performance on standard versions of ToM tasks and age appropriate performance on a nonverbal measure of ToM. Should they perform poorly on both verbal and nonverbal ToM tasks, further proof will have been garnered to demonstrate that language plays a role in acquiring a concept of other's mental states.

Along similar lines, it will be important for future studies to try and delineate exactly what elements a child needs to hold in mind in order to succeed on ToM tasks. More precise measures of working memory can then be chosen which reflect the information children need to retain and manipulate. Consider the "Smarties" task as an example. This requires one to keep a very distinct series of representations in mind to successfully complete the task: what it looks like the box contains, what the box has been shown to contain, to whom the representations belong (yourself, the examiner, an uninitiated other), and the order in which these representations occurred. The working memory demands of these elements when understood by Baddeley's (1981) model of working memory can be understood as calling upon both the visual-spatial scratch pad and articulatory loop to hold the information as well as the central executive to process the information. It seems plausible that the ability to recall a previous belief in the face of one's present belief requires verbal working memory. The child must be able to keep in mind through a verbal storage and rehearsal system, his or her previous false belief. Since language is the essential means by which a child can communicate this false belief, and the structural features of language are involved in the way mental states are conceptualized (Astington & Jenkins, 1999), the storage and rehearsal of one's own false belief is likely to primarily involve the verbal pathway. Therefore, it would be interesting to compare ability of different types of working memory tasks (visual-spatial, verbal) to predict performance on different ToM tasks.

Investigation into the relationship between the social skills and ToM ability of SLI children is also warranted given the lack of an association between behavioural deviancy and ToM ability in this study. The social deficits of young SLI children are well documented (Brinton et al., 1997; Craig & Washington, 1993; Hadley & Rice, 1991; Rice, Sell, & Hadley, 1991). It would be interesting to determine if poor social skills are a precursor to the behavioural problems generally found in older SLI children. It is suggested that ratings of social skills and behavioural deviancy be obtained from both parents and teachers in future studies in order to confirm or rule out the response bias suggested within the findings of this study. The absence of an association between ToM ability and behaviour, and between SLI children and behaviour suggests that the behaviour measure used in this study may have been inadequate or inappropriate in some way. Either the children were too young to be exhibiting the type of behavioural symptoms contained in this measure or the symptoms they were exhibiting were too subtle to be reflected in such a clinical scale. Incorporation of additional behaviour measures would have been useful. Specifically, a period of observation of the subjects that could be coded and then rated or a measure of social skill development might have captured less frank behavioural deviancy. It may be that the influence of ToM understanding is quite subtle and so will influence social skills but not influence behavioural functioning to the point of disorder or clinical severity.

Given that SLI children have difficulty processing rapidly presented auditory information (Tallal & Piercy, 1973a,b; Wright, Lombardino, King, Puranik, Leonard & Merzenich, 1997), it would be interesting to see if slowing down the presentation of the verbal ToM scripts would enhance the performance of SLI children. This seems somewhat unlikely considering that the SLI children did equally poorly on the nonverbal measure of ToM which was untimed and conducted more or less at their own pace.

It is also documented that SLI children possess limited processing capacity compared to their peer group (Gillam, Cowan, & Marler, 1998; Katz et al., 1992; Records et al., 1995). Bearing this in mind, devising versions of ToM tasks that vary the amount of information needing to be processed might reveal interesting variations in the performance of SLI children. Some preliminary evidence exists attesting to the influence processing load can have on the ToM performance of autistic (Tager-Flusberg & Sullivan, 1994b) and younger children (Sullivan et al., 1994).

Summary

The present study has clearly demonstrated that children with SLI are delayed in their acquisition of an understanding of others' minds. Their performance on a comprehensive battery of ToM tasks is significantly below that of same-aged normally developing children. Their performance deficit extends into the realm of nonverbal ToM tasks as well and suggests that language has a role to play in both the formation and performance of ToM understanding. The present study also showed that the ToM development of SLI children proceeds in the same sequence but at a diminished pace relative to their normally developing peer group. Support is thus gathered for the argument that their weaker ToM performance reflects a delay in development not a lack of ToM. An elevated rate of behaviour disorder was not found to be present in this sample of children. The relationship between ToM and behaviour was not supported; possibly due to the age of the sample and the measure employed. The influence of working memory on competent ToM performance in SLI children was revealed in this study. This finding is an extension of previous research which has not investigated the role of working memory on ToM performance outside of normally developing children. The importance of both language ability and processing capacity are thus evidenced in the successful development of a child's ability to understand other minds.

## References

Achenbach, T. M., & Edelbrock, C. S. (1991a). Manual for the Child Behaviour Checklist. Burlington, VT: University Of Vermont Press.

Achenbach, T., & Edelbrock, C. (1991b). Manual for the Teacher Report Form. Burlington, VT: University of Vermont Press.

Ackerman, B. P. (1981). Young children's understanding of a speaker's intentional use of a false utterance. Developmental Psychology, 17, 472-480.

Ackerman, B. P. (1986). Children's sensitivity to comprehension failure in interpreting a nonliteral use of an utterance. Child Development, 57, 485-497.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Anderson, W. B. (1998). An investigation of the relationship between literacy, verbal ability, and metarepresentational ability in normal young readers. Dissertation Abstracts International: Section B: The Sciences & Engineering, 58(11-B), 6251.

Astington, J. W., & Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. British Journal of Developmental Psychology, 9, 7-31.

Astington, J. W., Harris, P. L., & Olson, D. R. (Eds.). (1988). Developing theories of mind. New York: Cambridge University Press.

Astington, J.W., & Jenkins, J. M. (1995). Theory of mind development and social understanding. Cognition and Emotion, 9 (2/3), 151-165.

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

Azmitia, M., & Hesser, J. (1993). Why siblings are more important agents of cognitive development: A comparison of siblings and peers. Child Development, 64, 430-444.

Bachara, G. H., Raphael, J., & Phelan, W. J. III. (1980). Empathy development in deaf adolescents. American Annals of the Deaf, 125, 38-41.

Baddeley, A. (1981). The concept of working memory: A view of its current state and probable future development. Cognition, 10, 17-23.

