Theory of Mind and Pretend Play in Children with Specific Language Impairment

by

Melanie Stich

A thesis submitted in conformity with the requirements for the degree of Ph.D. (Doctor of Philosophy)

Department of Speech-Language Pathology
University of Toronto

© Copyright by Melanie Stich (2010)
Theory of Mind and Pretend Play in Children with Specific Language Impairment

Melanie Stich

Ph.D. (Doctor of Philosophy)

Department of Speech-Language Pathology
University of Toronto

2010
Abstract

The ability to represent the mental states of others (i.e., Theory of Mind, ToM) is vital for social interaction. There is limited information on ToM knowledge in children with specific language impairment (SLI). These children have deficient language abilities that cannot be explained by hearing, cognitive, or neurological problems. Furthermore, children with SLI experience difficulty in initiating and maintaining social pretend play. Language, pretend play, and ToM typically develop in concert, which may indicate that they share an underlying capacity for representation. Given that language is deficient in children with SLI, these children may have problems with ToM, which might be related to their social behaviors during pretend play. This study was the first to investigate the association between ToM and pretend play in children with and without SLI.

Twenty-two children with SLI and 22 with typical development (TD), between 48-71 months of age, participated in this study. Children engaged in a variety of ToM tasks and participated in two pretend play assessments: a standardized pretend play assessment and a role play activity. Children with SLI scored significantly lower on ToM tasks and engaged less often in some sophisticated forms of pretend play than their age-matched peers with TD. After controlling for language and SES, there were no significant associations between ToM and pretend play in children with and without SLI. When language groups were analyzed individually, different patterns of associations emerged for children with and without SLI. ToM was positively associated with pretend play in children with TD but negatively associated in children with SLI. Moreover, inconsistent patterns of associations were observed for some children with SLI (i.e., poor ToM understanding but sophisticated pretend play or vice versa).
This study demonstrated that children with SLI may also have concomitant problems in ToM and pretend play, which may have implications for clinical assessment and intervention. The study contributes to the literature by investigating the link between ToM and different forms of sophisticated pretend play in children with and without SLI. Given the different patterns of partial correlations, future investigation of the relationship between ToM and pretend play relationship is warranted.
I would like to express my deepest gratitude to those who supported me during my doctorate studies:

- To the members of my supervisory committee, Dr. Carla Johnson, Dr. Luigi Girolametto, and Dr. Joan Peskin for their invaluable guidance during design, implementation, and analysis of this study. In particular, I would like to thank Dr. Carla Johnson, my supervisor and Doktormutter, for her patience as well as for her personal and professional support. I greatly appreciate all the time and effort she dedicated to reviewing the various drafts to this manuscript.

- To all the parents who participated in this study and their children who reminded me how much fun it is to play. To all kindergarten directors, particularly to Petra Voigt, who opened their kindergartens to me and enabled the implementation of this study. Further, to Professor Ptok who generously provided financial and professional support during implementation of the study in his clinic.

- To my husband who was there for me with his love, support, and encouragement every single step of the way.

- To my family who was supportive of my studies and let their daughter move far away – they were always with me.

- To my friends Anny and Heidi who believed in me and who made this journey more fun.
# Table of Contents

Abstract .......................................................................................................................................... iii
Acknowledgments ................................................................................................................................ v
Table of Contents ........................................................................................................................... vi
List of Tables .................................................................................................................................. ix
List of Figures ................................................................................................................................... x
List of Appendices .......................................................................................................................... xi

1 Introduction ..................................................................................................................................... 1
  1.1 Theory of Mind .......................................................................................................................... 2
    1.1.1 What is Theory of Mind? ...................................................................................................... 2
    1.1.2 The Development of Theory of Mind .................................................................................. 2
    1.1.3 False Belief Understanding within Theory of Mind Development .................................... 4
    1.1.4 ToM and General Language Abilities ................................................................................ 6
    1.1.5 ToM and Specific Language Abilities ............................................................................... 7
    1.1.6 Summary ........................................................................................................................... 9
    1.1.7 Individual Differences ....................................................................................................... 9
      1.1.7.1 Siblings ....................................................................................................................... 9
      1.1.7.2 Family Discourse and Parenting Style ....................................................................... 11
      1.1.7.3 Socioeconomic Status ............................................................................................... 11
      1.1.7.4 Summary ................................................................................................................... 12
  1.2 Theory of Mind in Children with Specific Language Impairment ........................................... 13
    1.2.1 Theory of Mind in School-Age Children with Specific Language Impairment .................. 13
    1.2.2 Theory of Mind in Preschool Children with Specific Language Impairment ...................... 16
    1.2.3 Summary ......................................................................................................................... 18
  1.3 Pretend Play .............................................................................................................................. 21
    1.3.1 What is Pretend Play? ........................................................................................................ 21
    1.3.2 Development of Pretend Play ........................................................................................... 22
    1.3.3 Pretend Play and Language .............................................................................................. 26
    1.3.4 Pretend Play and Narrative Competence ........................................................................ 28
    1.3.5 Summary .......................................................................................................................... 29
    1.3.6 Pretend Play and Individual Differences in Socioeconomic Status .................................. 30
  1.4 Pretend Play in Children with Specific Language Impairment ................................................. 30
  1.5 Pretend Play and Theory of Mind ........................................................................................... 34
  1.6 Conclusions .............................................................................................................................. 37
  1.7 Research Aims and Hypotheses ............................................................................................... 40

2 Method ......................................................................................................................................... 41
  2.1 Power Analysis .......................................................................................................................... 41
  2.2 Participants ............................................................................................................................... 41
    2.2.1 Inclusion and Exclusion Criteria ......................................................................................... 42
      2.2.1.1 Coloured Pogressive Matrices (CPM) ........................................................................ 42
      2.2.1.2 Sprachentwicklungstest für 3-5jährige Kinder (SETK 3-5) ........................................ 42
2.2.1.3 Aktiver Wortschatztest – Revised (AWST-R) .................................................. 44
2.2.1.4 Questionnaire on Demographic Information .............................................. 44
2.2.2 Matching Criteria ............................................................................................ 44
2.2.3 Description of Participants ............................................................................. 45
2.3 Procedures .......................................................................................................... 48
2.3.1 Measures .......................................................................................................... 52
2.3.1.1 Theory of Mind ............................................................................................ 52
2.3.1.2 Child-Initiated Pretend Play Assessment ..................................................... 54
2.3.1.3 Role Play Activity ....................................................................................... 57
2.3.1.4 Pretend Play Questionnaire ....................................................................... 57
2.3.2 Coding Procedures .......................................................................................... 58
2.3.2.1 Child-Initiated Pretend Play Assessment ..................................................... 58
2.3.2.2 Role Play Activity ....................................................................................... 58
2.3.2.3 Validity and Reliability .............................................................................. 62

3 Results .................................................................................................................... 66
3.1 Theory of Mind ..................................................................................................... 66
3.2 Differences in Family Background and ToM Performance .................................. 73
3.3 Pretend Play .......................................................................................................... 73
3.3.1 Pretend Play Questionnaire ............................................................................ 74
3.3.2 Gender Differences on the Pretend Play Questionnaire .................................. 74
3.3.3 Child-Initiated Pretend Play Assessment ....................................................... 77
3.3.4 Gender Differences on the Child-Initiated Pretend Play Assessment .............. 80
3.3.5 Pretend Play Examples .................................................................................. 80
3.3.6 Role Play Activity ........................................................................................... 81
3.3.6.1 Gender Differences in the Role Play Activity .............................................. 85
3.3.6.2 Experience with Playmobil Toys as a Potential Confound ......................... 85
3.3.7 Pretend Play Examples .................................................................................. 85
3.3.8 Pretend Play Prompts ................................................................................... 86
3.4 Relationship between Theory of Mind and Pretend Play .................................. 87

4 Discussion ............................................................................................................... 94
4.1 Theory of Mind ..................................................................................................... 94
4.2 Pretend Play Measures ......................................................................................... 98
4.2.1 Child-Initiated Pretend Play Assessment ....................................................... 98
4.2.2 Role Play Activity ............................................................................................ 101
4.2.3 Pretend Play Questionnaire ......................................................................... 103
4.2.4 Relations of Gender and SES to Pretend Play ............................................. 104
4.3 Theory of Mind and Pretend Play ...................................................................... 105
4.4 General Implications of Findings ....................................................................... 108
4.4.1 Theoretical Implications .............................................................................. 108
4.4.2 Research Implications ................................................................................. 111
4.4.3 Practical Implications ................................................................................... 113
4.5 General Limitations ............................................................................................. 114
4.5.1 Limitations in Study Design .......................................................................... 114
4.5.2 Limitations in Generalizability ...................................................................... 115
4.5.3 Limitations in Measurement................................................................. 116
4.6 Contributions......................................................................................... 117

References ................................................................................................. 119
Appendices ................................................................................................. 135
List of Tables

Table 1  Task Descriptions for the Theory of Mind Scale ........................................4
Table 2  Summary of Studies Comparing Theory of Mind Performance in Children with and without SLI .................................................................19
Table 3  Milestones in the Development of Pretend Play in Children with Typical Development .................................................................23
Table 4  Milestones in the Development of Language, Theory of Mind (ToM), and Pretend Play .................................................................38
Table 5  Inter-Correlations of Language Subtests of the SETK 3-5 .................................................43
Table 6  Characteristics, Cognitive and Language Abilities for Children with and without SLI .................................................................47
Table 7  Description and Examples of Play Behavior Codes of the Child-Initiated Pretend Play Assessment .........................................................55
Table 8  Description of Role Play Activity Codes ..................................................................60
Table 9  Kappa Values for the Child-Initiated Pretend Play Assessment (ChIPPA) and Role Play Activity Codes .................................................................64
Table 10  Inter-Correlations of Theory of Mind Tasks .................................................................68
Table 11  Means and Standard Deviations (in parentheses) of Correctly Answered Theory of Mind Questions for Children with and without SLI .................................................................71
Table 12  Means and Standard Deviations (in parentheses) for Children with and without SLI on the Pretend Play Questionnaire .................................................................76
Table 13  Means, Standard Deviations (in parentheses), and Effect Sizes for Number of ChIPPA Segments Coded, and PEPA and NOS Standard Scores for Children with and without SLI .................................................................79
Table 14  Correlations and Partial Correlations for Theory of Mind (ToM), PEPA, Direct Role Play, Language, and SES in Children with and without SLI .................................................................89
Table 15  Correlations and Partial Correlations for Theory of Mind (ToM), PEPA, Direct Role Play, Language, and SES in Children with TD .................................................................91
Table 16  Correlations and Partial Correlations for Theory of Mind (ToM), PEPA, Direct Role Play, Language, and SES in Children with SLI .................................................................92
List of Figures

Figure 1  Testing and Matching Procedures of Participants………………………………..51
Figure 2  Percentages of Children with and without SLI who Passed Target and
          Control Questions on Theory of Mind (ToM) Tasks…………………………… .72
Figure 3  Percentages of Pretend Play Codes during Role Play Activity
          for Children with and without SLI……………………………………………….84
List of Appendices

Appendix A  English Translation of Questionnaire on Demographic Information……………135
Appendix B  English Translation of Pretend Play Questionnaire…………………………138
Appendix C  English Translation of Informed Consent Form (Clinic) ...........................141
Appendix D  English Translation of Informed Consent Form (Kindergarten) ..........146
Appendix E  English Translation of Verbal Assent Form (Child) .................................151
Appendix F  English Translation of Theory of Mind Tasks.........................................152
Appendix G  English Translation of Story Description for Role Play Activity……………157
1 Introduction

Early childhood is characterized by major developmental changes in areas such as language, pretend play and Theory of Mind (i.e., being aware of another person’s mental states). Language, pretend play, and Theory of Mind require cognitive resources (e.g., memory or processing skills) and may, therefore, be considered cognitive abilities (Rice & Kemper, 1984). Further, these abilities typically develop in concert, that is, along similar timelines. Language may play an important role in the development of ToM and pretend play because language enables the child to express desires and intentions (Bartsch & Wellman, 1995), and moreover, allows the child to take part in social interactions, such as pretend play (Farver, 1992). Others, however, have argued that these cognitive abilities may be distinct from each other, with few or no connections (Fodor, 1983).

Children with specific language impairment (SLI) have often been included in studies to investigate, for instance, relationships between language and cognition or language and pretend play. These children provide an interesting case because they experience severe limitations in their language abilities without other serious cognitive, neurological, or hearing problems (Leonard, 1998). This allows researchers to investigate how other cognitive abilities develop when language is deficient.

Most often, language, pretend play, and Theory of Mind (ToM) have been studied in isolation in children with typical development (TD). Other studies investigated single associations, such as language-play or language-ToM relations but considerably less research has investigated the interrelations of all three abilities. Despite the fact that children with SLI may provide valuable insights, little is known about how preschool children with SLI perform on ToM tasks and even less is known about how ToM relates to their social-cognitive behaviors, particularly pretend play. The aims of this study were to compare the ToM and pretend play performance of children with SLI with that of children with TD and to investigate possible associations between ToM and pretend play.

This chapter begins with a review of the literature regarding the development of ToM in children with TD. The relationships between general and specific aspects of language and ToM are described as well as factors that might contribute to individual differences in ToM.
performance. Subsequently, studies are discussed that investigated ToM in children with SLI. The second part of this chapter provides an overview of pretend play development in children with TD as well as information on how pretend play relates to language. Then, empirical evidence from children with SLI is discussed. The final part describes evidence concerning the relations between play and ToM.

1.1 Theory of Mind

1.1.1 What is Theory of Mind?
The term Theory of Mind (ToM) refers to a broad social-cognitive concept. It describes the ability to represent one’s own and other people’s mental states (i.e., their desires, emotions, beliefs, and intentions). The main assumption is that people act to fulfill their desires, given their beliefs (Wellman, 1990). Thus, by knowing their mental states we are able to predict and explain others’ behaviors and interactions, which is vital for social interaction (Astington & Baird, 2005). During the preschool years, children learn to ascribe thoughts and beliefs to themselves and others. Furthermore, children begin to understand that people will act upon their beliefs because they think they are true. Thus, children can then influence a person’s behavior by manipulating his beliefs (e.g., with tricks, lies, or secrets) (Astington, 2003).

1.1.2 The Development of Theory of Mind
It is important to realize that ToM is not either absent or present in a child. Children’s ability to interpret other people’s behaviors develops gradually during childhood (Wellman & Liu, 2004). Based on the analysis of spontaneous speech samples, Bartsch and Wellman (1995) described the development of ToM in three main phases.

Children’s utterances were coded for cognitive terms of thought and belief (e.g., remember, think, know, believe, expect, wonder, and dream) as well as for desire terms (e.g., want, hope, wish, care (about), and afraid (that)) (Bartsch & Wellman, 1995; Shatz, Wellman, & Silber, 1983). Bartsch and Wellman (1995) reported that in the first phase, two-year-old children mostly talked about desires. That is, their talk was about what they did (or did not) want or like with respect to objects or actions. Furthermore, young children understood that different people may have different desires. For example, 18-month-old children understood that another person might desire broccoli whereas they themselves would prefer crackers to broccoli (Repacholi &
Gopnik, 1997). Around children’s third birthdays, they entered the second phase. At this time, there was a relative decline in talk about desires whereas talk about beliefs started to increase (Bartsch & Wellman, 1995). However, children still used desire terms more frequently than belief terms and were more likely to explain behaviors based on desires rather than on beliefs. Children entered the third phase when they were approximately four years of age. Then their talk about belief terms showed a relative increase and they also began to explain other people’s behaviors based on their (true or false) beliefs (Bartsch & Wellman, 1995).

Bartsch and Wellman (1995) referred to these phases as desire psychology, desire-belief psychology, and belief-desire psychology. In each phase, children demonstrated an increasing understanding of other people’s minds. Children understood desires before they understood other people’s beliefs (e.g., Gopnik & Slaughter, 1991). Further, children came to understand that people actively and constantly form representations about the world (i.e., have beliefs) before they realized that these representations did not always mirror reality (i.e., these beliefs can be false). Thus, the preschool period was characterized by many important developmental changes in ToM understanding.

Some of these ToM aspects are captured in the Theory of Mind scale (Wellman & Liu, 2004). This scale assesses understanding of desires, beliefs, false beliefs, and emotions and, therefore, provides a comprehensive picture of ToM understanding in children. The scale is ordered from the easiest to the hardest tasks (see Table 1) and the percentages in parentheses show the correct performance of 75 typically developing children between the ages of 2;11 and 6;6 (years; months).
Table 1

Task Descriptions for the Theory of Mind Scale

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diverse Desires (95%)</td>
<td>Child judges that two persons (child vs. someone else) have different desires about the same objects</td>
</tr>
<tr>
<td>Diverse Beliefs (84%)</td>
<td>Child judges that two persons (child vs. someone else) have different beliefs about the same object, when the child does not know which belief is true or false</td>
</tr>
<tr>
<td>Knowledge Access (73%)</td>
<td>Child sees what is in a box and judges (yes/no) the knowledge of another person who does not see what is in a box</td>
</tr>
<tr>
<td>Contents False Belief (59%)</td>
<td>Child judges another person’s false belief about what is in a distinctive container when child knows what is in the container</td>
</tr>
<tr>
<td>Real-Apparent Emotion (32%)</td>
<td>Child judges that a person can feel one thing but display a different emotion</td>
</tr>
</tbody>
</table>


1.1.3 False Belief Understanding within Theory of Mind Development

Although ToM understanding shows a developmental progression (Wellman & Liu, 2004), the major focus in ToM research was on children’s ability to pass false belief tasks. The passing of false belief tasks is just one step in ToM understanding; however, it has long been viewed as the most important one because it marks the transition where children ascribe thoughts and beliefs
to themselves and others (Astington, 2003). To assess these first-order false beliefs (i.e., understanding one person’s belief), researchers often use two tasks: unexpected contents and change-of-location. In the former, the child sees a familiar box (e.g., Smarties) that actually contains unexpected contents (e.g., paper clips). After the child has experienced the false belief himself, he is asked what a friend would say when shown the box and asked to identify its contents before looking inside it. In the change-of-location task, one person leaves an object at place A and in his absence the object is moved to place B by another person. When the first person returns, the child is asked where he will look for the object.

Researchers tested whether one of the false belief tasks was more difficult than the other. However, it was concluded that this was not the case (Wellman, Cross, & Watson, 2001). Thus, the often replicated finding is that 3-year-olds told the true state of affairs, that is, they said what the box actually contained or where the object really was without appreciating that the other person had a false belief and did not know the true contents or location. On average, typically developing children began to pass false belief tasks when they were 4 years old (Wellman et al., 2001). That is, their performance showed a shift from usually incorrect answers at around 3 years of age to usually correct answers at around 4;6 years of age. The false belief tasks have been criticized for being linguistically too complicated because the use of language might mask the real ToM competence in children (e.g., Lewis & Osborne, 1990). This claim has generated literally hundreds of studies and the tasks have been manipulated in various ways. For instance, the mode of presentation (using dolls, pictures, or video clips), the phrasing of the questions (e.g., “Where will x look?” vs. “Where will x look first?”) or the amount of linguistic input have been altered to test if tasks can be made easier and thus, more likely for children to demonstrate their understanding of false belief.