Baker, L., & Cantwell, D. P. (1987). A prospective follow-up of children with speech/language disorders. Journal of the American Academy of Child and Adolescent Psychiatry, 26(4), 546-553.

Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. Journal of Child Psychology & Psychiatry, 30(2), 285-297.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioural and intentional understanding of picture stories in autistic children., British Journal of Developmental Psychology, 4(2), 113-125.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. (Eds.). (1993). Understanding other minds: Perspectives from autism. New York: Oxford University Press.

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., Inglis, A., & Lancee, W. (1996). Long-term consistency in speech/language profiles: II. Behavioral, emotional, and social outcomes. Journal of the American Academy of Child and Adolescent Psychiatry, 35(6), 815-825.

Beitchman, J. H., Wilson, B., Johnson, C. J., Atkinson, L., Young, A., Adlaf, E., Escobar, M., & Douglas, L. (2001). Fourteen-year follow up of speech/language-impaired and control children: Psychiatric outcome. Journal of the American Academy of Child and Adolescent Psychiatry, 40(1), 75-82.

Bishop, D. V. M., & Adams, C. (1992). Comprehension problems in children with specific language impairment: Literal and inferential meaning. Journal of Speech and Hearing Research, 35, 119-129.

Brimer, M. A., & Dunn, L. M. (1962). Manual for the English Picture Vocabulary Test. Educational Evaluation Enterprises, Bristol.

Brinton, B., Fujiki, M., & McKee, L. (1998). Negotiation skills of children with specific language impairment. Journal of Speech, Language, and Hearing Research, 41, 927-940.

Brinton, B., Fujiki, M., & Powell, J. M. (1997). The ability of children with language impairment to manipulate topic in a structured task. Language, Speech, and Hearing Services in Schools, 28, 3-11.

Brinton, F., Fujiki, M., Spencer, J., & Robinson, L. A. (1997). The ability of children with specific language impairment to access and participate in an ongoing interaction. Journal of Speech and Hearing Research, 40, 1011-1025.

Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. Child Development, 67, 836-849.

Burroughs, E. I., & Tomblin, J. B. (1990). Speech and language correlates of adults' judgments of children. Journal of Speech and Hearing Disorders, 55, 485-494.

Case, R., Kurland, M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory. Journal of Experimental Psychology, 33, 386-404.

Cohen, N. J., Barwick, M. A., Horodezky, N. A., Vallance, D. D., & Im, N. (1998). Language, achievement, and cognitive processing in psychiatrically disturbed children with previously identified and unsuspected language impairments. Journal of Child Psychology and Psychiatry and Allied Disciplines, 39(6), 865-877.

Cohen, N. J., Davine, M., Horodezky, N., Lipsett, L., & Isaacson, L. (1993). Unsuspected language impairment in psychiatrically disturbed children: Prevalence and language and behavioural characteristics. Journal of the American Academy of Child and Adolescent Psychiatry, 32(3), 595-603.

Cohen, N. J., Menna, R., Vallance, D. D., Barwick, M. A., Im, N., & Horodezky, N. B. (1998). Language, social cognitive processing, and behavioural characteristics of psychiatrically disturbed children with previously identified and unsuspected language impairments. Journal of Child Psychology and Psychiatry and Allied Disciplines, 39(6), 853-864.

Cole, K., & Mitchell, P. (2000). Siblings in the development of executive control and a theory of mind. British Journal of Developmental Psychology, 18(2), 279-295.

Craig, H. K. (1993). Social skills of children with specific language impairment: Peer relationships. CLINICAL FORUM: Language and social skills in the school-age population. Language, Speech, and Hearing Services in Schools, 24, 206-215.

Craig, H. K., & Gallagher, T. M. (1986). Interactive play: The frequency of related verbal responses. Journal of Speech and Hearing Research, 29, 375-383.

Craig, H. K., & Washington, J. A. (1993). Access behaviours of children with specific language impairment. Journal of Speech and Hearing Research, 36, 322-337.

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. Child Development, 70(4), 853-865.

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

Demorest, A., Meyer, C., Phelps, E., Gardner, H., & Winner, E. (1984). Words speak louder than actions: Understanding deliberately false remarks. Child Development, 55, 1527-1534.

Dunn, J. (1995). Children as psychologists: The later correlates of individual differences in understanding of emotions and other minds. Cognition and Emotion, 9(2/3), 187-201.

Dunn, J., & Brown, J. (1994). Affect expression in the family, children's understanding of emotion, and their interactions with others. Merrill-Palmer Quarterly, 40, 120-137.

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. Child Development, 62, 1352-1366.

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

Dunn, L. M., Dunn, L. M., Whetton, C., & Pintillie, D. (1982). British Picture Vocabulary Scale. London: NFER-Nelson.

Eisenmajer, R., & Prior, M. (1991). Cognitive linguistic correlates of "theory of mind" ability in autistic children. British Journal of Developmental Psychology, 9(2), 351-364.

Flavell, J. H., & Miller, P. H. (1998). Social cognition. In D. Kuhn & R. Siegler (Eds.) & W. Damon (Series Ed.), Handbook of child psychology, fifth edition. Cognition, perception and language (pp. 851-898). New York: John Wiley & Sons, Inc.

Folstein, S., & Rutter, M. (1978). A twin study of individuals with infantile autism. In M. Rutter & E. Schoper (Eds.), Autism: A reappraisal of concepts and treatment. New York: Plenum.

Frith, U., Happé, F., & Siddons, F. (1994). Autism and theory of mind in everyday life. Social Development, 3(2), 108-124.

Frith, U., Morton, J., & Leslie, A. M. (1991). The cognitive basis of a biological disorder: Autism. Trends in Neuroscience, 14, 433-438.

Fujiki, M., Brinton, B., & Todd, C. M. (1996). Social skills of children with specific language impairment. Language, Speech, and Hearing Services in Schools, 27, 195-202.

Garrison, W. M., Emerton, R. G., & Layne, C. A. (1978). Self-concept and social interaction in a deaf population. Paper Series, Department of Research and Development, National Technical Institute for the Deaf.

Gillam, R. B., Cowan, N., & Marler, J. A. (1998). Information processing by school-age children with specific language impairment: Evidence from a modality effect paradigm. Journal of Speech and Hearing Research, 41, 913-926.

Hadley, P. A., & Rice, M. L. (1991). Conversational responsiveness in speech and language-impaired children. Journal of Speech and Hearing Research, 34, 1308-1317.

Hadley, P. A., & Rice, M. L. (1991). Conversational responsiveness of speech- and language-impaired preschoolers. Journal of Speech and Hearing Research, 34, 1308-1317.

Halford, G. S., Maybery, M. T., & Bain, J. D. (1986). Capacity limitations in children's reasoning: A dual-task approach. Child Development, 57, 616-627.

Happé, F. G. E., (1993). Communicative competence and theory of mind in autism: A test of relevance theory. Cognition, 48, 101-119.

Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. Journal of Autism and Developmental Disorders, 24(2), 129-154.

Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. Child Development, 66, 843-855.

Happé, F., & Frith, U. (1996). Theory of mind and social impairment in children with conduct disorder. British Journal of Developmental Psychology, 14(4), 385-398.

Haynes, C., & Naidoo, S. (1991). Children with specific speech and language impairment. Clinics in Developmental Medicine, Volume 119. Oxford: Blackwell Scientific Publications Ltd.

Hresko, W. P., Reid, D. K., & Hammill, D. D. (1981). The Test of Early Language Development (TELD). Austin, TX: Pro-Ed.

Jenkins, J. M., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. Developmental Psychology, 32(1), 70-78.

Katz, W. F., Curtiss, S., & Tallal, P. (1992). Rapid automatized naming and gesture by normal and language-impaired children. Brain and Language, 43, 623-641.

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

Keenan, T. (1999, April). The role of working memory, inhibitory control and speed of processing in children's understanding of false belief. Paper/Poster presented at the biennial meeting of the Society for Research in Child Development (SRCD), Albuquerque, New Mexico.

Keenan, T. (2000). Working memory, 'holding in mind', and the child's acquisition of a theory of mind. In J. W. Astington (Ed.), Mind in the making (pp. 233-249). Oxford: Blackwell Publishers.

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

Keenan, T., & Quigley, K. (1999). The role of echoic mention in young children's understanding of sarcasm. British Journal of Developmental Psychology, 17, 83-96.

Konstantareas, M. M., & Beitchman, J. H., (1996). Comorbidity of autistic disorder and specific developmental language disorder: Existing evidence and some promising future directions. In J. H., Beitchman, N. J., Cohen, M. M., Konstantareas, & R. Tannock (Eds.), Language, learning and behavior disorders (pp. 178-196). New York: Cambridge University Press.

Kusche, C. A., Garfield, T. S., Greenberg, M. T. (1983). The understanding of emotional and social attributions in deaf adolescents. Journal of Clinical Child Psychology, 12 (2), 153-160.

Lalonde, C. E., & Chandler, M. J. (1995). False belief understanding goes to school: On the social-emotional consequences of coming early or late to a first theory of mind. Cognition and Emotion, 9 (2/3), 167-185.

Le Couter, A., Bailey, A., Goode, S., Pickles, A., Robertson, S., Gottesman, I., & Rutter, M. (1996). Broader phenotype of autism: The clinical spectrum in twins. Journal of Child Psychology and Psychiatry, 37(7), 785-801.

Leonard, L. B. (1998). Children with specific language impairment. Cambridge, Massachusetts: MIT Press.

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

Leslie, A. M., & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. British Journal of Developmental Psychology, 6(4), 315-324.

Luria, A. R., (1976). Cognitive Development. Cambridge, Mass: Harvard University Press.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. Psychological Bulletin, 95, 109-133.

Ozonoff, S., Pennington, B. F., & Rogers, S. J., (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. Journal of Child Psychology and Psychiatry and Allied Disciplines, 32(7), 1081-1105.

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: Knowledge, belief, and communication. Child Development, 60, 689-700.

Perner, J., Leekam, S., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief. The case for a conceptual deficit. British Journal of Developmental Psychology, 5, 125-137.

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

Perner, J., & Wimmer, H. (1985). "John *thinks* that Mary *thinks* that . . ." Attribution of second-order beliefs by 5 - 10 year old children. Journal of Experimental Child Psychology, 39, 437-471.

Peterson, C. C., & Siegal, M. (1995). Deafness, conversation and theory of mind. Journal of Child Psychology and Psychiatry and Allied Disciplines, 36(3), 459-474.

Peterson, C. C., & Siegal, M. (1997). Domain specificity and everyday biological, physical, and psychological thinking in normal, autistic and deaf children. In H. M., Wellman, & K., Inagaki (Eds.), The emergence ot core domains of thought: Children's reasoning about physical, psychological , and biological phenomena. New directions for child development, No. 75 (pp.55-70). San Francisco, CA: Jossey-Bass Inc. Publishers.

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

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. American Journal of Psychiatry, 154, 185-190.

Prior, M., Eisenmajer, R., Leekam, S., Wing, L., Gould, J., Ong, B., & Dowe, D. (1998). Are there subgroups within the autistic spectrum? A cluster analysis of a group of children with autistic spectrum disorders. Journal of Child Psychology and Psychiatry and Allied Disciplines, 39(6), 893-902.

Records, N. L., Tomblin, B., & Buckwalter, P. R. (1995). Auditory verbal learning and memory in young adults with specific language impairment. The Clinical Neuropsychologist, 9 (2), 187-193.

Rice, M. L., Hadley, P. A., & Alexander, A. L. (1993). Social biases toward children with speech and language impairments: A correlative causal model of language limitations. Applied Psycholinguistics, 14, 445-471.

Rice, M. L., Sell, M. A., & Hadley, P. A. (1991). Social interactions of speech- and language-impaired children. Journal of Speech and Hearing Research, 34, 1299-1307.

Richman, N., Stevenson, J., & Graham, P. (1982). Preschool to school. London: Academic Press.

Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. C. (1998). Older (but not younger) siblings facilitate false belief understanding. Developmental Psychology, 34(1), 161-174.

Ruffman, T., Slade, L., Clements, W., & Import, A. (1999). How language is related to verbal and nonverbal theory of mind. Unpublished manuscript, University of Sussex, East Sussex.

Russell, P. A., Hosie, J. A., Gray, C. D., Scott, C., & Hunter, N. (1998). The development of theory of mind in deaf children. Journal of Child Psychology and Psychiatry and Allied Disciplines, 39(6), 903-910.

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: Preliminary findings. Journal of Child Psychology and Psychiatry, 8, 1-11.

Rutter, M., Bailey, A., Bolton, P., & Le Couteur, A. (1993). Autism: Syndrome definition and possible genetic mechanisms. In R. Plomin & G. E. McClearn (Eds.), Nature, nurture & psychology. Washington DC: American Psychological Association.

Rutter, M., & Mawhood, L (1991). The long-term psychosocial sequelae of specific developmental disorders of speech and language. In M. Rutter & P. Casaer (Eds.), Biological risk factors for psychosical disorders (pp 233-259). Cambridge: Cambridge University Press.

Sattler, J. M. (1988). Assessment of children, third edition. San Diego: Jerome M. Sattler, Publisher.

Semel, E., Wiig, E. H., Secord, W. A. (1995). Clinical Evaluation of Language Fundamentals - Third Edition (CELF-3). San Antonio, Texas: The Psychological Corporation/ Harcourt Brace & Co.

Sodain, B., & Frith, U. (1992). Deception and sabotage in autistic, retarded and normal children. Journal of Child Psychology and Psychiatry, 33, 591-605.

Sparrow, S. S., Balla, D. A., Cicchetti, D. V. (1984). Vineland Adaptive Behaviour Scales. MN: American Guidance Service.

Stark, R., & Tallal, P. (1981). Selection of children with specific language deficits. Journal of Speech and Hearing Disorders, 46, 114-122.

Stevens, L., J., & Bliss, J. S. (1995). Conflict resolution abilities of children with specific language impairment and children with normal language. Journal of Speech and Hearing Research, 38, 599-611.

Stevenson, J. (1996). Developmental changes in the mechanisms linking language disabilities and behaviour disorders. In J. H., Beitchman, N. J. Cohen, M. M., Konstantareas, &

R. Tannock (Eds.), Language, learning and behavior disorders (pp. 78-99). New York: Cambridge University Press.

Stevenson, J., Richman, N., & Graham, P. (1985). Behaviour problems and language abilities at three years and behavioural deviance at eight years. Journal of Child Psychology and Psychiatry, 26(2), 215-230.

Sullivan, K., Winner, E., & Hopfield, N. (1995). How children tell a lie from a joke: The role of second-order mental state attributions. British Journal of Developmental Psychology, 13, 191-204.

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

Tager-Flusberg, H., & Sullivan, K. (1994a). Predicting and explaining behaviour: A comparison of autistic, mentally retarded and normal children. Journal of Child Psychology and Psychiatry and Allied Disciplines, 35(6), 1059-1075.

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

Tallal, P., Dukette, D., & Curtiss, S. (1989). Behavioural/emotional profiles of pre-school language-impaired children. Development and Psychopathology, 1(1), 51-67.

Tallal, P., & Piercy, M. (1973a). Developmental aphasia: Impaired rate of nonverbal processing as a function of sensory modality. Neuropsychologia, 11, 389-398.

Tallal, P., & Piercy, M. (1973b). Defects of nonverbal auditory perception in children with developmental aphasia. Nature, 241, 468-469.

Tanguay, P. E., Robertson, J., Derrick, A. (1998). A dimensional classification of autism spectrum disorder by social communication domains. Journal of the American Academy of Child and Adolescent Psychiatry, 37(3), 271-277.

Taylor, M. (1996). A theory of mind perspective on social cognitive development. In R. Gelman & T. Au (Eds.) & E. C. Carterette & M. P. Friedman (Gen. Eds.), Handbook of perception and cognition: Vol. 13, Perception and cognition development. New York: Academic Press.

Thorndike, R. L., Hagen, E. P., Sattler, J. M. (1986). The Stanford-Binet Intelligence Scale: Fourth Edition. Chicago, IL: The Riverside Publishing Company.

Upton, G. J. G. (1978). The analysis of cross-tabulated data. Chichester, NY: Wiley.

Wechsler, D. (1989). Wechsler Preschool and Primary Scale of Intelligence - Revised. New York: The Psychological Corporation.

Wechsler, D. (1991). Wechsler Intelligence Scale of Children - Third Edition. New York: The Psychological Corporation.

Wellen, C., & Broen, P. (1982). The interruption of young children's responses by older siblings. Journal of Speech and Hearing Disorders, 47, 204-210.

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

Wellman, H. M., & Inagaki, K. (Eds.). (1997). The emergence of core domains of thought: Children's reasoning about physical, psychological, and biological phenomena. San Francisco, CA: Jossey-Bass Publishers.

White, S. (2000). The link between preschoolers' theory of mind and executive functioning revisited. Unpublished master's thesis, University of Canterbury, Christchurch, New Zealand.

Wiig, E. H., Secord, W., & Semel, E., (1992). Clinical Evaluation of Language Fundamentals - Preschool (CELF-Preschool). San Antonio, Texas: The Psychological Corporation/Harcourt Brace Jovanovich, Inc.

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

Winner, E., & Leekam, S. (1991). Distinguishing irony from deception: Understanding the speaker's second-order intention. British Journal of Developmental Psychology, 9, 257-270.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. Nature, 387, 176-178.

Yirmiya, N., Osnat, E., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. Psychological Bulletin, 124(3), 283-307.

## Appendix A

### Cover Letter to Parents

Dear Parent,

I am conducting a research project through Lakehead University looking at the development of children's understanding of social interactions. Specifically, I am interested in their understanding that peoples' behaviours are affected by their beliefs and their goals, in other words, by what goes on inside their heads. The main purpose of the present study is to better understand what factors in the early years of life lead to theory of mind development. In order to do this, it is necessary to look at how different groups of children perform on a variety of tasks. This study will focus on children with language difficulties, however, it will also study children without any known language or learning difficulties.