However, a meta-analysis showed that none of the task modifications facilitated ToM performance enough that 3-year-olds would pass the false belief tasks at above-chance levels (Wellman et al., 2001). Furthermore, a cross-cultural meta-analysis reported that false belief tasks may measure the same construct across cultures (Liu, Wellman, Tardif, & Sabbagh, 2008). The meta-analysis revealed that Chinese children displayed a similar developmental trajectory of ToM understanding as has been demonstrated in North American children. However, the change from below-chance to above-chance performance became evident by 64 months, and thus, considerably later in Chinese children than in North American children, who showed that
change around 48 months. These differences might be explained by cultural differences in parenting styles or social backgrounds (Liu et al., 2008).

1.1.4 ToM and General Language Abilities

The relation between language abilities and ToM was the topic of a recent meta-analysis (Milligan, Astington, & Dack, 2007) involving more than 100 studies of English-speaking preschool children with TD. Aspects of language that were under investigation included general language ability, semantics, receptive vocabulary, and syntax (including memory for sentential complements). The assessment of memory for complements is based on the assumption that mastery of a specific linguistic structure (i.e., sentential complements) is important for ToM understanding. In this assessment, the child listens to a statement such as “He thought he found his ring, but it was really a bottle cap” and is then asked “What did he think?” Any reference to the ring instead of the bottle cap is scored as correct (J. de Villiers & Pyers, 2002). More information on sentential complements is provided in the following section, ToM and Specific Language Abilities.

Milligan et al. (2007) reported that all aspects of language were significantly related to false belief understanding. The effect sizes varied depending on the type of language ability from \( r^2 = .12 \) for receptive vocabulary, \( r^2 = .23 \) for semantics, \( r^2 = .27 \) for general language ability, to \( r^2 = .29 \) and \( r^2 = .44 \) for syntax and memory for complements, respectively. Furthermore, age was correlated with language and ToM \( (r = .43) \); however the relation between ToM and language remained significant even when age was partialled out \( (r = .31) \). Moreover, this analysis found a bi-directional effect between language and ToM. However, the correlation between earlier language competence and later false belief task performance was significantly stronger than that for earlier false belief and later language performance. This suggested that language may play a causal role in ToM development.

The meta-analysis included only studies that were conducted with English-speaking children with TD. However, cross-linguistic research (Cheung, Chen, & Yeung, 2009) and research with children with atypical development (e.g., children with autism or deafness) supported the finding that language is important for ToM to develop (Lind & Bowler, 2009; Peterson, Wellman, & Liu, 2005; Schick, P. de Villiers, J. de Villiers, & Hoffmeister, 2007). For instance, researchers compared the ToM performance of deaf children who have deaf parents and were
exposed to sign language from early on (‘native signers’) to that of deaf children of hearing parents, who were exposed to sign language late (‘late signers’) (Schick et al., 2007). They found that native signers performed significantly better on ToM tasks than their late-signing peers and they concluded that growing up deaf and having limited access to language had a negative impact on ToM development.

Some researchers argued that the association between ToM and language might only be an artifact of testing because language is the means of assessing ToM competence (Chandler, Fritz, & Hala, 1989). Others, however, have tried to find out what makes language so important, and which specific aspects of language might be related to ToM development. Therefore, researchers have investigated several aspects of semantics and syntax more closely (e.g., J. de Villiers, 2005; Ruffman, Slade, & Crow, 2002). J. de Villiers stated that the mastery of certain syntactic structures (i.e., sentential complements) was required for ToM development, whereas others (e.g., Peterson & Slaughter, 2003; Ruffman, Slade, Devitt, & Crowe, 2006) demonstrated a close relationship between specific lexical items (i.e., mental state verbs) and ToM and concluded that they are important for ToM development. Further, cross-linguistic research and studies with atypical populations were also conducted to shed some light on the language-ToM relationship. The aspects of semantics and syntax are highly intertwined because sentential complements are embedded under mental state terms. However, they are described individually in turn. Only a brief description is provided, as these specific linguistic aspects were not the research focus of this thesis.

1.1.5 ToM and Specific Language Abilities

Specific linguistic aspects that were considered to play an influential role in the development of ToM were the acquisition of mental state verbs as well as sentential complements (J. de Villiers, 2005; Ruffman et al., 2002). Mental states are part of people’s internal subjective experiences and are not directly observable. These abstract mental states can only be expressed and shared with others via language, that is, by the use of mental state verbs, such as think, remember, want, like, know (Bartsch & Wellman, 1995). Thus, one assumption is that the child’s use of these mental state verbs indicates that he or she has an understanding of ToM. The analysis of spontaneous speech samples showed that by approximately four years of age, children used belief terms in an adult-like manner and furthermore, began to explain other people’s behaviors
based on their (true or false) beliefs (Bartsch & Wellman, 1995). As this parallels the pattern of when children begin to pass false belief tasks, the hypothesis was that acquisition of mental state verbs was related to ToM development.

Further, mental state verbs of belief take a special syntactic structure that children master by the age of approximately four years (J. de Villiers, 2005). An example would be “Susie thinks that the chocolate is in the drawer”. Here, the mental verb *think* is the main verb, followed by a subordinate clause that describes the object (in italics). This syntactic structure (i.e., sentential complement) allows children to talk about mental states and compare them with reality. Even when the proposition is false – there is no chocolate – the sentence itself remains true that Susie thinks that the chocolate is there (J. de Villiers, 2005; J. de Villiers & Pyer, 2002). The acquisition of the sentential complement structure precedes the development of children’s success in false belief tasks. Thus, J. de Villiers claimed that this specific syntactic structure played a causal role in promoting false belief understanding.

There is more evidence that supports the assumption that mental state verbs are related to ToM development than there is for the sentential complements account. Evidence for the role of mental state verbs in ToM comes from correlational and longitudinal studies of children with TD (Cheung et al., 2009; Howard, Mayeux, & Naigles, 2008; Peterson & Slaughter, 2003; Ruffman et al., 2002; Ruffman et al., 2006; Symons, Peterson, Slaughter, Roche, & Doyle, 2005) as well as from children with atypical development (Moeller & Schick, 2006). However, there are few studies that support the sentential complement account (e.g., J. de Villiers & Pyers, 2002; Schick et al., 2007). A number of studies used variations of the memory for complements task (e.g., Cheung et al., 2009; Miller, 2004); however, given these methodological differences, it is difficult to compare results across the different studies and judge the validity of the account.

Moreover, both accounts face conflicting findings that contradict their main assumptions. For instance, a training study (Peskin & Astington, 2004) and cross-linguistic research indicated that mental state verbs may not be the only factor or a causal factor in ToM development (Liu et al., 2008; Shatz, Diesendruck, Martinez-Beck, & Akar, 2003). On the other hand, inconsistent results were reported for the sentential complements account in studies that included children with atypical development (Farrar, Johnson, Tompkins, Easters, Zilisi-Medus, & Benigno, 2009;
Lind & Bowler, 2009). These studies indicated that mastery of sentential complements might not be a prerequisite for false belief understanding.

1.1.6 Summary
There are strong associations between language and ToM. Although specific language aspects, such as mental state verbs or sentential complements, may provide scaffolding for ToM development, they did not exclusively influence the overall rate of ToM development or necessarily precede it (Hale & Tager-Flusberg, 2003; Peskin & Astington, 2004). On the other hand, training studies (Lohmann & Tomasello, 2003; Peskin & Astington, 2004) suggested that other language aspects, such as conversing about ambiguous objects or highlighting different perspectives, might facilitate ToM understanding.

It is conceivable that languages might affect ToM performance differently. For instance, English-speaking children benefited from an explicit temporal marker (“Where will Susie look first?”), whereas Chinese-speaking children performed more poorly when this marker was used (Liu et al., 2008). Thus, what might promote ToM performance in one language might not be helpful in another.

1.1.7 Individual Differences
Despite the general tendency for Western children with TD to pass false belief tasks when they are approximately 4 years old, substantial individual differences in young children’s ToM development have been reported (Wellman et al., 2001). For instance, some children initially passed false belief tasks when they were three years of age, others when they were almost six. Areas that seem to be associated with individual differences in ToM development in children are family environment and social background. Factors that have been investigated are the presence of siblings, family discourse and parenting style, as well as the socioeconomic status of a family.

1.1.7.1 Siblings
Children may benefit from having siblings, as children who had one or more siblings passed false belief tasks earlier than only children (McAlister & Peterson, 2007; Perner, Ruffman, & Leekam, 1994). However, the evidence is mixed. Some studies reported that only older siblings facilitated ToM understanding (Ruffman, Perner, Naito, Parkin, & Clemens, 1998); others found that younger siblings as well as twins also had a positive influence (Peterson, 2000).
Furthermore, Peterson (2000) found that ToM performance was better in children who had child-aged siblings (i.e., siblings between 12 months and 12 years of age), whereas children who had siblings who were very young infants or teenagers performed comparably to only children on ToM tasks. On the other hand, Jenkins and Astington (1996) reported that the number of siblings a child had was important and not their respective ages. Moreover, the authors found a stronger association between false belief understanding and the presence of siblings in children with lower language skills than in children with higher language skills.

Possible explanations for the facilitating effect might be that having siblings may increase the opportunities for social interaction and forms of communication that might require appreciation of another’s perspective (e.g., tricking, teasing, and deceiving) (McAlister & Peterson, 2007). Older siblings may provide more possibilities to engage in pretend play and to problem solve and negotiate conflicts, which may be beneficial for ToM development. An alternative explanation might be varying linguistic input (Jenkins, Turrell, Kogushi, Lollis, & Ross, 2003). Jenkins et al. (2003) found that 4-year-old children who had a 6-year-old sibling were exposed to more talk involving cognitive terms (think, know) than children without an older sibling. In addition to the cognitive talk directed to them, children with older siblings heard more talk directed to their sibling. As discussed earlier, there is evidence that talk about mental states is linked to ToM performance (Peterson & Slaughter, 2003).

On the other hand, some studies did not find a significant link between siblings and ToM (Cole & Mitchell, 2000; Pears & Moses, 2003; Peterson & Slaughter, 2003). Some of the differences between studies might be explained by differences in socioeconomic status (SES) of the samples. The children in the studies that reported a link came from middle-class backgrounds and had well-educated parents, whereas the other studies included a higher proportion of low-income and single-parent families. However, Taylor and Carlson (1997) also did not detect a significant association between ToM and siblings even though their children had middle-class backgrounds and they had a relatively large sample (N = 152). Further, the different findings on who exerts a positive influence (older vs. younger siblings) might be explained by varying sample characteristics of the siblings. Some siblings might have been too old or young to interact or there might have been large differences in age between the participating child and his nearest child-aged sibling. Differences might also be due to factors such as closeness or temperament of the siblings.
1.1.7.2 Family Discourse and Parenting Style

Familial discourse and parenting style have also been associated with preschoolers’ ToM understanding. Dunn, Brown, and Beardsall (1991) investigated the conversations of 36-month-old children with their mothers and siblings. Despite wide variations in conversations, they found that disputes were the most frequent reason to discuss the feelings of others. Three-year-old children who were frequently exposed to talk about feelings did significantly better in judging the emotions of an unfamiliar adult three years later than children with less exposure to feeling-state talk.

These results were supported by a study in which mothers were asked how they would deal with five disciplinary situations (Ruffman, Perner, & Parkin, 1999). Their answers were coded as to whether they would talk about feelings, have a general discussion, or reprimand their children. Children of mothers who discussed the feelings of the implicated person (how the victim might feel) had better ToM scores than children of mothers who would only reprimand the child. This finding was supported by studies that reported that children’s ToM performance benefited from mothers’ as well as fathers’ use of mental state words such as sad, happy, think, and know when describing a series of pictures, independent of the child’s language abilities (LaBounty, Wellman, Olson, Lagattuta, & Liu, 2008; Ruffman et al., 2002). These studies suggest that discussing social conflicts and the feelings of others may promote the development of ToM in children.

1.1.7.3 Socioeconomic Status

Family discourse and parenting practices are related to SES (Bornstein, Hynes, & Painter, 1998; Hart & Risley, 1999). Mothers with low SES background interacted and talked less (e.g., 600 words per hour) to their young child than mothers with middle-class (1,200 words per hour) or upper-class backgrounds (e.g., 2,100 words per hour) (Hart & Risley, 1999). Further, mothers with higher SES levels used a more diverse vocabulary and longer sentences when talking to their children than mothers with lower SES levels (Hoff-Ginsberg, 1991). Thus, children who had more language exposure (i.e., from higher SES backgrounds) had a greater vocabulary and were more talkative than children who had less exposure (i.e., from lower SES backgrounds) (Hart & Risley, 1999; Hoff & Tian, 2005).
Moreover, SES has been shown to correlate with ToM (Hoff-Ginsberg, Laursen, & Tardif, 1995; Hughes, Jaffee, Happé, Taylor, Caspi, & Moffitt, 2005; Jenkins et al., 2003; Moeller & Schick, 2006; Shatz, et al., 2003; Symons et al., 2005). Most ToM research has been conducted with children from middle-class backgrounds with well-educated parents. Studies that included children from more diverse SES backgrounds reported findings that differed from those for children with middle-class backgrounds.

Holmes, Black, and Miller (1996) recruited preschool children from the Head Start Program. This program was established to provide resources for low-income families to counteract the effects of poverty on developing children. These preschool children showed levels of ToM performance that were lower than typically found in the ToM literature (Holmes et al., 1996). Children with a mean age of 4;3 answered false belief questions correctly between 28% and 52% of the time, whereas children with a mean age of 5;3 did so between 52% and 84% of the time. Thus, there was a trend towards mastery of false belief, but later than typically reported (Wellman et al., 2001). It is difficult to conclude how much SES may have affected performance because this study did not include a middle-class comparison group and did not assess the language abilities of the participating children. A comparison group was included in a study conducted by Cutting and Dunn (1999). About half of the children had working-class backgrounds and the other half had middle-class backgrounds. Children from the working-class performed significantly worse on a receptive vocabulary test and on false belief tasks than children from the middle-class. The group differences in ToM performance remained even when language abilities were partialled out. Thus, there is some evidence that SES is linked to ToM development in children.

1.1.7.4 Summary

Several factors in the child’s social environment, such as siblings, parenting style, and SES, may be related to individual differences in ToM performance. All three factors share the commonality that children are exposed differentially to language. This strengthens the important role of language in ToM development as has been discussed in the previous section.
1.2 Theory of Mind in Children with Specific Language Impairment

Children with specific language impairment (SLI) experience significant limitations in their language abilities despite normal nonverbal intelligence and in the absence of hearing or neurological dysfunctions (Leonard, 1998). Given that language is strongly related to ToM performance, it seems reasonable to assume that children with SLI might show difficulties with false belief understanding and other aspects of ToM. On the other hand, children with SLI have relatively normal nonverbal cognitive abilities; therefore, predicting the behaviors of others, for instance by pointing, may not be a problem. The development of ToM has been well researched in children with TD. However, little is known about ToM and, in particular, false belief understanding in preschool children with SLI. Further, existing evidence is not conclusive.

1.2.1 Theory of Mind in School-Age Children with Specific Language Impairment

Different aspects of ToM understanding in school-age children with language impairments have been investigated. Thus, it is not always possible to compare study results directly because ToM tasks vary and participants may differ in their characteristics or age. The majority of ToM studies included school-age children with language impairments as a comparison group to children with autism (Colle, Baron-Cohen, & Hill, 2007; Perner, Frith, Leslie, & Leekam, 1989; Ziatas, Durkin, & Pratt, 1998). In general, children with language impairments performed better than children with autism and often comparably to the control groups of children with TD on different ToM tasks. However, there were methodological weaknesses that limit the generalizability of these results.

One weakness concerns the description of the samples, which makes it difficult to be confident that all participating children indeed had SLI. That is, inclusion or exclusion criteria to establish the SLI group were not explicitly stated (Gillott, Furniss, & Walter, 2004; Perner et al., 1989). For instance, in the study conducted by Gillott et al. (2004), speech therapists diagnosed the children prior to their participation in the study. Only children with SLI with phonologic-syntactic disabilities were eligible; however, information was missing on how this was established (i.e., inclusion or exclusion criteria) using standardized measures. Further, nonverbal intelligence and language comprehension were not assessed by the authors.
Another problem might be that children were diagnosed as having SLI when they were admitted to a special school for children with communication disorders. It was not clear how much time had passed from initial diagnosis to study participation (Colle et al., 2007; Perner et al., 1989) and thus, if children still met criteria for SLI. Moreover, in the study conducted by Colle et al., (2007), the only language testing that was done prior to study participation was the MacArthur Communicative Development Inventories (Fenson et al., 1993), completed by the parents. The MacArthur is usually used to assess the communicative abilities of infants and young children between the ages of 8 to 30 months and thus, may not be suitable to assess school-age children. Children were included in the study when their receptive and expressive language skills were equivalent to or less than those of children with typical development at two years of age. This considerable gap of approximately six years between chronological age and language ability raises the question whether these children may have had other underlying conditions, such as motor speech or cognitive problems. Therefore, it is difficult to interpret the result that children with language impairment passed 70% of the false belief tasks despite their very poor receptive and expressive language skills (Colle et al., 2007). Colle et al. concluded that their finding supports the theoretical assumption that language and ToM function independently of each other.

Finally, Perner’s et al. (1989) study did not include a control group of children with TD. Children with SLI were matched for mental age to children with autism and both groups participated in two false belief tasks (i.e., unexpected contents, change-of-location). Children with SLI performed at ceiling; however, their mental age was 6;6 years and it would have interesting to see their performance on a second-order false belief task that might have been more age-appropriate, as first-order tasks are passed by children with TD at around 4;6 years. First order false belief tasks are about the belief of another person (e.g., Susie thinks X) whereas second-order belief tasks are about what a person thinks about another person’s thoughts (e.g., Susie thinks that Mom thinks X) and are, therefore, more complicated.

However, there are also findings that indicate that school-age children with language impairments might have problems in ToM understanding (Farmer, 2000; Gillott et al., 2004; Norbury, 2005). Norbury (2005) tested children with language impairments and autism and compared them to children with TD matched for age and non-verbal ability. All children were assessed with two false belief tasks that tapped first-order as well as second-order beliefs. The
children with language impairment were on average 11;9 years old and 43% passed both first- and second-order belief tasks; however, 46% of them passed only the first-order false belief task. Children with TD performed at ceiling. This may indicate that children with language impairments experienced a delay in ToM development.