I would like to include you and your child in this study. The project has been reviewed by the head of your child's school/daycare/school/district health unit/integrated services program and we believe it poses no threat to children's welfare. The project has also been reviewed and approved by Lakehead University's Ethics Advisory Committee. In order to include you and your child, I need your written consent, which can be given by signing and dating page 3 of this letter (entitled "CONSENT TO PARTICIPATE") and returning it to your child's teacher/speech and language pathologist/Case Manager. Please keep this page for future reference.

Children will be seen individually by a female experimenter who has a great deal of experience working with children. The study will take place in a quiet place within your child's school/daycare/health unit/integrated services office. The children will complete 3 simple sets of tasks. The first task is a hiding game with dolls; the children are asked questions about where another doll will look for an object. Children's answers to these questions give us a better understanding as to what they know about other people's thoughts. The second task involves reading to the child and asking questions, and is designed to measure children's developing language skills. Finally, children will play a memory game using cards and coloured dots. This will help determine the role of memory in children's understanding of the mind.

In total, this testing should take between 30 and 45 minutes. Please note that at no time will your child be given negative feedback. All answers will be greeted with enthusiasm from the experimenter. Other experimenters have found that this approach results in children enjoying the "games" and benefiting from their interaction with a stranger in a positive way - it leaves them feeling happy and confident. I would also like to point out that your consent in no way means that your child will be forced to participate. If your child is reluctant to become involved their decision will be respected, and I will always make sure that the children know they can withdraw from the games at any time.

Aside from your child's participation in these tasks, I am also requesting that parents complete a behaviour checklist. This will help to discern any links between behavioural traits and performance on the above tasks. All information gathered during the course of this study will be kept confidential. No names or individual identifications will be used in publications that may arise as a result of this research. It is standard research policy that all raw data be retained for 7 years following an investigative study. Please be assured that the data will be under lock and key and that it will not contain any identifying information.

If you would agree to allow me to test your child, please sign the consent form on page 3 (entitled, "CONSENT TO PARTICIPATE") and return it to your child's teacher/speech and language pathologist/ISNC Case Manager. I also require the number of persons residing within your home, the birth date of your child, and would like to know the birth date of any of their siblings (this will allow me to see if the number of sibling makes any difference in children's understanding of other minds). This information can be provided in the appropriate spaces on the following page.

Thank you very much for your help with this study. If you have any questions about the study, please do not hesitate to contact me. Also, a brief write-up outlining the results of this study will be available upon its completion. This write-up will be available to you through your child's school/daycare/district health unit/integrated services unit or by contacting me at the address or phone number given below. Once again, thank you very much for your help with this study and for helping to extend the body of knowledge surrounding children's understanding of social interactions.

Sincerely,

Alana Holmes, M.A., Ph. D. Candidate

Psychology Department

Lakehead University

Thunder Bay, ON

P7B 5E1

(phone 343-8441)

## Appendix B

### CONSENT TO PARTICIPATE [Theory of Mind Study]

I have read and understood the description of the above-named project. On this basis, I agree to allow my child to participate as a subject in the project, and I consent to publication of the results of the project with the understanding that anonymity will be preserved. I understand also that at any time I may withdraw my child from the project, including the withdrawal of any information provided.

I agree to allow my child, ______________________________________ to participate in the study described on pages 1 & 2. I understand that the results will be confidential

Signed: ______________________________ Date: ______________________

Please print the following information:

- Child's Full Name: ________________________
- Child's Birthdate: __________________________
- Does this child have any siblings? ______________ If so, how many? ______________
- How many people in total live with you and your child?(e.g., grandparents, partner, spouse, step children) __________
- Of these people, how many are older than your child? ___________
- Please list below the names, and birth dates of any of your child's siblings:

    _______________________________________________________

    _______________________________________________________

    _______________________________________________________

    _______________________________________________________

## Appendix C

## Working Memory Spatial Task Script

Practice Trials:

"We are going to look at some houses. Each house will have a red dot in one of its rooms. The dot can be in *any* room in the house. When I show you a house, I want you to point to the red dot and then try to remember where it is. After a few seconds, I will show you an empty house and ask you to point to where the dot was."

Level 1:

"I am going to show you some more cards with pictures of houses on them. Each house will have a red dot in one of its rooms. When I show you a card, I want you to point to the red dot and try to remember where it is. I will then show you an empty house and ask you to point to where the red dot was."

Level 2:

"Now we are going to make the game a little harder. I am going to show you two cards in a row. I will show you one card and you are to point to the red dot and try to remember where it is. Then, I will show you another card and you are to point to the red dot and try to remember where it is. Then, I am going to show you two empty houses and ask you to show me where the red dot was in the first house, and where the red dot was in the second house."

Level 3:

"Now we are going to make the game a little harder again. I am going to show you three cards in a row. I will show you one card and you are to point to the red dot and try to remember where it is. Then, I will show you another card and you are to point to the red dot and try to remember where it is. Finally, I will show you one more card and you are to point to the red dot and try to remember where it is in the house. Then, I am going to show you three empty houses and ask you to show me where the red dots were in each house."

## Appendix D

## First-Order Unexpected Change of Location Task Script

1st Order False Belief Task 1 [ Ball Story]

"This is Calvin. He is tired of playing with his ball so he puts it away in the box and goes outside to the park. While Calvin is at the park, his Mum takes his ball and plays with it for a few minutes. When she is finished playing with the ball she puts it away in the basket."

*Memory Question:* Where did Calvin put the ball in the beginning?

*Reality Question:* Where is the ball now?

*Knowledge Question:* Did Calvin see his Mum move the ball?

"Now Calvin comes back from the park he wants to play with his ball again."

*Belief Question:* Where will Calvin look for his ball?

1st Order False Belief Task 2 [ Ring Story ]

"Today is Mum's birthday. For her birthday she got a very beautiful ring. Mum puts her ring in her orange jewelry box to keep it safe while she is out shopping. While Mum is out shopping, Dad comes to have a look at Mum's beautiful ring. By mistake, he puts it back into Mum's blue jewelry box."