Gillott et al. (2004) tested school-age children with SLI, autism, and typical development children using the Strange Stories (Happé, 1994). These short stories tap different motivations for everyday behaviors that are not literally true, such as joking, sarcasm, and white lies. Participants were asked to explain why a story character had said what he said. These answers were coded as either referring to physical states or mental states either correctly or incorrectly. There were no group differences in referring to physical states; however, groups differed in referring to mental states. Children with autism and SLI differed significantly from their peers with TD in their use of correct mental state attributions but children with SLI and TD did not differ in their use of incorrect mental state attributions (Gillott et al., 2004). This may indicate that children with SLI were limited by their linguistic abilities to express themselves more so than by knowledge about mental states per se.

Finally, Farmer (2000) tested children with SLI from two different educational backgrounds (segregated special school or integrated language unit) and compared them to two groups of children with TD, matched for chronological or mental age. Children were given two second-order false belief tasks and six of the Strange Stories (Happé, 1994) to assess their ToM understanding. Children with SLI in the special school group performed significantly worse on all ToM measures than both groups of children with TD. Interestingly, the language skills of the group from the language unit were lower than those from the group attending special schools. However, despite having lower language skills, children from the language unit group did not differ significantly in ToM performance from both groups of children with TD. Interpretation of these results, however, is difficult because of small sample sizes (i.e., only eight children per group). The small sample sizes may also explain why this study did not find a significant association between language abilities and ToM performance, whereas others did (Norbury, 2004, 2005). Another problem in interpreting the performance of the special school group is that 6 out of 8 children had elevated scores on two standardized measures assessing social behaviors, indicating that clinical intervention was necessary. This raises the possibility that behavioral problems may have led to poor ToM performance.
1.2.2 Theory of Mind in Preschool Children with Specific Language Impairment

So far, there are four published studies that have investigated ToM in preschool children with SLI (Farrant, Fletcher, & Maybery, 2006; Farrar et al., 2009; Miller, 2001, 2004). Both Farrar et al. (2009) and Miller (2001, 2004) examined possible relationships between language and ToM. However, Miller’s focus was on the effects of linguistic demands in solving change-of-location tasks, whereas Farrar et al. explored the relations of general and specific language aspects with ToM performance in children with SLI.

Miller (2001, 2004) assessed children with SLI on false belief tasks and compared their performances with those of age- and language-matched children with TD. In Miller (2001), the child had to predict the behavior of a returning puppet after an object was moved and the four experimental conditions were “think”, “look”, “show”, and “pretend”. Thus, the child was asked “Where does puppet think the toy is”, “Where will puppet look for the toy”, or “Show me what puppet will do”. In the last condition, the child and puppet pretended that a block was, for instance, a racecar, and, while the puppet was gone, the experimenter and the child pretended that the block was a drum. Then the child was asked “What does puppet think we’re pretending the block is?” The “think” and “pretend” conditions were considered to be linguistically more challenging, whereas the “look” and “show” conditions were thought to be less demanding (Miller, 2001). The group of children matched for language comprehension skills did not benefit from lower linguistic demands and performed at chance levels in all four conditions. Children with SLI performed comparably to the age-matched children in the “look” and “show” conditions but performed significantly worse than their peers in the “think” and “pretend” conditions. Thus, Miller (2001) concluded that higher linguistic complexity had a negative effect on ToM performance in children with SLI.

A subsequent study used the same conditions of “show”, “think”, and “pretend” but the “look” condition was replaced by a “less verbal” condition (Miller, 2004). In the “less verbal” condition, the child was presented with a brief, silent video and was asked to point to a picture that would complete the change-of-location story in the video (“What comes next?”). This task proved to be too difficult for all children, who performed at chance levels. Similar to the previous study, the group of children matched for language comprehension skills performed at chance levels in the other three conditions. However, in this study, there were no group
differences in the “pretend” condition. Moreover, children with SLI performed comparably to the age-matched children in the conditions of “show” and “think” (Miller, 2004). It is not clear why children with SLI showed inconsistent performances in the “think” condition because the wording of the question was exactly the same across the two studies. The difference was unlikely to be due to age, as the children who performed better in the more demanding “think” condition were on average 4;11 years old and thus younger than in the previous study (i.e., $M = 5;6$ years). A possible explanation for this inconsistency could be that children with SLI had better language abilities in the second study. However, language scores were not provided in the first study; therefore, it was not possible to compare the language skills across studies. It might also be that factors related to individual differences in children with SLI, such as familial SES or presence of siblings could account for the better ToM performance. Thus, these two studies provided inconsistent results on whether linguistic demands were related to false belief task performance in children with SLI.

Farrar et al. (2009) attempted to answer whether general or specific language abilities were predictive of false belief performance in children with SLI. They tested preschool children with SLI on different language measures (e.g., expressive grammar, receptive vocabulary, memory for complements) and a variety of common false belief tasks. Memory for complements did not make a significant contribution to ToM performance once grammar and vocabulary were accounted for in the regression analyses. Thus, general language abilities (e.g., grammar, vocabulary) rather than specific language abilities (e.g., sentential complements) seemed to be important for ToM development in children with SLI. This finding supports the result of a previous study conducted by Miller (2004), in which children with SLI tended to perform similarly to an age-matched group on false belief tasks but their performance on sentential complements was similar to that of language-matched children. Thus, both studies provided evidence against J. de Villiers’ (2005) claim that mastery of sentential complements is crucial for false belief performance. However, this result may be specific to children with SLI and it may not be possible to extend these findings to children with TD as this study did not include a control group of children with TD.

On the other hand, Farrant et al. (2006) investigated children’s understanding of different ToM concepts. The authors used the ToM scale of Wellman and Liu (2004), which includes assessment of desires, beliefs, false beliefs, and emotions and is ordered from the easiest to the
hardest tasks. The ToM scale included control questions and except for the last task, children with SLI did not differ in answering them from children with TD. However, the total ToM score of children with SLI was significantly lower than that of the comparison group. Children with SLI acquired earlier ToM concepts (i.e., desires and beliefs) at a level comparable to their peers but had significantly more difficulties in judging whether a person has access to certain knowledge or not (e.g., does a person know what is in a drawer when he has never seen it before?). Furthermore, children with SLI performed significantly worse on the unexpected contents task than their peers. The last task of the scale required the children to identify why a person might feel in one way but display a different emotion. Many children with SLI failed the control questions (e.g., “Why did he try to look [sad, okay, happy]?”) and thus, this task was omitted from further analysis (Farrant et al., 2006).

The format of the knowledge access task resembled the format of the contents false belief task. Perhaps children with SLI did not understand the concept of false belief and the similarity of the knowledge access and the false belief task prevented a better performance on the knowledge access task. The notion that children with SLI may not have understood the concept of false belief is supported by the finding that they answered the control questions correctly but failed the test question (“Where will he look for x?”). Thus, the authors argued that children with SLI were delayed in their ToM development.

1.2.3 Summary

Table 2 provides a summary of the most important studies that were relevant to this thesis and that compared ToM performance in children with and without language impairments. The top panel of Table 2 summarizes studies that investigated the ToM performance of school-age children, whereas the bottom panel summarizes studies conducted with preschool children. It seems that school-age children with language disorders perform relatively well on first-order false belief tasks. However, their level of ToM development may not be comparable to that of peers with TD. Children with SLI may follow a similar trajectory of ToM development but on a different, later time line than children with TD. This is, in part, supported by studies that showed that preschool children with SLI displayed problems in solving ToM tasks (Farrant et al., 2006). The delay might be explained by their deficient language skills. However, evidence with regard to this explanation is mixed (Miller, 2001, 2004).
As can be seen in Table 2, few studies investigated ToM understanding in preschool children with SLI and, as the review of the literature suggests, results were inconclusive. The current study may help to provide further evidence about ToM understanding in children with SLI by applying the ToM scale, using a low verbal input. Further, problems in understanding ToM may be associated with problems that children with SLI experience in social situations, such as in pretend play.

Table 2

<table>
<thead>
<tr>
<th>Studies</th>
<th>Participants</th>
<th>Measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>School-age children</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farmer, 2000</td>
<td>( n = 8 ) with SLI in special school (( M_{age} = 11;0 )), [SLI/SS]</td>
<td>1) second-order false belief</td>
<td>1) (SLI/SS &lt; TD/L &lt; TD/A) = SLI/IS</td>
</tr>
<tr>
<td></td>
<td>( n = 8 ) with SLI in integrated setting (( M_{age} = 10;8 )), [SLI/IS]</td>
<td>2) correct mental state attribution (Strange Stories)</td>
<td>2) (SLI/SS = TD/L) &lt; SLI/IS &lt; TD/A</td>
</tr>
<tr>
<td></td>
<td>( n = 8 ) with TD age-matched (( M_{age} = 10;3 ), [TD/A]</td>
<td>3) incorrect mental state attribution (Strange Stories)</td>
<td>3) SLI/SS &lt; (SLI/IS = TD/L) &lt; TD/A</td>
</tr>
<tr>
<td>Gillott et al., 2004*</td>
<td>( n = 15 ) with SLI, ( M_{age} ) across all groups 10;3</td>
<td>1) correct mental state attribution (Strange Stories)</td>
<td>1) SLI &lt; TD/A</td>
</tr>
<tr>
<td></td>
<td>( n = 15 ) with TD/A, ( M_{age} ) across all groups 10;3</td>
<td>2) incorrect mental state attribution (Strange Stories)</td>
<td>2) SLI = TD/A</td>
</tr>
<tr>
<td>Studies</td>
<td>Participants</td>
<td>Measures</td>
<td>Results</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td><strong>School-age children</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ziatas &amp; Durkin, 1998</td>
<td>a) $n = 12$ with SLI ($M = 6;1$)</td>
<td>1) first-order false belief</td>
<td>1) a) SLI = TD/L</td>
</tr>
<tr>
<td></td>
<td>b) $n = 12$ with TD/L ($M = 6;1$) AND</td>
<td></td>
<td>1) b) SLI = TD/A</td>
</tr>
<tr>
<td></td>
<td>$n = 12$ with SLI ($M = 6;11$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b) $n = 12$ with TD/L ($M = 6;5$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Preschool children</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farrant et al., 2006</td>
<td>$n = 20$ with SLI ($M_{\text{age}} = 5;2$)</td>
<td>1) ToM scale</td>
<td>1) SLI &lt; TD/A</td>
</tr>
<tr>
<td></td>
<td>$n = 20$ with TD/A ($M_{\text{age}} = 5;1$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miller, 2001</td>
<td>$n = 10$ with SLI ($M_{\text{age}} = 5;6$)</td>
<td>1) “look”</td>
<td>1) + 2) TD/L &lt; (SLI = TD/A)</td>
</tr>
<tr>
<td></td>
<td>$n = 10$ with TD/A ($M_{\text{age}} = 5;6$)</td>
<td>2) “show”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$n = 9$ with TD/L ($M_{\text{age}} = 3;9$)</td>
<td>3) “think”</td>
<td>3) + 4) (TD/L = SLI) &lt; TD/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4) “pretend”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(first-order false belief)</td>
<td></td>
</tr>
</tbody>
</table>
Studies | Participants | Measures | Results
---|---|---|---
Miller, 2004 | \(n = 15\) with SLI (\(M_{\text{age}} = 4;11\)) | 1) “show” | 1) SLI = (TD/A = TD/L) |
| | \(n = 15\) with TD/A (\(M_{\text{age}} = 5;0\)) | 2) “think” | 2) SLI = (TD/A > TD/L) |
| | \(n = 15\) with TD/L (\(M_{\text{age}} = 3;8\)) | 3) “pretend” | 3) + 4) SLI = TD/A = TD/L |
| | | 4) “less verbal” (first-order false belief) | |

Note. * Studies also included groups of children with autism.

1.3 Pretend Play

1.3.1 What is Pretend Play?

Many different terms have been used in the literature to refer to pretend play. Some of these terms are *symbolic* or *fantasy play*, *make-believe*, or *imaginative play* (Fein, 1981). Although the terms vary, they have in common that a child behaves in a nonliteral way, acting ‘as if’ inanimate objects were alive or had absent properties. Criteria to identify pretend play are (a) inanimate objects are active agents (e.g., doll getting ready for birthday party), (b) the child performs actions usually done by others (e.g., cooking, fixing) or (c) actions may not have a typical outcome (e.g., preparing dinner but not having dinner). Further, the child (d) substitutes an object to represent another object (e.g., box as a car), (e) performs a symbolic action to attribute properties (e.g., hunger), or (f) represents absent objects (e.g., taking off a non-existing hat) (McCune-Nicolich, 1981; Stagnitti, Unsworth, & Rodger, 2000). In this study the term *pretend play* is used. Further, the terms *solitary* or *social pretend play* are used, respectively, to describe that the child is engaging in pretend play either alone or with others.
1.3.2 Development of Pretend Play

In the last century, play has been a topic of interest and has been investigated from different theoretical perspectives and disciplines (e.g., psychology, pedagogy). The following is a brief overview of pretend play development highlighting the most important developmental changes between the ages of 2 to 6. During this period, pretend play becomes increasingly sophisticated on various levels, such as object play, enacted play themes, sequencing of play actions, play with others, and role play (Stagnitti, 1998). Around the age of six, obvious forms of pretend play behaviors decrease and other forms of play (e.g., games with rules) become more important (Fein, 1981). An overview is proved in Table 3 of milestones in the development of play in children with TD.

Beginning in infancy, pretend play skills develop gradually during interactions with caregivers and later on with siblings and peers (Fiese, 1990; Youngblade & Dunn, 1995). Toward the end of the second year, the first pretend play behaviors emerge, which are simple and brief, such as the child “drinking” from a cup or “feeding” a doll (McCune-Nicolich, 1981). By 24 months, most children are able to use a realistic looking object to stand for something else (e.g., using a toy telephone to “call” someone). Howes and Matheson (1992) observed more complex pretend play behaviors around 30 months (e.g., doll is “hungry”, “drinks” from cup and “eats”). Even though the ability to pretend with substitute objects improves greatly during preschool years, most three-year olds prefer replica objects in their pretend play (McCune-Nicolich, 1981). That is, they use miniature toy versions of the real objects, such as a toy car representing a real car. By four years, children can use less realistic objects for the pretend object and can even represent absent objects (Casby, 1987). The ability to represent absent objects is mastered in the early school years, for instance, by pretending to have a “pistol” by using index finger and thumb (Fein, 1981).
<table>
<thead>
<tr>
<th>Age</th>
<th>Developmental Milestones</th>
</tr>
</thead>
</table>
| 13 – 24 months | Play is object-centered, objects are explored and manipulated  
               Shows simple pretend behaviors, e.g., combing hair, “drinking” from a cup, “feeding” a doll or toy animal  
               Uses replica objects (i.e., object substitution)  
               Uses inanimate toys (e.g., dolls, toy animals) in simple play actions  
               Engages in solitary play                                                                                                                                   |
| 25 – 36 months | Plays with ambiguous objects but prefers replica objects  
               Pretend play behaviors are more complex, play actions more detailed and organized in a logical sequence  
               Plays with dolls and toy animals as active agents  
               Play themes are domestic themes and those of personal experience (e.g., shopping, birthday)  
               Engages in parallel play  
               Shows an awareness of social roles                                                                                                                       |
| 37 – 48 months | Pretend play is very complex  
               Play actions take place in space and time  
               Story plots are planned, may be more than one plot  
               Engages in social pretend play  
               Enacts roles of familiar or fictional characters  
               Begins to engage in role play with other children, may have difficulty maintaining role play  
               Begins to use meta-communication about play                                                                                                             |
<table>
<thead>
<tr>
<th>Age</th>
<th>Developmental Milestones</th>
</tr>
</thead>
<tbody>
<tr>
<td>(cont.)</td>
<td></td>
</tr>
<tr>
<td>37 – 48 months</td>
<td>Shows gender differences in play</td>
</tr>
<tr>
<td>49 – 60 months</td>
<td>Takes extensive time for ‘setting up the play scene’</td>
</tr>
<tr>
<td></td>
<td>Play themes go beyond personal experience (e.g., dragon hunt)</td>
</tr>
<tr>
<td></td>
<td>Story plots are complex, organized, and elaborate with sub-plots</td>
</tr>
<tr>
<td></td>
<td>Engages in elaborate role play in social pretend play; roles relate to each other and are compatible with play theme</td>
</tr>
<tr>
<td>61 – 72 months</td>
<td>Play themes are elaborate and story plots are complex</td>
</tr>
<tr>
<td></td>
<td>Cooperates with others in ‘setting up the play scene’</td>
</tr>
<tr>
<td></td>
<td>Play may be influenced by hobbies</td>
</tr>
<tr>
<td></td>
<td>Plays games with rules (e.g., board games)</td>
</tr>
<tr>
<td></td>
<td>Shows a decline in overt pretend play behaviors</td>
</tr>
</tbody>
</table>

*Note. Adapted from Pronin Fromberg & Bergen (2006); Stagnitti (1998).*
Around two years, toddlers also begin to play with inanimate objects, such as dolls or stuffed animals (Stagnitti, 1998). At first, pretend play is simple and may only include one action, such as placing a doll in a chair or bed. However, during the third year more play actions can be performed and play actions become more detailed and organized in a logical sequence (e.g., parking a car, driving the car, having a crash) (Garner & Bergen, 2006). Further, the child increasingly starts to treat inanimate toys as active agents; the doll or hero-figurine is made to act on its own (Fein, 1981). Play themes become more elaborate and typically revolve around familiar home-like themes (e.g., fixing the car, shopping) and personal experiences (e.g., birthday, sickness) (Garner & Bergen, 2006; Lyttinen, 1991). Older preschool children are able to enact play themes that they have never personally experienced, such as ghosts, dragon hunt, or pirates (Stagnitti, 1998).

As the child becomes more proficient in pretending with objects and toy figurines and engages in more imaginative play themes, the child also becomes more able to coordinate pretend play with other children. At around 3 years of age, children start to show a shift from solitary to simple social pretend play (Fein, 1981; Howes & Matheson, 1992). That is, a child engages in a similar play activity beside another child and begins to acknowledge the peer by smiling or offering toys. However, there is still little coordination or negotiation.

Three-year-olds also become increasingly aware of social roles and role play becomes a part of their social pretend play as well (Howes, Unger, & Seidner, 1989). Role play or role enactment is defined as “behavior in which the child simulates the identity or characteristics of another person” (Fein, 1981, p. 1101). These social roles may either be complementary (e.g., mother-child, doctor-patient) or behavioral. That is, the play behaviors for the roles do not depend on a particular partner but derive their meaning from the play theme (e.g., pirates on a treasure hunt) (Andresen, 2002; Fein, 1981). For three-year-olds, role play is a challenging social activity because children have to choose roles that relate to one another and are compatible with the chosen play theme. Young three-year-olds may begin a play session with complementary role play but soon the children enact their roles independent of each other (Fein, 1981). Between the ages of three and five years, children become more proficient in coordination of role play. That is, children spend more time in role play and increase their ability to sustain role play and enact diverse roles (Andresen, 2005).
1.3.3 Pretend Play and Language

Both first words and simple pretend play behaviors emerge around the second year of life. Further, children who began to combine words also began to demonstrate simple play combinations (McCune-Nicolich, 1981). However, it is of interest whether the two abilities only co-occur or whether they are related and share underlying capacities, for instance, for symbolic functioning (Casby, 1987; McCune-Nicolich, 1981). Some correlational studies support the assumption that language and pretend play share the general capacity for symbolic representation. For instance, Casby and Della Corte (1987) tested children between 19 and 32 months using an object substitution pretend play task (e.g., play ‘hat’ with similar or dissimilar looking objects). Children at the single-word utterance level and those who produced word combinations were significantly different in their ability to substitute objects. Those with single-word utterances were less able to represent a dissimilar looking object as a hat or pretend to have a hat when no object was present. The ability to perform these pretend play behaviors was significantly correlated with children’s mean length of utterance (MLU) in morphemes. The correlation between pretend play and MLU was stronger than the correlation between pretend play and chronological age. Moreover, the relation between pretend play and MLU remained significant independent of age (r = .73). Thus, the authors concluded that there is an association between early language development (early word combinations/syntax) and pretend play.

Casby and McCormack (1985) tested children between three and six years who had severe hearing loss. Parents, teachers, and clinicians were asked to rate the children’s communication levels (including signs, gestures, and words) and children were then grouped as having either low or high communication levels. The children were introduced to familiar pretend play themes (e.g., race cars, feed baby) and were then asked to perform pretend play with the same play themes but a conventional toy (e.g., car) was replaced by an unconventional object (e.g., block). Children with low communication levels did significantly more poorly on the pretend play task than children with higher communication levels. Further, all children in the low communication group were only able to perform the pretend actions after the experimenter modeled them. Casby and McCormack (1985) found a significant correlation between communication level and pretend play independent of chronological age (r = .77). A limitation of this study was that the authors did not use a standardized measure to assess the language abilities of the children, which are therefore difficult to estimate. Further, they did not include a control group of children with
TD, either matched for chronological or mental age. However, given the development of play as described earlier, it seems likely that children with TD would have performed well on this task.

Most studies that investigated the relationship between play and language were conducted with young toddlers. However, Lyytinen (1990) observed the relationship continuing into the preschool years. She conducted a cross-sectional study with children between 2 and 6 years of age and found significant correlations between language and pretend play, independent of age. Lyytinen noticed that as children were playing by themselves, three-year-olds relied heavily on language to talk about imagined objects in pretend play, whereas five-year-olds were playing more silently. Further, in each age group, there were some children who were engaged in pretend play but rarely verbalized their actions. It has been suggested that this may reflect personal style (Fein, 1981; Lyytinen, 1990) or indicate that the role of language during pretend play changes during the preschool period (Ungerer, Zelazo, Kearsley, & O’Leary, 1981). This may make it more difficult to detect an association between pretend play and language in preschoolers, as some researchers did not find a link (Astington & Jenkins, 1995; Schwebel, Rosen, & Singer, 1999; Youngblade & Dunn, 1995). It may well be that language and pretend play are only temporarily related in young toddlers. Nevertheless, language becomes increasingly important when children start to play with each other and engage in role play (Andresen, 2002; Farver, 1992).

Farver (1992) investigated the communication strategies that children between the ages of two and five years used to establish and maintain social pretend play. She found that two-year-olds showed short and simple social pretend play episodes with a peer. The children initiated play by calling for attention (“Hey!”, “Look!”) and relied on paralinguistic cues (e.g., sounds of toy cars) to indicate their intentions to play. With increasing language abilities, children then verbalized actions or behaviors that indicated pretend play behaviors, which were increasingly independent of the context or the availability of objects (McCune-Nicolich, 1981). These verbalization abilities were important in social pretend play to assign roles, act them out, or negotiate play scripts (Farver, 1992). Children between 4 and 5 years of age engaged in long play episodes and used many communicative strategies to organize their complex social pretend play. They described their actions, directed peers’ play, and used conversational tags to elicit a response or acknowledgment (e.g., “right?”,” “OK?”). Farver pointed out that the developmental changes in communicative strategies were most apparent in the group of 3-year-olds. They
began to use their voices (high or low pitch, intonation) to animate objects and to mark role enactment. They also started to add new information to a playmate’s previous utterance instead of just repeating it. This highlights the importance of language during social pretend play.

Moreover, children used a ‘special language’ to mark their play behaviors in social pretend play as ‘unreal’ (Andresen, 2005; Garvey & Kramer, 1989). Garvey and Kramer (1989) investigated social pretend play in child-dyads between 3 and 6 years and found that talk within social pretend play was characterized by a higher frequency of certain linguistic features than talk outside of pretend play. Talk inside and outside a pretend play frame was coded for linguistic features such as sentence complexity, temporal expressions, as well as future, past tense and modal verb forms. Garvey and Kramer (1989) found that within social pretend play, children’s sentence structure was more complex with higher proportions of future auxiliary and modal verb forms (“you are going to”, “you have to”) as well as higher proportions of temporal expressions (“soon”, “before”, “in the morning”) than outside of pretend play. These findings were supported by Andresen (2005), who reported that German-speaking children used phrases such as “pretend”, “for fun” (German “wohl”, “aus Spass”, “im Spiel”) or statements such as “it’s only play”, “it’s not real” (German “das ist nur Spiel”, “das ist nicht in echt”) to mark the boundary between pretend play and reality.

1.3.4 Pretend Play and Narrative Competence

Language may also be important for pretend play because language provides the means for narrative competence, which, in turn, scaffolds pretend play (Nicolopoulou, 2005; Pellegrini, 1985). Narrative competence refers to the ability to comprehend and tell stories (i.e., narratives), and it also involves an understanding that the actions of a story character are temporally and causally motivated (Pellegrini, 1985). Narrative competence is further viewed to be part of the broader category of literate behavior, that is, the ability to use and comprehend decontextualized language in general (Pellegrini, 1985). Using decontextualized language means that the child has to use language itself (either oral language or text) to convey meaning and to resolve ambiguity. In social pretend play, a child cannot rely on context or shared knowledge as objects may represent something different or play themes may change. Thus, the child has to use language to explain the new ideas and meanings (e.g., “This is the food. He gives it to the animals because they are hungry”) because ambiguity may otherwise cause a disruption in the
social pretend play sequence. In this sense, pretend play is narrative-like (Pellegrini, 1985) or like an enacted narrative (Nicolopoulou, 2005).

Pellegrini (1982) found that preschoolers varied their use of explicit language depending on the play context. For instance, children provided more cohesive information (e.g., linking play events to create meaning) while playing in a housekeeping corner than while playing with blocks. In the latter, children were more likely to rely on contextual cues, such as “give me this”, where “this” was defined by a gesture and did not convey meaning by itself.

There are a few longitudinal (e.g., Pellegrini, 1980) and experimental studies (e.g., Baumer, Ferholt, & Lecusay, 2005; Pellegrini & Galda, 1982) that investigated the relationship between pretend play and narrative competence (Pellegrini, 1985). Baumer et al. (2005) conducted an intervention study, and they showed that children who enacted stories in role play activities were significantly better in organizing pictures of an unfamiliar story and telling the story than children who did not participate in the role play activity. Similar results were reported by Pellegrini and Galda (1982), where only children in the play condition (but not the discussion or drawing conditions) demonstrated better story comprehension and story telling.

Narrative skills emerge during preschool and continue to develop into later school age (Engel, 2005; Ilgaz & Aksu-Koç, 2005). Pretend play, specifically role play, may facilitate narrative competence in school-age children (Baumer et al., 2005; Pellegrini & Galda, 1982). However, a limitation in the experimental studies might be that children required some form of narrative competence (i.e., story comprehension) to begin with. That is, children were read to (Pellegrini & Galda, 1982) or watched the researchers enacting stories (Baumer et al., 2005) before they were asked to engage in pretend play. Further, children’s narrative abilities were assessed pre- and post-test but pretend play abilities were not assessed. Thus, narrative skills may have influenced children’s pretend play skills as well and the relationship may be reciprocal and not unidirectional.

1.3.5 Summary
There is some evidence that language and pretend play are temporarily related during toddlerhood. However, the relationship seems to undergo changes during the preschool years. Nevertheless, as children become more interested in playing with other children, language again becomes an important tool that allows children to engage in social pretend play. Language helps
children to establish, structure, and maintain social pretend play. Further, play-language ‘markers’ (e.g., “for fun”, “it’s not real”) help children to identify when a playmate is still playing or is not any longer, which also facilitates social pretend play. Social pretend play might also provide a facilitative context for the development of narrative skills.

1.3.6 Pretend Play and Individual Differences in Socioeconomic Status

Given that language is related to pretend play and SES, it is conceivable that pretend play is also associated with SES. Studies that investigated the influence of family SES on pretend play abilities in children with TD provided inconclusive evidence (McLoyd, 1982). For instance, Rubin, Maioni, and Hornung (1976) found significant differences in children’s pretend play abilities favoring children from higher SES backgrounds. That is, children from lower class backgrounds were more likely to engage in functional play and to play alone or engage in parallel play than children from middle class backgrounds. The authors proposed that children with TD from lower-income classes might have different access to toys or limited space or time to engage in play, which might explain their lower levels of sophisticated pretend play. Others, however, did not find differences in pretend play behaviors in children from different social classes (Fein & Stork, 1981; McLoyd, 1982). In her systematic review, McLoyd (1982) suggested that pretend play differences might also be due to cultural differences or due to methodological weaknesses. That is, children from different cultural backgrounds may differ in their play ideas when they play “going shopping” based on their different experiences. Similarly, children from lower SES backgrounds might be at disadvantage when asked to play “house” when these children came predominantly from single parent families whereas children from higher SES backgrounds came from intact families. Thus, other factors might also be involved and the influence of family SES on children’s pretend play remains unclear.

1.4 Pretend Play in Children with Specific Language Impairment

Acquisition of language emerges parallel to pretend play in children with TD and language and play are related at some stages in their development (Casby & Della Corte, 1987; Fein, 1981; McCune-Nicolich, 1981). Studying pretend play in children with SLI may be informative because it allows investigation of the development of pretend play when language development is delayed. If both language and play are delayed to a similar extent, one could argue that they depend on a general representational capacity. If, however, pretend play develops on a similar
trajectory to that seen in children with TD but language does not, then play and language are not strongly related (Roth & Clark, 1987). Studies that investigated play in children with SLI have provided mixed evidence.

Rescorla and Goossens (1992) investigated play behaviors in toddlers with and without SLI matched for age, gender, and SES. They reported significant differences in the quality and type of play. Toddlers with SLI were more likely to manipulate and group objects than their peers with TD during sessions of solitary free play. They also tended to spend more time in functional play (i.e., relate two objects in a conventional manner) and less time in advanced pretend play than their peers. That is, children with SLI were less likely to substitute objects, pretend with absent objects, or animate a doll or toy animal than their peers. Further, both groups showed better pretend play performance when playing with an experimenter. However, most children with SLI needed to observe the experimenter modeling pretend play actions before they performed them, whereas children with TD performed them spontaneously or with verbal instructions (Rescorla & Goossens, 1992). Effect sizes were medium to large ($d = .56$ to $1.10$). Thus, children with SLI were not only delayed in their language abilities but also in their development of pretend play relative to age-matched peers.

Others reported no statistical differences between 30-month-old children with and without language impairment on different play measures (Lombardino, Stein, Kricos, & Wolf, 1986). However, this may have been due to small sample sizes ($n = 5$ in each group), which resulted in a lack of statistical power. In general, children with SLI tend to perform worse on most measures when compared to age-matched controls (Leonard, 1998). Therefore, in some pretend play studies, children with SLI were also matched to control children for their linguistic abilities (Roth & Clark, 1987; Terrell, Schwartz, Prelock, & Messick, 1984). Thereby, it was possible to see whether children with SLI followed the normal developmental pattern but on a delayed timeline due to their deficient language skills. Terrell et al. (1984) matched children with language impairment ($M_{age} = 2;11$) for expressive vocabulary to children with TD ($M_{age} = 1;7$), with both groups producing no word combinations. The children’s pretend play abilities were tested with the Symbolic Play Test (SPT) (Lowe & Costello, 1976) in which children are presented with different sets of toys and are observed for their engagement with the toys. Children with language impairment performed significantly better on the SPT than the language-matched group ($d = 1.69$). Interestingly, although it had been reported that the
emergence of word combinations is related to simple play combinations (McCune-Nicolich, 1981), children with SLI in Terrell et al.’s study combined play actions even though they were still in the one-word stage. Thus, their language abilities were more delayed than their pretend play skills. One limitation of this study was that children with SLI were older and may have had more experience with different play materials than the younger control group. The inclusion of an age-matched control group would therefore have been interesting.

On the other hand, Roth and Clark (1987), who also matched children with and without SLI for their language abilities, reported different results. The SLI group was on average 6;7 years old and the TD group was on average 2;9 years. The children were observed either playing alone with toy sets from the SPT, interacting with an experimenter, or playing with another child (dyads were formed within each language group). In contrast to Terrell et al.’s finding, children with SLI performed significantly more poorly on the SPT than the language-matched control children ($d = 1.33$). That is, children with SLI did not sequence their pretend play behaviors (e.g., put a doll to bed and cover it with a blanket). They were also less able to perform and structure pretend play actions around a theme as requested by the experimenter than the children with TD (e.g., “Show me how Mommy drives the car” or “Let’s make a birthday party for the doll. What do we need?”) ($d = 1.53$). Typically for their age, children with TD engaged in solitary and parallel play when play was observed in child-child dyads. Children with SLI however, spent significantly more time in nonplay behaviors than children with TD ($d = 1.94$). Despite being considerably older, children with SLI displayed even less sophisticated pretend play than the language-matched group on all play measures. A weakness of the study was that some of the instructions were provided verbally and children with SLI had problems with language comprehension, as measured with Test of Auditory Comprehension of Language (Carrow, 1973). This may have put them at a disadvantage as they may not have fully understood the experimenter’s instructions. However, the SPT does not involve language. It might be that the toy sets provided were less attractive for the older children with SLI, given their age range from 5 to 8 years compared to the younger children with TD (2;8 to 3;2 years).

Social pretend play in preschoolers with SLI has received little attention in the literature and may not have been the central focus of investigation when investigating social interactions. Some studies investigated social behaviors, such as play on the playground (e.g., Fujiki, Brinton, Isaacson, & Summers, 2001). Fujiki et al. (2001) found that school-age children with
SLI displayed more solitary play and withdrawn behaviors on the playground than their age- and gender-matched peers with TD (e.g., watching others without joining, being unoccupied). Others have investigated the ability of school children with SLI to access ongoing social interactions (Liiva & Cleave, 2005). Children with SLI tended to be passive and were more likely to be invited to play by their unfamiliar play partners instead of initiating play themselves. One of the few studies that investigated social pretend play in preschool children with language impairment was conducted by DeKroon, Kyte, and Johnson (2002). Three children with language impairment (LI) were observed when playing with familiar peers either with or without LI. The quality and quantity of social pretend play varied with the different play partners. In LI-dyads, children were often content to engage in solitary or parallel play. When they started to interact with each other, they showed a limited variety of play themes. Difficulties in initiating and maintaining role play sequences often resulted in abandoning the play ideas and engaging in solitary play. Children with LI engaged in more play themes in mixed LI-TD dyads. However, these themes were most often initiated by the peer with TD. Moreover, children with LI relied on the peers to maintain, structure, and expand the play themes. Thus, children with LI were less able to initiate and maintain social pretend play than their peers, which may be due to their deficient language skills. However, language did not seem to be a strong predictor of social pretend play (DeKroon et al., 2002), which is consistent with previous studies (Leonard, 1998).

Family SES has received little attention in studies investigating pretend play performances in children with SLI. It was proposed that SES might be related to children’s pretend play abilities because children might have different access to toys or limited space or time to engage in play depending on their social class background (Ramsey, 2006; Rubin et al., 1976). However, the majority of studies on pretend play in children with SLI did not provide information on SES (e.g., Fujiki et al., 2001; Roth & Clark, 1987; Terrell et al., 1984). Thus, it is not clear if SES might have contributed to the pretend play differences in children with and without language impairments.

To summarize, children with SLI tend to display less sophisticated pretend play than their peers with TD. However, when children with SLI are matched to children with similar language levels, the evidence is mixed. Some results are in favor of children with SLI and others in favor
of the language-matched group (Leonard, 1998). Further, there is some evidence that children with SLI differ in their social play behaviors from their peers with TD. It might be that children with SLI are limited in their possibilities for participation due to their deficient language skills. Thus, language deficits might limit the number of possibilities children with SLI have to express play intentions and play scripts (Casby, 1997). This is likely to restrict their ability to participate in pretend play scripts with peers which, in turn, might put the further development of pretend play skills at risk (Rescorla & Goossens, 1992). Limited pretend play skills are, furthermore, likely to negatively affect social interactions, friendships, and learning experiences (Stagnitti & Unsworth, 2000). However, language skills do not necessarily predict play performance; i.e., language scores can be relatively high and yet children with SLI may experience unsuccessful play interactions (DeKroon et al., 2002; Leonard, 1998). Another explanation might be that limited ToM understanding hinders children with SLI from active participation in social pretend play. However, to date there is no published research that investigated the association between ToM and pretend play in children with SLI.

1.5 Pretend Play and Theory of Mind

Some researchers state that early pretend play is a precursor to ToM (e.g., Leslie, 1987). As a child becomes able to engage in pretend play, she also becomes able to differentiate between real-world and pretend play behaviors in others (Leslie, 1987; Rakoczy & Tomasello, 2006). Other researchers however, have refuted that claim, arguing that preschool children do not attribute mental states to the playmate during pretend play (Harris, 1994; Lillard, 1993; Perner, Baker, & Hutton, 1994). Thus, there is a debate on the extent to which young children show an early awareness of pretense as a reference to mental states in others. Although it remains unclear how early pretend play may be linked to ToM, researchers have investigated whether social pretend play is associated with ToM (Astington & Jenkins, 1995; Jenkins & Astington, 2000; Nielsen & Dissanayake, 2000; Schwebel et al., 1999; Youngblade & Dunn, 1995). The assumption is that in social pretend play, children have to inform each other about play ideas, the changed meaning of objects, and the roles to be enacted. Thus, a child needs to be aware of what his playmate knows or what he needs to know in order to successfully join the social pretend play. It was hypothesized that a child who knew that he had to provide meaningful
information about the play should be a better playmate than somebody who did not (Astington & Jenkins, 1995).