*Memory Question:* Where did Mum put the ring in the beginning?

*Reality Question:* Where is the ring now?

*Knowledge Question:* Did Mum see Dad move the ring?

"Now Mom comes home and wants to put her ring back on."

*Belief Question:* Where will Mum look for her ring?

<u>1st Order False Belief Task 3 [ Pencil Story ]</u>

"This is Dad. He is done writing with his pencil, so he puts it away in the blue box and goes outside. While Dad is outside, Calvin comes and uses the pencil to draw a picture. Then Calvin puts the pencil away in the white desk drawer."

*Memory Question:* Where did Dad put the pencil in the beginning?

*Reality Question:* Where is the pencil now?

*Knowledge Question:* Did Dad see Calvin move the pencil?

"Now Dad comes back inside and want to start writing again."

*Belief Question:* Where will Dad look for his pencil?

## Appendix E

## First-Order Unexpected Contents Task Script

<u>1st Order False Belief Task 4 [Smarties]</u>

Show box to child with closed lid.

*Control Question 1:* "What do you think is in here?"

Open box and show contents. Close box.

*Memory/Own Belief Question:* "What did you think was inside this box *before* I opened it?"

*Forced Choice Alternative:* "Did you think there was nothing inside it, or did you think there were Smarties inside it?"

*Memory/Reality Question (Control Question 2):* "What is really inside the box?"

*Other's Belief Question:* "(Name of friend) has never looked inside this box. What will (name of friend) *think* is inside this closed box before we take the top off?"

*Forced Choice Alternative:* "(Name of friend) has never looked inside this box. Will (name of friend) think there are Smarties or pencils inside the box?"

## Appendix F

## First- and Second-Order Sarcasm Task Script

Sarcasm Story Text #1: [Football Story]

One day, Charlie Brown was walking through the park. It was a nice sunny day. As he walked along, Charlie Brown saw Lucy sitting down and holding a ball. Lucy called to Charlie Brown and said, "Hey, Charlie Brown! Come and kick the ball." Charlie Brown said, "OK! That will be fun." So Charlie Brown ran very fast to kick the ball. But, just as he tried to kick the ball, Lucy pulled the ball away and Charlie Brown fell down and he landed on his head. When Charlie Brown got up, he looked at Lucy and said, *"Boy. That was the most fun I've had all day."*

*Fact Question 1:* What did Charlie Brown want to do?

*Fact Question 2:* Did Charlie Brown kick the ball?

*First-order (attitude) Question:* When Charlie Brown said, "Boy, that was the most fun I've had all day", what do you think he meant?

*Second-order (intention) Question:* When Charlie Brown said, "Boy that was the most fun I've had all day," what did he want Lucy to think?

Sarcasm Story Text #2: [Dog Story]

One day, Charlie Brown and Sally were trying to teach their dog Snoopy to do a trick. Sally asked, "Can Snoopy fetch a ball?" Charlie Brown said, "Snoopy's a really smart dog and can do any trick." So, Charlie Brown got a ball, showed it to Snoopy, and then threw it for Snoopy to chase. Snoopy just stood there looking at the ball. He didn't chase it at all. When Sally saw this she looked at Charlie Brown and said, *"He sure is a smart dog."*

*Fact Question 1:* What did Sally want Snoopy to do?

*Fact Question 2:* Did Snoopy chase after the ball?

*First-order (attitude) Question:* When Sally said, "He sure is a smart dog", what do you think she meant?

*Second-order (intention) Question:* When Sally said, "He sure is a smart dog", what did she want Charlie Brown to think?

<u>Sarcasm Story Text #3: [Cookies Story]</u>

One day Lucy decided to bake some cookies for her brother Linus. She said, "Linus, I'm going to make chocolate chip cookies for you." Linus said, "Oh great, I'm really hungry! <u>Chocolate chip cookies are the best</u>." So Lucy got busy and made the cookies. As she was baking, the phone rang and Lucy answered it. When she came back, the cookies were all burned. She took them into Linus and said, "Here are you cookies!" Linus looked at the cookies and said, *"Yum. These cookies are the best."*

*Fact Question 1:* Who made the cookies?

*Fact Question 2:* Did the cookies turn out all right?

*First-order (attitude) Question:* When Linus said, "Yum. These cookies are the best", what do you think he meant?

*Second-order (intention) Question:* When Linus said, "Yum. These cookies are the best", what did he want Lucy to think?

# Appendix G

## Second-Order Unexpected Change of Location and Ignorance Task Script

$2^{nd}$ Order False Belief & Ignorance Task 1 [puppy]

This is a story about a boy named Peter and his Mum.

This is Peter, this is his Mum, this is the kitchen, and this is the garage in their house.

Tonight is Peter's birthday and Mum is surprising him with a puppy. She has hidden the puppy in the garage. Peter days, "Mom I really hope you get me a puppy for my birthday." Remember, Mum wants to surprise Peter with the puppy. So, instead of telling Peter she got him a puppy, Mum says, "Sorry Peter, I did not get you a puppy for your birthday. I got you a really great toy instead."

*Reality Control Question:* What did Mum really get Peter for his birthday?

Now, Peter says to Mum, "I'm going outside to play." On his way outside Peter goes to the garage to fetch his ball. In the garage, Peter finds the birthday puppy! Peter says to himself, "Wow, Mum didn't get me a toy, she really got me a puppy for my birthday." Mum does NOT see Peter go to the garage and find his birthday puppy.

*1st order Ignorance Control Question:* Does Peter know that his Mum got him a puppy for his birthday?

*Linguistic Control Question:* Does Mum know that Peter saw the birthday puppy in the garage?

Now, the telephone rings, brring-brring! Peter' grandma calls to find out what time the birthday party is. Grandma asks Mum on the phone, "Does Peter know what you really got him for his birthday?"

*$2^{nd}$ order Ignorance Question:* What does Mum say to Grandma?

Now, remember, Mum does not know that Peter went into the garage. Then Grandma says to Mum, "What does Peter think you got him for his birthday?"