Astington and Jenkins (1995) administered a series of false belief tasks to preschoolers and observed their play during free play in groups. Children’s play was coded for the presence of pretend play (e.g., object substitution, using replica objects). Pretend play was further coded for Joint Proposal (i.e., child refers to self and another child within the same speaker turn) and Role Assignment (i.e., child verbally and explicitly assigns a pretend role to themselves or to another child). There was no significant relationship between pretend play and ToM. However, Joint Proposal and Role Assignment were significantly related to ToM, independent of age and language abilities ($r = .49$ and $r = .37$). Thus, children who scored higher on false belief tasks were significantly more likely to display more sophisticated forms of pretend play than children who scored lower (Astington & Jenkins, 1995).

The hypothesis that an understanding of thoughts and beliefs of another person might be most important during pretend play was supported by other studies that reported a significant relation between false belief and pretend play skills (Nielsen & Dissanayake, 2000; Taylor & Carlson, 1997; Youngblade & Dunn, 1995). However, the studies showed inconsistent findings with respect to which forms of pretend play (i.e., solitary or social) and aspects of play (i.e., general pretend play or role play) were related to ToM. In two studies, Schwebel et al. (1999) observed preschool children playing alone and playing with other children. Only social pretend play was significantly related to ToM ($r = .36$) but in both studies, solitary pretend play was not ($r = .27$ and $r = .07$).

In contrast, Taylor and Carlson (1997) observed children’s solitary play with different sets of toys and found a weak but significant correlation between solitary play and ToM ($r = .26$). In this study children were classified as either ‘High Fantasy’ (e.g., having imaginary companion, preferring fantasy oriented toys or stories) or ‘Low Fantasy’ (e.g., having no imaginary companion, preferring reality oriented toys or stories). Children who were in the High Fantasy group were significantly more likely to demonstrate higher levels of pretend play and a better ToM understanding than children who were in the Low Fantasy group. Thus, different findings with respect to solitary play might be explained by individual differences in children’s ability to engage in fantasy activities (Taylor & Carlson, 1997).
Three studies applied a similar coding scheme to analyze social pretend play (Astington & Jenkins, 1995; Nielsen & Dissanayake, 2000; Youngblade & Dunn, 1995). Social pretend play was coded for general pretend play (e.g., object substitution) and for different aspects of role play (e.g., assigning roles, enacting roles). However, studies differed with respect to what aspects of social pretend play were associated with ToM. For instance, in one study general pretend play was significantly related to ToM \( (r = .35) \) (Nielsen & Dissanayake, 2000), whereas others did not find a significant association (Astington & Jenkins, 1995; Youngblade & Dunn, 1995). Two studies found that the child’s ability in role assignment was significantly related to ToM performance, independent of age and language abilities; however, Youngblade and Dunn (1995) did not report this relation.

Further, two longitudinal studies reported different predictors of the pretend play – ToM relationship at different ages in children’s development (Jenkins & Astington, 2000; Youngblade & Dunn, 1995). Jenkins and Astington (2000) tested preschool children three times over a period of seven months, i.e., at ages 40, 43.5, and 47 months. As in previous studies, children were given ToM tasks and videotaped while playing with a peer. Earlier ToM performance predicted later social pretend play abilities (i.e., role assignment), which suggested that ToM may facilitate role play abilities. On the other hand, Youngblade and Dunn (1995) tested children when they were 33 and 40 months old. They found that early social pretend play abilities (i.e., role enactment) predicted later ToM performance. These findings suggest that the relationship between ToM and social pretend play is not unidirectional but reciprocal. That is, in young preschoolers social pretend play promotes ToM development, whereas in older preschoolers ToM development promotes social pretend play. Thus, the direction of effect seems to change over time during preschool children’s development.

Across these studies, children’s social pretend play was observed with different play partners in different environments. Children in Youngblade and Dunn’s study (1995) were observed in their homes where play occurred naturally with their siblings. In contrast, children were videotaped in a university lab while playing with their parents in Nielsen and Dissanayake’s study (2000). Finally, Astington and Jenkins (1995) observed children playing with familiar peers in their preschool. Thus, it is difficult to compare the studies as play happened in different contexts, which may have influenced children’s play behaviors differently.
1.6 Conclusions

The review of the literature has demonstrated that language, play, and ToM develop in concert in children with TD. This is further illustrated in Table 4, where the merging of cells indicates overlap between two or more of the three cognitive abilities. Major developmental changes take place between three and four years of age. The child becomes able to construct longer and more complex sentences using conjunctions and sentential complements. As language becomes more skilled, the child is also able to express more complex thoughts and ideas, as can be seen by the increased use of cognitive state terms. Further, language allows the child to create sophisticated pretend play scenarios that extend beyond personal experiences and that are shared in social pretend play. In social pretend play, children take on roles of familiar or fictional characters. In order to establish boundaries within and outside play, children use their language skills to talk about their pretend play (i.e., metacommunication). For example, two children can negotiate play in which one gets punished by ‘his mom’ even though the ‘mom’ is only a child without the right to punish someone. Thus, for metacommunication, the child needs both linguistic ability and the awareness that this form of communication is necessary to engage in successful social pretend play.

In contrast, the pattern is not that clear for children with SLI. In these children, the development of language comprehension and/or production is delayed. This means that between the ages of three and four years, sentences are less likely to be complex. This may influence the child’s ability to express himself, which in turn may affect the development of ToM and/or pretend play. The question is if all three cognitive abilities are affected or show delays to a similar extent or if they are distinct and unrelated. Single associations (e.g., language – play or language – ToM) have been investigated and inconsistent results have been reported. Further, to date there is no published study that has investigated how the cognitive abilities of language, ToM, and pretend play may interrelate in children with SLI. Thus, the aims of this study were to compare the ToM and pretend play performances of children with SLI with those of children with TD and to determine how ToM performance is associated with the ability to engage in pretend play.
Table 4

*Milestones in the Development of Language, Theory of Mind (ToM), and Pretend Play*

<table>
<thead>
<tr>
<th>Age</th>
<th>Language</th>
<th>ToM</th>
<th>Pretend Play</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 – 24</td>
<td>Uses 50 words by 18 months</td>
<td>Shows awareness of other people’s likes/dislikes</td>
<td>Shows simple pretend behaviors, e.g., “drinking” from a cup</td>
</tr>
<tr>
<td></td>
<td>Shows vocabulary spurt</td>
<td></td>
<td>Uses inanimate toys (e.g., dolls, toy animals) in simple play actions</td>
</tr>
<tr>
<td></td>
<td>Uses two-word utterances</td>
<td></td>
<td>Engages in solitary play</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Uses desire terms</td>
</tr>
<tr>
<td>25 – 36</td>
<td>Uses 2-3 word sentences</td>
<td>Knows the difference between “pretending” and “trying” to perform an action</td>
<td>Pretend play actions are more detailed and organized in a logical sequence</td>
</tr>
<tr>
<td></td>
<td>Refers to self by name, “me”, or “mine”</td>
<td></td>
<td>Play themes are domestic themes and those of personal experience (e.g., shopping, birthday)</td>
</tr>
<tr>
<td></td>
<td>Understands and uses Why-questions</td>
<td></td>
<td>Engages in parallel play</td>
</tr>
<tr>
<td></td>
<td>Uses embedded sentences</td>
<td>Holds two different representations of an ambiguous object in mind</td>
<td>Shows an awareness of social roles</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Expresses desires and feelings</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Language</td>
<td>ToM</td>
<td>Pretend Play</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>----------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>37 – 48</td>
<td>Uses 4-5 word sentences</td>
<td>Shows awareness of other people’s knowledge</td>
<td>Pretend play is very complex</td>
</tr>
<tr>
<td>months</td>
<td></td>
<td></td>
<td>Story plots are planned, may be more than one plot</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Enacts roles of familiar or fictional characters</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Engages in social pretend play</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uses conjunctions (and, because)</td>
<td>Begins to engage social role play</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uses proportionally fewer desire terms and more cognitive terms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Masters sentential complements</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49 – 60</td>
<td>Uses 5-6 word sentences</td>
<td>Passes false belief tasks consistently</td>
<td>Takes extensive time for “setting up the play scene”</td>
</tr>
<tr>
<td>months</td>
<td></td>
<td></td>
<td>Play themes go beyond personal experience (e.g., dragon hunt)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Story plots are complex; there are sub-plots</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Engages in elaborate social role play</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uses cognitive terms in adult-like manner</td>
<td></td>
</tr>
<tr>
<td>61 – 72</td>
<td>Uses complex sentences</td>
<td>Begins to pass second-order false belief tasks</td>
<td>Play themes are elaborate and story plots are complex</td>
</tr>
<tr>
<td>months</td>
<td></td>
<td></td>
<td>Decline of overt pretend play behaviors</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note. Merged cells indicate an overlap between two or more cognitive abilities. Adapted from Bartsch & Wellman (1995); Paul (2001); Pronin Fromberg & Bergen (2006); Stagnitti (1998).

1.7 Research Aims and Hypotheses
This study had three aims. The first aim was to investigate how children with SLI performed on a variety of ToM tasks relative to their peers with TD. Research has shown that language abilities were closely related to ToM and that early language abilities predicted later ToM task performance in preschool children (Milligan et al., 2007). Given that children with SLI experience a deficit in language abilities it seems very likely that they also display difficulties in solving ToM tasks. However, results of the few studies that assessed ToM performance in children with SLI provided inconsistent results (Miller, 2004; Farrant et al., 2006). In the current study, ToM was assessed with the ToM scale, an additional change-of-location false belief task, and two additional faux-pas tasks. Thereby, this study adds to the existing research, which may help to clarify whether children with SLI experience difficulties with ToM or not. It was hypothesized that children with SLI would perform more poorly on ToM tasks relative to their peers with TD.

The second aim was to compare children with and without SLI on measures of pretend play. It has been reported that children with SLI differed in their social play behaviors from peers with TD. This study extended previous research by using a variety of different measures, including a parental questionnaire, a standardized pretend play test, and a role play activity. It was hypothesized that children with SLI would show less sophisticated forms of pretend play relative to their peers with TD.

The last aim was to investigate whether there was a relation between ToM task performance and pretend play skills in children with and without SLI. Studies have shown that ToM is positively related to pretend play skills in children with TD and that children with higher ToM scores show more sophisticated forms of pretend play (Astington & Jenkins 1995; Youngblade & Dunn, 1995). However, studies have not included children with SLI and this study was the first to explore possible correlations among ToM and pretend play. It was hypothesized that children with lower scores on ToM tasks would also display less sophisticated forms of pretend play.
2 Method

2.1 Power Analysis

A power analysis was conducted to determine the required sample size. The study of Farrant et al. (2006) aided in calculating the effect size for ToM performance in children with and without SLI. The means and standard deviations on the false belief task were 0.35 and 0.67 for children with SLI and 1.45 and 0.83 for children with TD, yielding an effect size of \( d = 1.46 \). Based on Cohen’s \( d \) (1988), a sample size of 12 children per group was needed to achieve a power of .90 (two-tailed). Furthermore, the power to detect possible correlations between pretend play and ToM was calculated and a correlation coefficient of \( r = .40 \) was estimated (Astington & Jenkins, 1995; Nielsen & Dissanayake, 2000). With a sample size of 40 children, a power of .83 (one-tailed) would be attained (Cohen, 1988). Therefore, the aim was to recruit at least a total of 40 children (i.e., 20 children with and 20 without SLI).

2.2 Participants

Forty-four children between the ages of 4;0 and 5;11 (years; months) participated in this study. All children were monolingual German speakers and came from Lower Saxony and from Hannover. Of these children, 22 had SLI and 22 had typical development.

There were two test sessions. In the first session, inclusion and exclusion criteria were confirmed and in the second session, study-relevant testing was conducted. Children were included when they passed a pure-tone hearing screening of 500, 1000, 2000, and 4000 Hz at 25 dB and achieved a score of at least 80 on the Coloured Progressive Matrices (CPM; Bulheller & Häcker, 2002), a nonverbal intelligence test. Children with TD had to score within the normal range on the Sprachentwicklungstest für 3-5jährige Kinder (SETK 3-5; Grimm, Aktaş, & Frevert, 2001) [Test of language development for children between 3-5 years of age]. That is, they had to achieve a \( T \)-score between 40 and 60 on each of the four subtests of the SETK 3-5. Based on criteria established in an epidemiological study (Tomblin, Records, & Zhang, 1996), children were identified as SLI when they performed at least -1.25 standard deviations below the mean on at least two out of four subtests. All standardized language measures in this study used \( T \)-scores, for which means are 50 with standard deviations of 10. Thus, a \( T \)-score of 37 is equal to a standard deviation of -1.25. In addition, all children were assessed with the Aktiver
Wortschatztest-Revised (AWST-R; Kiese-Himmel, 2005) [active vocabulary test] because vocabulary skills were not covered by the SETK 3-5. AWST-R scores were only used to estimate vocabulary skills and not to determine study eligibility. This was done because the AWST-R is standardized only for children up to the age of 5;5 (years; months), thus not providing T-scores for older study participants. Children were excluded when they had been diagnosed with a developmental or neurological disorder, such as autism or epilepsy, or were being raised bilingual (by parent report). Brief descriptions of the cognition and language measures are provided below.

2.2.1   Inclusion and Exclusion Criteria

2.2.1.1   Coloured Pogressive Matrices (CPM)
This test is based on the English version of the CPM (Raven, 2003) and assesses nonverbal cognitive abilities in children between the ages of 4 and 11 years. In a multiple-choice format, the child is asked to identify a missing pattern to complete a larger visual pattern. The test contains 36 items. The CPM was administered and scored by the experimenter as outlined in the CPM manual.

2.2.1.2   Sprachentwicklungstest für 3-5jährige Kinder (SETK 3-5)
The SETK 3-5 (Grimm et al., 2001) assesses language development in children between 3 and 5;11 years of age. It consists of five subtests that investigate receptive as well as expressive language abilities. These subtests include: 1) language comprehension [Verstehen von Sätzen], 2) acquisition of plurals [morphologische Regelbildung], 3) nonword repetition [phonologisches Gedächtnis für Nichtwörter], 4) sentence repetition [Satzgedächtnis], and 5) auditory memory span for words [Gedächtnisspanne für Wortfolgen]. The child is asked to label pictures, act out instructions or repeat words and sentences presented orally by the experimenter. Performance on the last subtest was not considered to establish inclusion because it provides only informal information on a pass/fail basis. The SETK 3-5 was administered and scored by the experimenter as outlined in the SETK 3-5 manual. As the SETK 3-5 does not provide a composite score, the T-scores of the four subtests were averaged and used for statistical analyses. The internal consistency of the four subtest scale was Cronbach’s α = .92, which is
considered to be reliable (Portney & Watkins, 2000). The inter-correlations of the four subtests are shown in Table 5. As expected for a reliable scale, all tasks were positively correlated.

Table 5

*Inter-Correlations of Language Subtests of the SETK 3-5*

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Language comprehension</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Plural acquisition</td>
<td>.626***</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>3. Nonword repetition</td>
<td>.640***</td>
<td>.806***</td>
<td>--</td>
</tr>
<tr>
<td>4. Sentence repetition</td>
<td>.690***</td>
<td>.807***</td>
<td>.888***</td>
</tr>
</tbody>
</table>

*Note. SETK 3-5 = Test of language development for children between 3-5 years of age.*

***$p < .001$.***
2.2.1.3 Aktiver Wortschatztest – Revised (AWST-R)
The AWST-R (Kiese-Himmel, 2005) assesses expressive vocabulary skills in children between 3 and 5;5 years of age. The child is presented with 75 pictures, representing 68% nouns and 32% verbs. The child is presented with one picture at a time and is asked to label it. The AWST-R was administered and scored by the experimenter as outlined in the AWST-R manual.

2.2.1.4 Questionnaire on Demographic Information
Parents were asked to provide information on their child’s medical history to ensure that inclusion criteria were fulfilled. Further, parents indicated their highest level of education and occupation to enable matching and to estimate their socioeconomic status based on the Winkler-Index (Winkler, 1998). Finally, information on the composition of the family was collected to facilitate interpretation of the study. Please see Appendix A for an English translation of this questionnaire. The actual questionnaire completed by parents was written in German.

2.2.2 Matching Criteria
The children with TD were matched to those with SLI for gender and chronological age (+ or – 2 months). Further, an attempt was made to match the children for the presence of siblings and the highest level of parental education as an estimate of family socio-economic status (SES). These matching criteria were chosen because ToM performance was associated with siblings and familial SES in previous studies (e.g., Jenkins et al., 2003; Perner et al., 1994; Peterson, 2000).

Information on SES was collected using the Winkler-Index (Winkler, 1998). The index is a composite score of the three variables of parental education, current occupation, and income. Each variable is rated on a scale from 1 to 7 (i.e., 1 being the lowest level and 7 the highest level). The scores of the three variables are then added and this combined score, the SES index, allows an estimate of social class. That is, an SES index of 3 – 8 is indicative of the lower class, 9 – 14 of the middle class, and 15 – 21 of the upper class.

The variable education is composed of the two variables “highest level of education” and “formal job training or university degree” (Winkler, 1998). The German educational system is quite different from the North American system. It is therefore important to mention that formal
job training requires, in general, three years of vocational education and training. During this time, the apprentice receives practical training at the workplace and, in addition, it is also mandatory to attend part time a vocational school specific to the career (Bundesministerium für Bildung und Forschung, n.d.).

Furthermore, there are three distinct school types that differ in the level of qualification that students can achieve. That is, students graduating after ninth grade are most likely to receive vocational training in the area of trade (e.g., plumber, painter, baker, or hairdresser). Students graduating after 10th grade are likely to receive training in the area of industry and commerce/mercantile trade (e.g., clerks in insurance companies, travel agencies, banks, or lawyer offices). Students graduating after 12th or 13th grade are eligible to apply for further education at a university.

In this study, the scale of the education variable was modified in order to make it more comprehensible for parents and to facilitate checking of the correct response. Thus, parents indicated their highest level of education on a scale with a range from 0-7; 0 being the lowest level (no school diploma + no formal job training) and 7 being the highest (high school diploma + university degree). Parents did not report any problems with checking their highest level of education. Further, parents provided their current job titles and these were then categorized by the experimenter into one of seven categories according to the descriptions in the index. The categories ranged from “unskilled labor” (1) to employees with a status of director or judge (7). However, information on family income was not gathered and the missing variable was estimated using the mean of the other two variables (i.e., education, occupation) as recommended by Winkler (1998). The scores from each variable were added, giving a range from 2-21.