*2nd order false belief:* What does Mum say to Grandma?

<u>2nd Order Ignorance & False Belief Task 2: [chocolate bar]</u>

This is a story about a boy named Ben and a girl name Sarah.

This is Ben, this is Sarah, this is Mum, and this is the kitchen in their house. Now, their mum leaves them a big chocolate bar to share. Sarah eats some of the chocolate bar and leaves the rest on the kitchen table. Then, Sarah leaves the kitchen. Ben wants to make sure that he gets some chocolate too, so he hides the rest in the kitchen cupboard. Now Sarah comes back into the kitchen. Sarah says, "Hey Ben, where's the rest of the chocolate bar?" Remember, Ben wants to have some chocolate for himself. So, instead of telling Sarah the chocolate bar is in the cupboard, he says to Sarah, "I put the rest of the chocolate in my room." Sarah says, "OK, maybe I'll have some later." And the she leaves.

*Reality Control Question:* Where did Peter really put the chocolate?

Now, Be wants to eat the chocolate so he opens the kitchen cupboard. Just then, Sarah walks by the kitchen and sees Ben getting the chocolate from the kitchen cupboard. Sarah says to herself, "Oh, Ben did not put the chocolate in his bedroom, he really hid it in the kitchen cupboard." Ben does not see Sarah peeking in the kitchen door.

*1st Order Ignorance Control Question:* Does Sarah know that Ben hid the chocolate in the cupboard?

*Linguistic Control Question:* Does Ben know that Sarah saw him getting the chocolate from the kitchen cupboard?

Now Ben puts the chocolate back into the cupboard. Then Mum comes into the kitchen. She says to Ben, "Does Sarah know where you really put the chocolate?"

*2nd Order Ignorance:* What does Ben say to his Mum?

Now remember, Ben did not see Sarah peeking through the kitchen door. Then, Mum says, "Where does Sarah think you put the chocolate?"

*2nd Order False Belief:* What does Ben say to Mum?

Table 2

Pearson-Product Moment Correlation Among Variables

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Age | 1.00 | -.03 | .16 | .05 | .06 | .45*** | .00 | .06 | .47*** | .49*** | -.11 | .54*** | .17+ | -.05 | .57*** | -.09 |
| 2. CBCL Ext. | | 1.00 | .51*** | .76*** | .06 | -.00 | .09 | -.11 | -.16 | .10 | -.23* | -.03 | .12 | -.26** | .13 | .08 |
| 3. CBCL Int. | | | 1.00 | .86*** | .05 | .10 | .09 | -.12 | -.13 | .14 | -.23* | .14 | -.01 | -.12 | .20* | -.02 |
| 4. CBCL Tot. | | | | 1.00 | .00 | -.04 | .05 | -.06 | -.18+ | .07 | -.29** | .09 | .07 | -.25** | .14 | -.05 |
| 5. CELF | | | | | 1.00 | .49*** | .15 | -.65*** | .22* | .29** | -.17+ | .43*** | .11 | -.29*** | .51*** | .64*** |
| 6. $1^{st}$ ToM | | | | | | 1.00 | -.06 | -.41*** | .29** | .40*** | -.02 | .46*** | .08 | -.06 | .68*** | .34*** |
| 7. Gender | | | | | | | 1.00 | -.15 | -.10 | .12 | .00 | .01 | .04 | -.01 | .04 | .04 |
| 8. Group | | | | | | | | 1.00 | -.02 | -.34*** | .13 | -.33*** | -.05 | .27** | -.39*** | -.60*** |
| 9. WM | | | | | | | | | 1.00 | .35*** | -.08 | .42*** | .23* | .03 | .34*** | .16 |
| 10. NV ToM | | | | | | | | | | 1.00 | -.25* | .47*** | .19+ | -.19+ | .51*** | .24* |
| 11. Older Sib's | | | | | | | | | | | 1.00 | -.11 | -.14 | .73*** | -.19+ | -.16 |
| 12. $1^{st}$ Sarc. | | | | | | | | | | | | 1.00 | .27** | -.16 | .63*** | .29** |
| 13. $2^{nd}$ Sarc. | | | | | | | | | | | | | 1.00 | -.15 | .19+ | -.08 |
| 14. Sib's | | | | | | | | | | | | | | 1.00 | -.17 | -.21* |
| 15. $2^{nd}$ ToM | | | | | | | | | | | | | | | 1.00 | .33*** |
| 16. WISC | | | | | | | | | | | | | | | | 1.00 |

Note. $1^{st}$ ToM = first-order ToM; WM= working memory; NV ToM = nonverbal ToM; $1^{st}$ Sarc. = first-order sarcasm; $2^{nd}$ Sarc. = second-order sarcasm; $2^{nd}$ ToM = second-order ToM

+$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$. (2-tailed significance)

Table 3

Means and Standard Deviations for Age-Corrected Scale Scores by Group and Age on Language Measures

| Age Group | N | Diagnostic Group | CELF[a] | WISC[a] |
|---|---|---|---|---|
| Younger | 23 | SLI | 5.70 (2.69) | 6.91 (2.31) |
| | 24 | NSLI | 10.13 (3.08) | 11.00 (2.19) |
| Older | 20 | SLI | 5.75 (2.34) | 6.95 (2.48) |
| | 25 | NSLI | 10.28 (2.54) | 10.48 (3.10) |
| All Ages | 43 | SLI | 5.72 (2.50) | 6.93 (2.36) |
| | 49 | NSLI | 10.20 (2.79) | 10.73 (2.68) |
| All Younger | 47 | SLI + NSLI | 7.96 (3.64) | 9.00 (3.04) |
| All Older | 45 | SLI + NSLI | 8.27 (3.32) | 8.91 (3.32) |

[a] Subtest Mean and Standard Deviation = 10.00 (3.00)