2.2.3 Description of Participants

The first panel of Table 6 provides an overview of the participants’ characteristics. In both groups, there were 59% (13) boys and 41% (9) girls. The group of children with SLI ranged in age from 49 to 69 months and the group of children with TD ranged from 48 to 71 months. Both groups had a mean age of 59 months. Most pairs were also matched for the presence of siblings; that is, 91% (20) of children with TD and 86% (19) of the children with SLI had one or more siblings, $t(42) = .159, p = .875$. 

45
Consistent with instructions for the Winkler-Index (Winkler, 1998), the person with the higher level of occupation was used to calculate the family SES. Thus, in the majority of cases, the scores of the fathers were used to estimate familial SES, as 41% of the mothers stayed at home or were on maternity leave. With respect to parental education, there were no significant differences between mothers’ and fathers’ highest level of education, ($M_{mother} = 4.47, SD = 1.87; M_{father} = 4.93, SD = 1.95$), $t(81) = .43, p = .667$.

An attempt was made to match the two groups on familial SES, but this proved impossible, given the families who volunteered to participate. Parents of children with TD had significantly higher levels of education than parents of children with SLI ($M_{TD} = 6.09; M_{SLI} = 4.38$), $t(41) = 3.99, p < .001$. That is, 59% of parents of children with TD had a university degree as opposed to approximately 10% of parents of children with SLI. Comparing this information with federal census data showed that parents of children with TD were more likely to have a university degree (59%) than the federal average (14%) of the adult population in 2007 (Statistische Ämter des Bundes und der Länder, 2009). In contrast, educational levels of parents of children with SLI were more representative of the population. That is, 57% of parents of children with SLI received a formal job training compared to 50% of the federal average.

Further, parents of children with TD had a significantly higher SES index than parents of children with SLI, $t(41) = 4.18, p < .001$. That is, parents of the TD group were significantly more likely to come from the upper class (i.e., have an SES index between 15 – 21), whereas parents of the SLI group were more likely to come from the middle class (i.e., have an SES index between 9 – 14). These significant group differences reflected, in part, an unexpectedly high number of children who were thought to be typically developing but who did not meet the inclusion criteria. That is, children who might have been matches for children with SLI with respect to SES had to be excluded because their cognitive or language performances were not within the normal range.
Table 6

*Characteristics, Cognitive and Language Abilities for Children with and without SLI*

<table>
<thead>
<tr>
<th>Participants’ characteristics</th>
<th>TD (n = 22)</th>
<th>SLI (n = 22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Boys</td>
<td>59%</td>
<td>59%</td>
</tr>
<tr>
<td>Age (in months)</td>
<td>59.41 (6.99)</td>
<td>59.41 (6.86)</td>
</tr>
<tr>
<td>% Siblings</td>
<td>91%</td>
<td>86%</td>
</tr>
<tr>
<td>SES index*</td>
<td>16.98 (3.91)</td>
<td>11.71 (4.34)</td>
</tr>
<tr>
<td>Parental education</td>
<td>6.09 (1.34)</td>
<td>4.38 (1.47)</td>
</tr>
<tr>
<td>Parental occupation</td>
<td>5.23 (1.34)</td>
<td>3.48 (1.81)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cognitive and language abilities</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CPM</td>
<td>100.59 (10.99)</td>
<td>100.18 (9.64)</td>
</tr>
<tr>
<td>SETK 3-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language comprehension</td>
<td>52.27 (6.44)</td>
<td>40.91 (8.84)</td>
</tr>
<tr>
<td>Language production</td>
<td>55.82 (7.85)</td>
<td>38.23 (6.72)</td>
</tr>
<tr>
<td>Nonword repetition</td>
<td>52.32 (6.58)</td>
<td>29.32 (4.81)</td>
</tr>
<tr>
<td>Sentence repetition</td>
<td>54.64 (8.42)</td>
<td>30.68 (6.13)</td>
</tr>
<tr>
<td>SEKT Mean*</td>
<td>53.76 (5.74)</td>
<td>34.78 (4.62)</td>
</tr>
<tr>
<td>AWST-R*</td>
<td>54.23 (9.23)</td>
<td>36.30 (7.55)</td>
</tr>
</tbody>
</table>

*Note.* Unless otherwise indicated, all table entries are group means followed by standard deviations in parentheses.

CPM = Colored Progressive Matrices. SETK 3-5 = Sprachentwicklungstest für 3-5 jährige Kinder [Test of language development in 3-5 year old children]. AWST-R = Aktiver Wortschatztest – Revised [Test of active vocabulary].

*a* n = 21; *b* n = 20.

*p* < .001.
It is important to point out, however, that although children were not matched deliberately for nonverbal cognitive abilities, the groups’ nonverbal abilities were, nonetheless, comparable. As can be seen in the second panel of Table 6, the two groups had similar mean scores of 100 on the CPM. Thus, there were no significant group differences, $t(42) = .131, p = .896$. These scores indicated that children’s language problems were not due to cognitive deficits, therefore fulfilling the definition of SLI. Moreover, this equivalence facilitates group comparisons on the measures of ToM and pretend play as nonverbal cognitive ability is often a confounding factor (e.g., Tomblin et al., 1997).

As expected, there were significant group differences on all standardized language measures. Children with TD performed significantly better on the SETK 3-5 than children with SLI ($M_{TD} = 53.76; M_{SLI} = 34.78$), $t(42) = 12.08, p < .001$. All children with TD but only 20 out of the 22 children with SLI completed the vocabulary assessment with the AWST-R. The scores for the two children with SLI were not available because, during testing, the children refused to label the items and testing was discontinued. Again, the expressive vocabulary of children with TD was significantly better than that of children with SLI ($M_{TD} = 54.23; M_{SLI} = 36.30$), $t(40) = 6.85, p < .001$ (see Table 6).

2.3 Procedures

Figure 1 provides a summary of the testing and matching procedures for the participants. Children with SLI were recruited from three different facilities: a teaching hospital, a kindergarten for children with special language needs, and a private speech therapy practice. The majority of children with SLI (16 children) were recruited from the Clinic of Phoniatrics and Pedaudiology (PHO), which is part of a teaching hospital in Hannover (Medizinische Hochschule Hannover). The clientele of the PHO are mainly outpatients between the ages of 0 to 16 years, who are referred by their pediatricians for a thorough diagnostic procedure due to suspected hearing, speech, and/or language disorders. In addition, children were also recruited from a kindergarten for children with special language needs. In this kindergarten, six children were eligible and their parents were contacted. Five families gave consent and one family gave consent under the condition that they would get the original video files at the day of testing. It would not have been feasible to analyze the videos and obtain the experimental data within that time because of the time consuming coding procedures. Therefore, the latter child was excluded.
and five children participated. Further, several speech-language pathologists working in private speech practices were contacted. However, most of the children in their caseloads were not in the required age range, were bilingual, or had only articulation problems. Five children were eligible and one family gave consent.

Children with TD were recruited from five kindergartens in the city of Hannover and Hannover Region respectively. It is important to note that a German kindergarten is conceptually different from a kindergarten in North America. In Germany, children attend a kindergarten between the ages of three to six and then they enter primary school. Furthermore, there is no curriculum in place, that is, children are not being taught how to read or write.

In two kindergartens with high numbers of bilingual children, parents were informed directly about the study by the director. In the other kindergarten, parents were informed during a parent meeting about the possibility to participate in the study. Sixty-nine parents gave consent. However, of these, 12 children were not in the required age range and 11 failed language or other testing. Parents of children who failed testing were informed about the test results and advised to consult their pediatrician.

Children with TD or attending the kindergarten for children with special language needs were tested in the kindergarten. Children who had an appointment in the PHO were tested in the clinic. Prior to testing, parents received an informed consent form that described the purpose and procedures of the study. Further, parents had the opportunity to contact the examiner or the PHO to ask questions or talk about concerns that they had. The children were assessed individually in a quiet room that was equipped with a video camera. If desired by parents or children, children were accompanied by either a parent or a kindergarten teacher during assessment. Children were offered a break or scheduled for a second appointment, depending on their needs. Parents or children could withdraw consent at any time.

There were two test sessions and testing for both sessions combined was completed in approximately 90 minutes. All tasks were administered in German. In the first session, the child’s hearing, cognition, and language were tested to ensure eligibility, as described in the inclusion section. Furthermore, parents were asked to fill out two questionnaires. The first questionnaire gathered demographic information to establish inclusion criteria and to facilitate interpretation of the study (e.g., parental education and occupation, child’s medical history). In
the second questionnaire, parents were asked to provide information regarding their child’s play behaviors (e.g., favorite play themes or play partners). It was assumed that by assessing the children directly and using a parental questionnaire it would be possible to gain a more comprehensive picture of the children’s pretend play skills. An English translation of the play questionnaire can be found in Appendix B. The actual questionnaire completed by parents was written in German. During the second session, the child’s performance was assessed on a ToM measure and two pretend play measures, to be described further below. The second session was videotaped for analysis and to establish reliability. After the second session was completed, the children received stickers and a DVD with the role play activity recording as tokens of gratitude.

However, there were two different procedures when written consent and verbal assent was obtained. Parents of children with TD and of those attending special kindergarten gave permission before any testing began. Similarly, the children were asked if they wanted to participate before they were tested. English translations of the information letters and informed consent forms for parents as well as the assent form for children can be found in Appendices C, D, and E. German versions of these documents were used in the actual study. Children seen in the clinic first underwent the hearing, cognition, and language testing, which is part of the standard procedure in the PHO. When the doctor discussed the test results, parents of children who were eligible were informed about the study and were asked whether they were interested in participating. Parents received the informed consent form and were given the time to read it and ask questions. Depending on the distance parents had to travel, children of parents who gave written consent were either tested the same day after having a break or scheduled for a second appointment. Children’s verbal assent was obtained before testing for the second session began.
Testing and Matching Procedures of Participants

TEST SESSION 1

- 12 children did not meet language test criteria
- 4 children did not meet other testing criteria

22 children with SLI

Matched for age and gender

46 children with TD

22 children with SLI, 22 children with TD

TEST SESSION 2
2.3.1 Measures
All children participated in a number of tasks that assessed their ToM understanding and three different measures were used to capture their abilities to engage in pretend play: the Child-Initiated Pretend Play Assessment (Stagnitti, 2007), a role play activity, and a parental questionnaire on pretend play. By assessing the children directly and using a parental questionnaire, it was possible to gain a more comprehensive picture of the children’s pretend play skills than just a single measure would allow. Both the ChIPPA and the role play activity were videorecorded for later analysis. In turn, each of these measures will be described.

2.3.1.1 Theory of Mind
Wellman and Liu (2004) proposed a ToM scale that captures the gradual development of ToM in children between 2;11 to 6;6 years of age. The scale examines five ToM aspects (desires, beliefs, knowledge access, false belief, and emotion) in the sequence they are mastered by children with TD. That is, children understand desires before beliefs and they understand beliefs before they realize that these beliefs can be false. Thus, the ToM scale was used in this study to gain a comprehensive picture of the overall ToM development in children with SLI.

The ToM scale is a relatively recent measure and has not been used in studies investigating potential links between ToM performance and social behaviors. However, there are reasons to assume that performance on the ToM scale may also be associated with language or social behaviors, such as pretend play. The ToM scale assesses children’s understanding of different mental states (e.g., desires, beliefs, and knowledge) using mental state verbs, which are associated with language abilities. Moreover, a representation of other people’s mental states is vital for social interactions and, therefore, might also be important during social pretend play.

The ToM scale has been translated and tested with German-speaking children (Hofer & Aschersleben, 2004; Kristen, Thoermer, Hofer, Aschersleben, & Sodian, 2006). However, the original ToM scale includes a so-called belief-emotion task, which has not been translated or tested in the German version of the ToM scale. The authors did not provide an explanation for the omission of this task in the German scale (Kristen et al., 2006).

The tasks of the ToM scale have low verbal input with brief and simple task narratives, which was thought to provide children with SLI with the best opportunity to answer them correctly. Previous studies investigating children’s ToM understanding mainly used false belief tasks (i.e.,
unexpected contents and change-of-location). Therefore, a change-of-location task was also administered in addition to the five tasks of the ToM scale. Thus, children would receive the two false belief tasks that have most often been used in previous ToM research (Wellman et al., 2001). In the change-of-location task, a person leaves an object at place A and, in his absence, the object is moved to place B by another person. When the first person returns, the child is asked where he will look for the object. This task has been used numerous times and is well established in the literature (Wellman et al., 2001). In order to minimize inattentiveness or frustration in children with SLI, no other false belief task was added. All tasks had either one pretest or two control questions (e.g., “What is in the box?”) and one or more target questions (e.g., “What does the boy think is in the box?”). To present the tasks, toy dolls and pictures were used as props and the structure of the tasks offered the experimenter the possibility to prompt the child once (forced choice).

Pilot testing showed that the ToM scale with the added change-of-location task might lead to ceiling effects in children with TD. Therefore, two syntactically and conceptually more demanding faux pas tasks were added (Banerjee, 2000). These faux pas tasks were chosen because they tapped into first-order knowledge as well as second-order belief and could assess higher order ToM understanding in a timely manner.

In these faux pas tasks, the child listens to a short story using pictures in which a character unknowingly commits a faux pas. That is, the character makes, unintentionally, a negative statement, which upsets someone. In the control questions, the child is asked to identify the faux pas and who committed it, whereas in the target questions, the child has to answer different questions that tap into various aspects of ToM understanding. That is, the child has to judge what the story character knows and what he thinks about another person’s thoughts. The faux pas tasks were scored as correct when the child answered all control questions as well as the second-order belief question correctly (e.g., “What does Sven think that Annika thinks about his painting?”). Every answer that contained a correct negative statement was given 1 point (e.g., “The painting is stupid/bad/ugly”).

There were eight tasks in total and the children participated in as many as possible. However, testing was discontinued when children became inattentive or frustrated. A correctly answered target question was given 1 point if all corresponding control questions were correctly answered.
as well, as is standard in the literature (Wellman et al., 2001). Thus, a range of 0-8 points was possible. For task descriptions please see Appendix F.

2.3.1.2 Child-Initiated Pretend Play Assessment

The children’s pretend play behaviors were assessed with the Child Initiated Pretend Play Assessment (ChIPPA, Stagnitti, 2007). The ChIPPA is divided into two segments of 15 minutes of testing: (a) a conventional imaginative session with gender-neutral, conventional toys (i.e., a farm set, truck, and dolls) and (b) a symbolic play session with unstructured play materials (e.g., boxes, cone, sticks, and cloth dolls). Both the experimenter and the child are sitting on the floor and the child is introduced to the toys. The child is encouraged to play with the toys in whatever manner she/he wants and the experimenter does not initiate or facilitate the play.

Each of the 15 minute play segments is divided into three 5 minute sections. In the first and last five minutes, the child plays alone, whereas in the middle section, the examiner randomly models five play actions. For the farm set, these play actions are: walk doll, wave doll’s hand, doll pats the cow, doll fixes the truck’s wheel, and doll drives the truck. In the symbolic session, the cloth doll is made to walk, wave, drive, go to bed, and drink from the cone. Modeling of the actions may start earlier if the child does not engage with the toys. Testing in each session ends either after 15 minutes or after the child has been prompted once (i.e., “Is there anything else that you can play with these toys?”).

Scoring of the ChIPPA is based on nonverbal behaviors and each behavior is coded as either “behavioral”, “repetitive”, “functional”, or “elaborate”. In addition, it is noted how often a child uses objects to stand for something else and imitates the modeled play actions of the examiner. Table 7 provides a more detailed description of the coding scheme with examples. The ChIPPA is norm-referenced and provides standard scores for the number of object substitutions and for the percentage of elaborate pretend play actions (PEPA) for both play sessions and a composite score. For this test, the mean standard score is 100 with a standard deviation of 15.
<table>
<thead>
<tr>
<th>Play behavior codes</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>B = Behavioral</td>
<td>Child is not engaged with the play materials, non-play behaviors</td>
<td>Fiddles with toys, mouths objects, daydreams</td>
</tr>
<tr>
<td>R = Repetitive</td>
<td>Child repeats an action or series of actions more than twice. The third time, the action or series is scored ‘R’</td>
<td>Builds the same tower three times and destroys it in the same way</td>
</tr>
<tr>
<td>F = Functional</td>
<td>Child uses the play materials in their conventional ways (i.e., functionally)</td>
<td>Stands up animals, demonstrates how objects work, pushes truck back and forth</td>
</tr>
<tr>
<td>E = Elaborate</td>
<td>Child performs functional actions that are executed carefully in a play context or sequence with attention to detail and uses controlled, fluid movements</td>
<td>Sequence of doll loads the truck, drives truck to farm, and feeds animals</td>
</tr>
<tr>
<td>NOS = Number of Object Substitutions</td>
<td>Each object that is substituted for something else is recorded as an object substitution. To be counted as an object substitution, the object must be deliberately used in a play action in a non-conventional manner</td>
<td>Trailer used as a boat, fence pieces to make a slide, or boxes used as drums</td>
</tr>
<tr>
<td>Play behavior codes</td>
<td>Description</td>
<td>Examples</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>NIA = Number of Imitated Actions</td>
<td>Imitated actions are only scored in the middle 5 minute segment of both the conventional imaginative and symbolic play session. They have to resemble the actions of the experimenter and have to occur within two actions of the modeled action</td>
<td>Experimenter models play action (e.g., walk doll) and child immediately imitates action</td>
</tr>
</tbody>
</table>

*Note.* Adapted from Stagnitti (2007).
2.3.1.3  Role Play Activity
The ChIPPA focused on nonverbal play behaviors; therefore, children participated in another pretend play activity to assess their verbal behaviors and role play abilities. The children were introduced to a box that contained toys of the brand ‘Playmobil’. The plastic figurines represented ‘children’ and ‘adults’ that were male and female. Play materials came from different play themes so that children could engage in either social themes (e.g., house) or in adventurous play themes (e.g., wild life).

Children’s role play abilities were observed in child-experimenter dyads. This form of dyad was chosen because children with SLI were recruited from a clinic with a broad catchment area and familiar peers were not readily available. Moreover, the play interactions of child-adult dyads (i.e., parents) were observed in a previous study that reported a link between ToM and pretend play (Nielsen & Dissanayake, 2000).

The experimenter initiated role play by modeling a short story that involved role assignments given to the child and different figurines. The same short story was enacted with each child individually and a detailed description of the story can be found in Appendix G. Briefly, the experimenter suggested playing zoo. A zoo scene was set up and toy figurines, taking complementary roles (e.g., parent-child), went to the zoo. The experimenter’s figurine brought along a baby chick that later got lost. The experimenter asked the child to help find the baby chick. When the chick was found, the short story was over. Then the child was asked to think of another story that both the experimenter and the child could enact. This could mean either extending the zoo story or to setting up something new. The child decided on the play script and was prompted if necessary.

2.3.1.4  Pretend Play Questionnaire
Parents were asked to fill out a questionnaire regarding their child’s play behaviors at home, e.g., whether the child prefers to play alone or with friends. These questions were based on the literature regarding pretend play (Farver & Shin, 1997; Howes & Smith, 1995; Stagnitti, 1998) and the format of response required parents to check the most appropriate answer from four response choices. A higher score indicated higher levels of pretend play skills; the possible range of scores was 8-32. Response choices for each question were presented to parents in
random order. However, scores depicting the developmental progression from 1-4 are provided in the copy of the questionnaire, which can be found in Appendix B.

2.3.2 Coding Procedures

2.3.2.1 Child-Initiated Pretend Play Assessment

The ChIPPA was designed as a test that could be scored online, i.e., the examiner could score the child’s play behaviors while the child is playing (Stagnitti, 2007). However, this was not feasible for the purpose of reliability testing. Therefore, the ChIPPA session was videotaped and analyzed later on. Using a similar approach to that as described by Stagnitti et al. (2000), the videos were segmented into 10-second time intervals, using an MPEG splitter software (KAKSOFT Studio, 2008). Each time interval was then analyzed and every completed play behavior was coded according to the manual (see Table 6 for description of codes). It was possible for a time interval to receive several codes or to stay ‘empty’ when the behavior was continuous and was coded in the previous interval.

2.3.2.2 Role Play Activity

For practical reasons, the video files of the role play activity were split into time intervals and each interval was coded. The video files of the 15-minute role play activity were segmented into intervals of 6 seconds with an MPEG splitter software (KAKSOFT Studio, 2008). Intervals of 6 seconds were chosen because the child’s play action was likely to be completed within this time segment; thus, there was little overlap of different play codes.

Coding was done in one pass. Each segment received only one code per person (i.e., the child and the experimenter), coding the highest level of the child’s play and whether the experimenter prompted the child. The coding scheme was based on research by Andresen (2005), Astington and Jenkins (1995), and Youngblade and Dunn (1995) and consisted of four codes that were hierarchically organized. The code Basic Pretend Play indicated a basic level of pretend play, whereas a code of Directing Play indicated advanced pretend play abilities. A description of the codes and examples are provided in Table 8. If a segment did not contain any pretend play or prompting (e.g., the child was moving from the toy box to the play scene), then the segment was coded as “empty”. Raw scores as well as proportions (e.g., pretend play segments over the total of all segments) were calculated for each category.
It took approximately 30 minutes per child to prepare the video file (i.e., splitting the file, establishing beginning and end of play). In addition, about 45 minutes were needed to code and analyze the complete video file. Thus, it took approximately 55 hours to analyze all 44 video files of the role play activity.
### Table 8

**Description of Role Play Activity Codes**

<table>
<thead>
<tr>
<th>Role play codes</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empty</td>
<td>Child is not engaged in pretend play behaviors</td>
<td>Daydreams, looks at toys, or questions materials (“What is that?”)</td>
</tr>
<tr>
<td>Basic Pretend</td>
<td>Child engages in verbal or nonverbal behavior concerned with pretend play. It includes:</td>
<td>Makes animals walk, sits figurines in a toy car and makes car drive, or produces sounds (onomatopoeia) that indicate that the toy car had a crash</td>
</tr>
<tr>
<td></td>
<td>- using replica objects as real objects,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- substituting one object for another,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- representing absent objects or properties,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- setting up an imaginary play scene</td>
<td></td>
</tr>
<tr>
<td>Role Enactment</td>
<td>Child is acting at being someone else by using the character’s voice, language, and/or behavior.</td>
<td>“Children, be careful!”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Husband, come!”</td>
</tr>
<tr>
<td>Role Assignment</td>
<td>Child assigns a pretend role to himself or herself, the experimenter, a toy figurine, or an ambiguous toy. This assignment has to be done verbally and explicitly.</td>
<td>“I’m the baby”, “You are the mommy”, This is the sister”, or “This is a climbing wall”</td>
</tr>
<tr>
<td>Directing Play</td>
<td>Child verbally directs/explains the events that unfold during pretend play, verbally explains the changed meaning of toys, or suggests play ideas.</td>
<td>“Then the girl falls down and she breaks her leg”, “You mow the lawn now” (child hands over a toy that was previously used as a stroller), “At first, it’s closed, OK? And then you open it”</td>
</tr>
</tbody>
</table>

Linguistic indicators might be the change of tense (i.e., use of subjunctive) [Konjunktiv] or use of words, such as “wohl”, “zum Spass” [for fun], “OK?”
<table>
<thead>
<tr>
<th>Role play codes</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretend Play</td>
<td>Questions or statements from the experimenter that motivate a child to start pretend play or introduce a new pretend play idea. Prompts can also be questions/suggestions that help the child to develop pretend play ideas and/or help to maintain a pretend play sequence.</td>
<td>“We could play zoo”, “Who comes with us to the zoo?”, or “Who can I be/play?”</td>
</tr>
</tbody>
</table>

*Note. Adapted from Andresen (2005); Astington & Jenkins (1995).*
Validity and Reliability

The validity and reliability of the different play measures was established with different raters for each individual play measure. All raters were native German-speaking and were not informed about the specific aims of the study. With respect to the validity of the pretend play questionnaire, five raters were asked to rate the response choices of the questionnaire in the order of their developmental progression. Three of these raters were kindergarten teachers and two were Speech-Language Pathologists. All raters were asked to assign a score from 1-4 to each response choice per question. A score of one was indicative of an earlier developmental stage and a score of four was indicative of a later, more mature developmental stage with respect to children’s pretend play abilities.

At least 4 out of 5 raters rated the order of the response choices consistent with the developmental progression originally developed by the experimenter for six out of the eight questions. The two questions for which the raters disagreed with the original developmental progression were question 4 (verbal activity) and question 8 (role play alone). At least 4 of 5 raters suggested that the last two response choices in question 4 and 8 should be reversed. These suggestions were followed and the resulting scoring scheme is shown in Appendix B.

To ensure that coding of the two pretend play measures (i.e., ChIPPA and role play activity) was reliable, two raters were trained to independently rate 20% of randomly selected videos of the ChIPPA or the role play activity. The second rater for the ChIPPA videos had a diploma in International Business and Foreign Languages and no background in speech-language pathology. She studied the comprehensive coding manual and received 20 hours of training on identifying completed pretend play actions and coding. The second rater for the role play activity videos was a Speech-Language Pathologist and she had also been a second rater during pilot testing. She received six hours of training on coding. Both raters were blind to the language status and ToM performance of the children.

To establish inter-rater reliability, the Kappa statistic was used to correct for agreements expected by chance. An agreement of $k = .80$ or more was desired, which is considered to be an excellent agreement (Portney & Watkins, 2000). Table 9 provides a summary of the individual kappa values for the ChIPPA and role play activity, respectively. The desired kappa was not
achieved for all codes. However, all values indicated substantial to excellent levels of agreement. Furthermore, the reported kappa values are comparable to previous reliability reports in the literature (Jenkins & Astington, 2000; Swindells & Stagnitti, 2006).
Table 9

*Kappa Values for the Child-Initiated Pretend Play Assessment (ChIPPA) and Role Play Activity Codes*

<table>
<thead>
<tr>
<th></th>
<th>ChIPPA</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PEPA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional</td>
<td></td>
<td>.77</td>
</tr>
<tr>
<td>Symbolic</td>
<td></td>
<td>.75</td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td>.76</td>
</tr>
<tr>
<td><strong>NOS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Symbolic</td>
<td></td>
<td>.90</td>
</tr>
<tr>
<td><strong>NIA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional</td>
<td></td>
<td>n/a*</td>
</tr>
<tr>
<td>Symbolic</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

|                  |        |       |
| **Role Play Activity** |    |     |
| Child related      |        |       |
| Empty Segments    |        | .86   |
| Pretend Play      |        | .76   |
| Role Assignment   |        | .80   |
| Direct Play       |        | .78   |
| Role Enactment    |        | .77   |
| Overall           |        | .79   |
### Role Play Activity Kappa

<table>
<thead>
<tr>
<th>Role Play Activity</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimenter related</td>
<td></td>
</tr>
<tr>
<td>Prompts</td>
<td>.86</td>
</tr>
</tbody>
</table>

*Note.* no child imitated.

PEPA = Percentage of elaborate pretend play actions. NOS = Number of object substitutions. NIA = Number of imitated actions.
3 Results

The section addresses the research hypotheses in the order they were posed. Thus, group differences in ToM performance are reported first, followed by the analyses of group performances on the different play measures (i.e., pretend play questionnaire, ChIPPA, and role play activity). As a primary analysis for each dependent measure, an analysis of covariance (ANCOVA) with SES index as covariate was conducted. This was done because children with and without SLI differed significantly from each other with respect to family SES and SES was, therefore, a potential confound for all dependent measures (Jenkins et al., 2003; Ramsey, 2006). Finally, the associations between ToM and play are described. This was an exploratory study; therefore, a relatively liberal alpha level of .05 was employed for all statistical tests unless otherwise noted. Further, all statistical analyses were based on two-tailed testing.

3.1 Theory of Mind

As described earlier, ToM was tested using the five tasks of the ToM scale (Wellman & Liu, 2004), an additional change-of-location task, and two faux-pas tasks (Banerjee, 2000). The eight tasks became increasingly difficult. Thus, children differed in how many tasks they participated in, depending on their attention span and frustration level. All 22 children with TD participated in the ToM scale, 21 participated in the first faux-pas task, and 15 continued in the second faux-pas task. Of the 22 children with SLI, 15 participated in all questions of the ToM scale, but seven did not complete the last task of the scale (i.e., emotion task). Of the 15 children with SLI who completed the ToM scale, nine participated in the first faux-pas task and four also participated in the second faux-pas task.

As is standard in the literature (Wellman et al., 2001), the target question of each task was scored as correct if all corresponding control questions were answered correctly as well. When children did not receive credit for the target question, a record was kept of whether they answered the control questions for each task correctly. The latter measure was thought to be an indication of children’s attentiveness and their ability to understand the task narratives.

The internal consistency of the eight ToM tasks was Cronbach’s $\alpha = .70$, which is considered to be acceptable (Field, 2005). The inter-correlations of the eight ToM tasks are shown in Table
10. As expected for a reliable scale, most tasks were positively, but modestly, correlated with each other. The one exception was the belief task, which showed slight negative correlations with the unexpected contents and the emotion tasks. The exclusion of the belief task from the composite ToM score led to an increased internal consistency of Cronbach’s $\alpha = .73$. However, the exclusion of the belief task from the composite ToM score did not alter the statistical results for the primary research questions. Therefore, because the belief task was part of the ToM scale as originally published, it was included and the composite ToM score of the eight ToM tasks was used for further statistical analyses.
### Table 10

*Inter-Correlations of Theory of Mind Tasks*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Desires</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Beliefs</td>
<td>.322*</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. KA</td>
<td>.202</td>
<td>.174</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. FB-UC</td>
<td>.237</td>
<td>-.158</td>
<td>.311*</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. FB-CL</td>
<td>.289</td>
<td>.163</td>
<td>.546***</td>
<td>.391**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Emotion</td>
<td>.182</td>
<td>-.102</td>
<td>.122</td>
<td>.392**</td>
<td>.232</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. FP-V</td>
<td>.137</td>
<td>.166</td>
<td>.204</td>
<td>.296</td>
<td>.175</td>
<td>.187</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>8. FP-R</td>
<td>.182</td>
<td>.058</td>
<td>.271</td>
<td>.392**</td>
<td>.232</td>
<td>.323*</td>
<td>.755***</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* KA = knowledge access. FB-UC = false belief unexpected contents. FB-CL = false belief change-of-location. FP-V = faux pas violin. FP-R = faux pas rocket.

* *p < .05, ** *p < .01, *** *p < .001.
The first research hypothesis was that children with SLI would perform more poorly on ToM tasks than their peers with TD. A composite ToM score was calculated, which was the sum of all correctly answered target questions for which the corresponding control questions were also answered correctly. Children with TD ($M = 4.59, SD = 1.79$) had a broader range of scores (i.e., 0-8) than children with SLI ($M = 2.64, SD = 1.36$, range $= 0-5$). There were two children with SLI who performed relatively well on the ToM tasks and answered 5 out of 8 tasks correctly. These two children with SLI were the only ones in the SLI group who passed almost all tasks of the ToM scale (including both false belief tasks) and failed only the last task ‘appearance-reality emotion’. An ANCOVA with SES index as the covariate revealed that the covariate was not significantly related to ToM, $F(1, 40) = 2.95$, $p = .094$, $r = .26$. Further, consistent with the hypothesis, the composite ToM score was significantly lower for children with SLI ($M_{adj} = 3.03$, $SD_{adj} = 1.68$) than for children with TD ($M_{adj} = 4.34$, $SD_{adj} = 1.67$), $F(1, 40) = 5.59$, $p = .023$, $r = .35$ after controlling for family SES.

The means and standard deviations for the group performances on each ToM task are provided in Table 11. Children who did not participate in a task were included and were treated as failing the task. As not all means were normally distributed, a series of independent $t$-tests (i.e., first 5 tasks) as well as Mann-Whitney tests (i.e., last 3 tasks) were conducted to compare ToM performance in children with and without SLI. A Bonferroni correction was applied; thus, all results are reported at a $p$-level of .006 (i.e., $0.05 / 8$). Only two comparisons were statistically significant. Children with TD performed significantly better on both false belief tasks than children with SLI; $t(42) = 3.33$, $p = .002$ and $t(33.52) = 3.25$, $p = .003$ for unexpected contents and change-of-location, respectively. Further, when age was also considered, only 5 out of 13 four-year-old children with SLI passed one or both of the false belief tasks, whereas 11 out of 13 four-year-old children with TD passed one or both of the false belief tasks. With respect to the five-year-old children, 7 out of 9 children with SLI passed one or both false belief tasks whereas 9 out of 9 children with TD passed one or both false belief tasks.

In addition, a task-by-task analysis of the performance of the two groups is provided in Figure 2. Some of the bar graphs do not add up to 100% because some children did not complete the task or failed on one of the control or target questions. These children were not included in the bars. As can be seen in Figure 2, the overall performance of both groups was consistent with expected developmental trends. Children performed better on the desire task than on the false belief tasks.
(Wellman & Liu, 2004). Further, children with SLI showed floor effects for the last three ToM tasks. Only one child with SLI answered the control questions of the emotion and two faux-pas tasks correctly but failed all target questions.

Interestingly, both groups performed better on the change-of-location task than on the unexpected contents task. That is, children with TD showed a difference in correct performance from 68% (content) to 91% (change) which was significant, $t(21) = 2.49, p = .021$. Children with SLI showed a difference in correct performance from 23% (content) to 50% (change) which was significant as well, $t(21) = 2.94, p = .008$. This indicates that the latter task was easier for both groups.
Table 11

*Means and Standard Deviations (in parentheses) of Correctly Answered Theory of Mind Questions for Children with and without SLI*

<table>
<thead>
<tr>
<th>ToM tasks</th>
<th>TD (n = 22)</th>
<th>SLI (n = 22)</th>
<th>p</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not-own desire</td>
<td>.91 (.29)</td>
<td>.68 (.48)</td>
<td>.065</td>
<td>.60</td>
</tr>
<tr>
<td>Not-own belief</td>
<td>.73 (.46)</td>
<td>.73 (.46)</td>
<td>1.00</td>
<td>.00</td>
</tr>
<tr>
<td>Knowledge access</td>
<td>.77 (.43)</td>
<td>.50 (.51)</td>
<td>.062</td>
<td>.57</td>
</tr>
<tr>
<td>FB unexpected contents</td>
<td>.68 (.48)</td>
<td>.23 (.43)</td>
<td>.002</td>
<td>.99</td>
</tr>
<tr>
<td>FB change location</td>
<td>.91 (.29)</td>
<td>.50 (.51)</td>
<td>.003</td>
<td>1.03</td>
</tr>
<tr>
<td>Emotion</td>
<td>.23 (.43)</td>
<td>.00 (.00)</td>
<td>.019</td>
<td>1.07</td>
</tr>
<tr>
<td>FP violin</td>
<td>.14 (.35)</td>
<td>.00 (.00)</td>
<td>.076</td>
<td>.80</td>
</tr>
<tr>
<td>FP rocket</td>
<td>.23 (.43)</td>
<td>.00 (.00)</td>
<td>.019</td>
<td>1.07</td>
</tr>
</tbody>
</table>

*Note.* FB = false belief. FP = faux-pas.

Maximum score for each task = 1.
Figure 2

Percentages of Children with and without SLI who Passed Target and Control Questions on Theory of Mind (ToM) Tasks

**Note.** The values represent percentages.

3.2 Differences in Family Background and ToM Performance

Research on individual differences in ToM performance suggested that children who had siblings would perform better on false belief tasks than children with no siblings; in particular, children with lower language levels might benefit from having siblings (Jenkins & Astington, 1996; Perner et al., 1994). In this study, five participants were only children, 24 children had one sibling, 13 had two siblings, one child had four siblings, and one had five siblings. Previous studies investigated the relation between false belief understanding and the presence of siblings. Therefore, a composite false belief score of the two false belief tasks was created ($r = .39$ and Cronbach’s $\alpha = .58$). The false belief composite score was significantly correlated with the ToM composite, $r = .804$. Children with TD had a mean composite false belief score of 1.59 with a standard deviation of .67 and children with SLI had a mean composite score of .73 with a standard deviation of .70. Correlational analyses using Pearson product moment correlations were conducted for each language group individually, using the composite false belief score as well as the composite ToM score.

There were no significant correlations between the number of siblings and the composite false belief score or the composite ToM score in children with TD, $r = .019$, $p = .932$ and $r = -.125$, $p = .578$, respectively. There was no significant correlation between the number of siblings and the composite false belief score in children with SLI, $r = .069$, $p = .760$. The association between the composite ToM score and number of siblings showed a trend in the expected direction; however, it did not reach significance, $r = .412$, $p = .057$.

3.3 Pretend Play

The second research hypothesis was that children with SLI would show less sophisticated forms of pretend play than their peers with TD. To test this hypothesis, three pretend play measures were used: a pretend play questionnaire completed by the parents, the Child-Initiated Pretend Play Assessment (ChIPPA, Stagnitti, 2007), and a role play activity. Descriptive statistics are provided for each measure in turn, followed by analyses of group performances on the different play measures, using ANCOVA with SES index as the covariate. Further, examples of children’s play behaviors are provided to illustrate the different levels of play abilities observed in this study.
3.3.1 Pretend Play Questionnaire

All parents completed a pretend play questionnaire that contained eight questions. The format of response required parents to check the most appropriate answer from four response choices. A higher score indicated higher levels of pretend play skills. The possible range of scores was 8-32. As mentioned earlier, Appendix B contains the questionnaire items and information on the scoring of the response choices.