Table 4

Means and Standard Deviations for Correct Responses by Group and Age on First-Order ToM

| Age Group | N | Diagnostic Group | 1st-order ToM[a] | Control[b] | Nonverbal ToM[c] |
|---|---|---|---|---|---|
| Younger | 23 | SLI | 2.09 (1.68) | 9.57 (2.00) | 1.96 (1.85) |
| | 24 | NSLI | 3.80 (1.50) | 10.88 (0.45) | 3.33 (2.01) |
| Older | 20 | SLI | 4.00 (1.52) | 10.90 (0.31) | 4.00 (1.81) |
| | 25 | NSLI | 4.92 (0.28) | 11.00 (0.00) | 5.28 (1.10) |
| All Ages | 43 | SLI | 2.98 (1.86) | 10.19 (1.61) | 2.91 (2.08) |
| | 49 | NSLI | 4.37 (1.20) | 10.94 (0.32) | 4.33 (1.88) |
| All Younger | 47 | SLI + NSLI | 2.96 (1.80) | 10.23 (1.56) | 2.66 (2.04) |
| All Older | 45 | SLI + NSLI | 4.51 (1.12) | 10.96 (0.21) | 4.71 (1.58) |

[a] Maximum score = 5.00. [b] Maximum score = 11.00. [c] Maximum score = 6.00.

Table 5

Means and Standard Deviations for Correct Responses by Group and Age on Second-Order ToM

| Age Group | N | Diagnostic Group | $2^{nd}$-order ToM [a] | Control [b] |
|---|---|---|---|---|
| Younger | 23 | SLI | 0.96 (1.19) | 3.87 (1.10) |
| | 24 | NSLI | 1.83 (1.24) | 4.83 (0.87) |
| Older | 20 | SLI | 2.50 (1.10) | 4.85 (1.09) |
| | 25 | NSLI | 3.76 (0.52) | 5.64 (0.76) |
| All Ages | 43 | SLI | 1.67 (1.38) | 4.33 (1.19) |
| | 49 | NSLI | 2.82 (1.35) | 5.25 (0.90) |
| All Younger | 47 | SLI + NSLI | 1.40 (1.28) | 4.36 (1.09) |
| All Older | 45 | SLI + NSLI | 3.20 (1.04) | 5.29 (0.99) |

[a] Maximum score = 4.00. [b] Maximum score = 6.00.

Table 6

Means and Standard Deviations for Correct Responses by Group and Age for Sarcasm and ToM Composite

| Age Group | N | Diagnostic Group | 1st-order Sarcasm [a] | 2nd-order Sarcasm [a] | Sarcasm Control [b] | ToM Composite [c] |
|---|---|---|---|---|---|---|
| Younger | 23 | SLI | 0.65 (1.07) | 0.30 (0.56) | 5.57 (0.73) | 5.96 (4.28) |
| | 24 | NSLI | 1.25 (1.22) | 0.54 (0.72) | 5.83 (0.82) | 10.75 (3.94) |
| Older | 20 | SLI | 1.65 (1.23) | 0.70 (0.86) | 5.90 (0.31) | 12.85 (4.22) |
| | 25 | NSLI | 2.68 (0.85) | 0.60 (0.96) | 5.92 (0.28) | 17.24 (1.59) |
| All Ages | 43 | SLI | 1.12 (1.24) | 0.49 (0.74) | 5.72 (0.59) | 9.16 (5.46) |
| | 49 | NSLI | 1.98 (1.27) | 0.57 (0.84) | 5.88 (0.60) | 14.06 (4.41) |
| All Younger | 47 | SLI + NSLI | 0.96 (1.18) | 0.43 (0.65) | 5.70 (0.78) | 8.40 (4.73) |
| All Older | 45 | SLI + NSLI | 2.22 (1.15) | 0.64 (0.91) | 5.91 (0.29) | 15.29 (3.73) |

[a] Maximum score = 3.00. [b] Maximum score = 6.00. [c] Maximum score = 21.00

Table 7

Means and Standard Deviations by Group and Age for Working Memory and Sibling Measures

| Age Group | N | Diagnostic Group | Working Memory Span [a] | Siblings in Home | Older Siblings in Home |
|---|---|---|---|---|---|
| Younger | 23 | SLI | 3.30 (1.40) | 1.78 (1.35) | 1.17 (1.19) |
| | 24 | NSLI | 3.46 (1.67) | 1.46 (1.02) | 1.13 (1.08) |
| Older | 20 | SLI | 5.05 (2.21) | 1.95 (1.19) | 1.05 (1.10) |
| | 25 | NSLI | 4.88 (2.35) | 1.04 (0.73) | 0.56 (0.71) |
| All Ages | 43 | SLI | 4.12 (2.00) | 1.86 (1.26) | 1.12 (1.14) |
| | 49 | NSLI | 4.18 (2.15) | 1.25 (0.90) | 0.84 (0.94) |
| All Younger | 47 | SLI + NSLI | 3.38 (1.53) | 1.62 (1.19) | 1.15 (1.12) |
| All Older | 45 | SLI + NSLI | 4.96 (2.27) | 1.44 (1.06) | 0.78 (0.93) |

[a] Maximum score = 9.00

Table 8

Means and Standard Deviations for T Scores by Group and Age and CBCL

| Age Group | N | Diagnostic Group | Internalizing [a] | Externalizing [a] | Total [a] |
|---|---|---|---|---|---|
| Younger | 23 | SLI | 48.52 (11.03) | 49.43 (9.44) | 52.00 (9.65) |
| | 24 | NSLI | 47.17 (9.58) | 51.38 (9.09) | 51.00 (10.33) |
| Older | 20 | SLI | 48.35 (8.99) | 48.50 (11.41) | 51.50 (10.16) |
| | 25 | NSLI | 54.48 (10.61) | 51.24 (13.80) | 54.76 (10.81) |
| All Ages | 43 | SLI | 48.44 (10.02) | 49.00 (10.29) | 51.77 (9.78) |
| | 49 | NSLI | 50.90 (10.67) | 51.30 (11.61) | 52.92 (10.64) |
| All Younger | 47 | SLI + NSLI | 47.83 (10.23) | 50.43 (9.21) | 51.49 (9.91) |
| All Older | 45 | SLI + NSLI | 51.76 (10.28) | 50.02 (12.72) | 53.31 (10.64) |

[a] Scale mean and standard deviation = 50.00 (10.00)