Eight parents of children with TD and eight parents of children with SLI checked more than one answer on one or more questions of the questionnaire. In these cases, the scores of multiple answers were averaged and the mean was used to calculate the overall score. Three parents of children with TD and two parents of children with SLI did not check any answer for one or more questions. Here, the child’s individual mean of the questionnaire was substituted in order to calculate the total score. The internal consistency reliability of the pretend play questionnaire was Cronbach’s $\alpha = .72$, which is considered to be acceptable (Field, 2005).

Table 12 provides an item-by-item summary of the questionnaire responses. On most of the questions, parents of children with SLI rated their children slightly lower ($M = 21.67, SD = 4.61$) than parents of children with TD did ($M = 23.81, SD = 2.99$). The biggest group difference can be seen in item three. That is, parents of children with SLI indicated that their children spent less time in play with other children after kindergarten. The overall pretend play questionnaire score was lower for children with SLI ($M_{adj} = 21.94, SD_{adj} = 4.38$) than for children with TD ($M_{adj} = 23.64, SD_{adj} = 4.36$). The ANCOVA revealed that the covariate, SES index, was not significantly related to the pretend play questionnaire, $F(1, 40) = .21, p = .650, r = .07$, and group differences on the questionnaire were not significant after controlling for the effects of SES, $F(1, 40) = 1.38, p = .247, r = .18$. Thus, parents in the two groups did not differ in how they evaluated their children’s play behaviors.

3.3.2 Gender Differences on the Pretend Play Questionnaire

Ayres and Levé (2006) reported that there might be gender differences in play behaviors. Therefore, the effect of gender was investigated in a supplementary analysis. Children were matched for gender; thus, both groups consisted of 13 boys and 9 girls. An independent $t$-test was conducted to assess gender differences. The overall pretend play questionnaire score was higher for boys ($M = 23.64, SD = 4.22$) than for girls ($M = 21.44, SD = 3.32$); however,
differences were not significant, $t(42) = 1.85, p = .072, d = .58$. Thus, there were no gender differences on the pretend play questionnaire.
Table 12

Means and Standard Deviations (in parentheses) for Children with and without SLI on the
Pretend Play Questionnaire

<table>
<thead>
<tr>
<th>Questionnaire Item</th>
<th>TD (n = 22)</th>
<th>SLI (n = 22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Books</td>
<td>2.86 (.79)</td>
<td>2.89 (.90)</td>
</tr>
<tr>
<td>2. Favorite play partner</td>
<td>3.43 (.84)</td>
<td>3.25 (.90)</td>
</tr>
<tr>
<td>3. Time in play after kindergarten</td>
<td>3.18 (.73)</td>
<td>2.49 (.93)</td>
</tr>
<tr>
<td>4. Language use in play</td>
<td>3.36 (.71)</td>
<td>3.23 (1.11)</td>
</tr>
<tr>
<td>5. Favorite play themes</td>
<td>2.70 (1.03)</td>
<td>2.51 (1.01)</td>
</tr>
<tr>
<td>6. Play ideas</td>
<td>2.98 (1.01)</td>
<td>2.35 (1.27)</td>
</tr>
<tr>
<td>7. Role play with children</td>
<td>2.88 (1.08)</td>
<td>2.38 (1.06)</td>
</tr>
<tr>
<td>8. Role play alone</td>
<td>2.42 (1.06)</td>
<td>2.58 (1.16)</td>
</tr>
</tbody>
</table>

Note. Maximum score for each item = 4.
3.3.3 Child-Initiated Pretend Play Assessment

All children participated in the two play sessions of the ChIPPA (i.e., conventional play session and imaginative play session). Both sessions were videotaped and the video was then segmented into intervals of 10 seconds. According to the manual, each play session should last for 15 minutes (Stagnitti, 2007). However, not all children were able to play for 15 minutes in one or both play sessions. Therefore, the number of segments available for coding varied. The first panel of Table 13 provides an overview of the means and standard deviations for the number of coded segments in each ChIPPA play session for both language groups. The number of coded segments for the conventional play session violated the assumption of a normal distribution; therefore, a Mann-Whitney test was performed. The group difference in coded segments was not statistically significant for the conventional play session, \( U = 205.00, p = .377 \). The number of coded segments for the imaginative play session was normally distributed and no group differences were found, \( t(42) = 1.46, p = .152 \).

The ChIPPA manual does not provide standard scores for the number of object substitutions (NOS) during the conventional session because the majority of children did not use the conventional toys in object substitution (Stagnitti, 2007). Further, as suggested by the manual, the number of actions that were imitated by the child during modeling was recorded. However, only a few children imitated occasionally; therefore, no further analysis was conducted. No child was excluded and statistical analyses were based on the whole sample. The second panel of Table 13 provides a summary of the standard scores for percentage of pretend play actions (PEPA) and number of object substitutions (NOS) for both groups.

As expected, children with SLI had lower combined PEPA scores (\( M_{adj} = 71.08, SD_{adj} = 17.61 \)) and NOS scores (\( M_{adj} = 84.64, SD_{adj} = 7.72 \)) than children with TD (\( M_{adj} = 92.70, SD_{adj} = 17.54 \) and \( M_{adj} = 88.71, SD_{adj} = 7.69 \), respectively). To assess the main hypothesis, two ANCOVAs were conducted. In these analyses, the dependent variables were the combined PEPA standard score and the combined NOS standard score, respectively. The covariate was the SES index.

The covariate, SES index, was not significantly related to PEPA, \( F(1, 40) = .00, p = .998, r < .001 \) or NOS, \( F(1, 40) = .05, p = .831, r = .03 \). After controlling for the effect of SES, there was a significant group difference for PEPA, \( F(1, 40) = 13.83, p = .001, r = .50 \); however, the groups
did not differ for NOS, $F(1, 40) = 2.55, p = .18, r = .24$. Thus, children with SLI showed significantly less sophisticated pretend play than their peers; however, children with SLI did not differ from their peers in their number of object substitutions when family SES was taken into account.
Table 13

Means, Standard Deviations (in parentheses), and Effect Sizes for Number of ChIPPA Segments Coded, and PEPA and NOS Standard Scores for Children with and without SLI

<table>
<thead>
<tr>
<th>Session Type</th>
<th>TD (n = 22)</th>
<th>SLI (n = 22)</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of coded segments</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional</td>
<td>83.68 (9.60)</td>
<td>76.27 (19.73)</td>
<td>.51</td>
</tr>
<tr>
<td>Symbolic</td>
<td>65.64 (23.02)</td>
<td>55.68 (22.18)</td>
<td>.44</td>
</tr>
<tr>
<td>Standard scores on ChIPPA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEPA</td>
<td>93.72 (14.32)</td>
<td>72.79 (21.14)</td>
<td>1.18</td>
</tr>
<tr>
<td>NOS</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Symbolic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEPA</td>
<td>93.95 (13.99)</td>
<td>77.40 (11.85)</td>
<td>1.28</td>
</tr>
<tr>
<td>NOS</td>
<td>88.58 (8.71)</td>
<td>84.33 (3.99 )</td>
<td>.67</td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEPA</td>
<td>92.71 (14.97)</td>
<td>70.81 (16.20)</td>
<td>1.41</td>
</tr>
<tr>
<td>NOS</td>
<td>88.86 (8.74)</td>
<td>84.45 (4.13)</td>
<td>.69</td>
</tr>
</tbody>
</table>

*Note.* Maximum number of segments = 90. n/a = not applicable because conventional NOS scores are not provided by the manual.

ChIPPA = Child-Initiated Pretend Play Assessment. PEPA = Percentage of elaborate pretend play actions. NOS = Number of object substitutions.
3.3.4 Gender Differences on the Child-Initiated Pretend Play Assessment
As previously stated, gender differences might occur in pretend play behaviors (Ayres & Levé, 2006). Therefore, the effect of gender on elaborate pretend play was investigated in supplementary analyses. The ChIPPA standard scores of some age groups are different for boys and girls to account for gender differences in play. Thus, PEPA and NOS raw scores were used in two independent t-tests. The PEPA raw score was higher for boys ($M = 86.03$, $SD = 35.95$) than for girls ($M = 78.61$, $SD = 44.74$), whereas the NOS raw score was higher for girls ($M = 6.17$, $SD = 7.09$) than for boys ($M = 5.27$, $SD = 4.05$). However, these differences were not significant, $t(42) = .61$, $p = .546$, $d = .18$ and $t(42) = .53$, $p = .596$, $d = .16$, respectively. Thus, there were no gender differences on the ChIPPA.

3.3.5 Pretend Play Examples
In the conventional play session, the children received farm set toys, that is, a tractor and a trailer, fences, a doll, and various farm animals (e.g., cows, horses, pigs). In both groups there were a few children who did not play or only minimally engaged with the toys. However, the majority of children in both groups engaged with these toys in a conventional way. Thus, they set up the fences to make a barn or range and the farmer (i.e., the doll) drove the animals into the fence area. Some girls used the fence pieces as hurdles in a horse tournament. After setting up the scene, children with TD then started to create a story line, such as the animals were hungry and needed to be fed or it became night and the animals needed shelter. In contrast, many of the children with SLI remained in functional play. They grouped and regrouped the fence pieces or animals several times without starting to play in an organized sequence.

However, there were also exceptions. One child with SLI set up the farm scene and then the farmer and his wife had to feed the hungry animals. Suddenly, the child completely changed the play theme. He pretended to be a dinosaur, who escaped from “Dino-Land”, which is a fun and adventure park in Germany. He would stomp and roar very loudly because the dinosaur was hungry. The dinosaur caught some animals and brought them into his nest to eat them later. When the dinosaur returned to the farm, the farmer wanted to protect his animals and then was himself eaten by the dinosaur. Thus, the wife came to rescue her husband but the dinosaur was stronger and won. This example illustrates that some children with SLI were capable of engaging in an elaborate pretend play sequence in spite of their deficient language skills.
In the symbolic play session, children received ambiguous materials such as boxes, sticks, pebbles, a tin, and a cloth doll. There were children in both groups who had no idea what to play with these materials. Some children with and without SLI were happy to repeatedly drop the pebbles through a cone into the tin from various heights. However, many children made a series of constructions and they built rockets, airplanes, or ships with cannons. One child with TD enacted an extraordinary story with the materials. The shoebox became the attic and the cloth doll was a ghost. The ghost had to clean up his room (i.e., the attic) and then went to sleep. Then the child took the experimenter’s cloth doll used for modeling. The doll became another ghost who lived in another attic. However, this ghost was very unfortunate and had no toys, food, or friends. Finally, the two ghosts met and became friends. From then on, they lived happily ever after and always cooked, ate, and played together.

One child with SLI also showed imaginative skills, pretending to be a magician. The wooden stick was his magician wand and a tea towel was a magic towel ("pocus-blanket"). He invented different spells to transform the play materials and objects in the room into wild animals and cookies. After the magician invented an imaginary house for the experimenter, he was tired and fell asleep.

3.3.6 Role Play Activity
All children participated in the role play activity. The role play activity took 15 minutes and was videotaped. The video was then segmented into intervals of 6 seconds. Due to experimenter error, the beginning of the role play activity was not recorded for two children (i.e., for one child with and one child without SLI). When the errors were detected, the children were given the opportunity to play longer to make up for the lost recording time. The child with SLI did not play longer and only 42 segments were available for coding. This was more than two standard deviations below the SLI group mean (122.95 segments ($SD = 31.06$)). Therefore, this child was excluded. The child with TD continued to play so that 144 segments (out of a possible 150) were available for coding; therefore, the child was included.

A descriptive analysis of coded segments for the remaining children with TD showed that on average 133.36 segments ($SD = 33.55$) were available with a range of 26 – 150 segments. A frequency analysis revealed that one child had only 26 segments, which was more than two standard deviations below the group mean and was therefore excluded as an outlier.
exclusion of the two outliers, the means were 138.48 (24.04) segments and 126.81 (25.88) segments for children with TD and SLI, respectively. Because the data violated the assumption of a normal distribution (i.e., 15 out of 21 TD children had the maximum segments of 150), a Mann-Whitney test was performed. The difference in coded segments was statistically significant, \( U = 124.50, p = .011 \). Moreover, some individual children played longer than others did. Therefore, the pretend play codes were calculated as a percentage of each child’s total number of segments.

The coding category of Role Assignment during the role play activity was relatively low in occurrence for both groups (\( M_{TD} = 3.25, SD = 2.71, M_{SLI} = 2.94, SD = 2.17 \)). Role Assignment and Directing Play were categories in which the child informs the play partner about changes in play ideas and how the play unfolds. Therefore, these two categories were collapsed to form one Direct Role Play category. In this study, many children created roles and directed activities for characters to engage in without explicitly verbalizing the assigned roles. Therefore, it seemed that no information was lost when the two categories were combined.

Figure 3 provides an overview of the percentages of segments that received the different pretend play codes during the role play activity for both groups, including the new Direct Role Play category. As can be seen, children with SLI spent more time in Basic Pretend Play (\( M = 32.25, SD = 6.94 \)) and Role Enactment (\( M = 5.86, SD = 5.68 \)) than children with TD (\( M = 28.49, SD = 6.78 \) and \( M = 3.71, SD = 3.21 \), respectively). Further, children with TD were more likely to talk about how the play unfolds (\( M = 26.04, SD = 12.25 \)) than children with SLI (\( M = 17.91, SD = 8.64 \)). However, the variables could not be entered into parametric analyses because they may violate the assumption of normal distribution (Pennington, James, McNally, Pay, & McConachie, 2009; Smith, 2006). By converting the raw frequencies of the different play behavior codes into percentages, the four pretend play codes became dependent on each other as they always added up to 100%. Parametric procedures assume variables to be independent; therefore, a more stringent compositional analysis was conducted. Each variable was transformed with a logarithm to form a new variable (i.e., log-ratios) as described by Pennington et al. (2009). These log-ratios were then assumed to be normally distributed and were submitted to a multivariate analysis of covariance (MANCOVA). For more information on compositional analysis please see Aitchison (1986) for theoretical and Pennington et al. (2009) for applied aspects.
A MANCOVA was conducted to compare the different pretend play codes used by the two groups during the role play activity. The independent variable was language group (SLI, TD), the dependent measures were the four play codes (i.e., log-ratios of Empty, Basic Pretend Play, Role Enactment, and Direct Role Play), and the covariate was SES index. The MANCOVA showed that the covariate was not significantly related to the pretend play codes, Wilk’s Λ = .99, $F(4, 35) = .09, p = .985$. Further, there were no differences between children with and without SLI in the role play activity, Wilk’s Λ = .81, $F(4, 35) = 2.12, p = .099$. Thus, children with SLI did not differ in the distribution of their role play behaviors from children with TD.
Figure 3

Percentages of Pretend Play Codes during Role Play Activity for Children with and without SLI

Note. The values represent mean percentages. The error bars represent standard deviations.

Pret play = Basic Pretend Play. Role enact = Role Enactment. Direct play = Direct Role Play
3.3.6.1 Gender Differences in the Role Play Activity
The effect of gender was investigated in a supplementary analysis as boys and girls may differ in their pretend play behaviors (Ayres & Levé, 2006). A MANOVA was conducted with gender as the independent variable and the four pretend play codes (i.e., log-ratios of Empty, Basic Pretend Play, Role Enactment, and Direct Role Play) as dependent variables. There were no group differences with respect to gender, Wilk’s Λ = .94, \( F(4, 37) = .57, p = .686 \).

3.3.6.2 Experience with Playmobil Toys as a Potential Confound
It is theoretically possible that children who were not familiar with Playmobil toys may have performed differently in the role play activity than children who were familiar with them. A supplementary analysis was conducted to rule out the possible confound of experience with Playmobil toys. According to the parents, two children had no experience with Playmobil, 16 children had experience with Playmobil but preferred different toys, and 24 children liked to play with Playmobil toys. For this analysis, the two children with no Playmobil experience were added to those who had some experience but preferred to play with other toys. A MANOVA was conducted to investigate if children’s performance in the role play activity differed based on their Playmobil experience. The dependent measures were Empty, Basic Pretend Play, Role Enactment, and Direct Role Play (log-ratios) and the independent variable was Playmobil experience (prefers different toys, plays with Playmobil). Children’s role play did not differ based on their experience with Playmobil, Wilk’s Λ = .92, \( F(4, 37) = .79, p = .540 \). Thus, Playmobil experience did not seem to have influenced children’s role play performance.

3.3.7 Pretend Play Examples
In the role play activity, children could choose from different toys (e.g., animals, vehicles, fences, and ambiguous materials). The experimenter suggested playing zoo, a topic that was assumed to be familiar to all children. As reported in the previous section on the ChIPPA, a range of different play behaviors was observed for children with and without SLI. In both groups, there were a few children who were very reluctant to engage in the role play activity and who mainly performed basic pretend play actions. However, most children actively engaged in the role play activity and the zoo topic. There were children in both groups who had nice
pretend play ideas but they did not verbally inform the experimenter about them. This made it difficult for the experimenter to join the play. On the other hand, there were children who took the lead during role play and added funny and imaginative details or story plots to the zoo theme.

For instance, one child with SLI labeled one part of the zoo as a petting zoo area and, in the surrounding area there was a free running lion and a tiger. Both were wild but kind because they were just born. He assigned complementary roles as father (played by the child) and son (played by the experimenter). The father worked in the zoo and sold the tickets. But then, all of the tickets were blown away and everyone could get into the zoo without paying admission. As there were no tickets to be sold, the father could join his son on the tour to the petting zoo, which was guided by the lion. The lion only pretended to know the zoo well, when in fact he did not. Soon he had to climb on a tower where he could see better. But then, he could not get down and the father had to help him. Because the ladder was stolen by villains, the father took his flying sled. The father saved the lion and then flew everyone home.

A child with TD began to set up the zoo scene and then she decided that the zoo had a playground. The ambiguous toys were made to be a slide, a climbing wall, and a carrousel. In front of the slide was a little boy who did not dare to slide without his older sister. Thus, a long line up happened and the other children became angry with the little boy. One child was so angry that he closed the door of the slide and thereby made the little boy slide. Now all children could finally slide. Then the little boy and his sister went to their parents, who had arrived with their sled and were waiting next to it. The two children asked for money because the sled could also be a carrousel if money was put in. While the children were in the carrousel, the mother went to her friend who was climbing the climbing wall. Then the zoo was about to close but suddenly the friend got stuck in the wall and could not get down. The father and the rest of the family had to help her to get down. Then the zoo closed and they went all home.

3.3.8 Pretend Play Prompts
During the role play activity, prompts by the experimenter were coded. Modeling the initial short story was also coded in Pretend Play Prompts. This was done because it was assumed that some children might need more modeling than others to start role playing. Further, prompts could also be questions or suggestions that helped the child to develop pretend play ideas and/or
helped to maintain a pretend play sequence. Children with SLI were prompted more frequently ($M = 44.19, SD = 11.18$) than children with TD ($M = 41.44, SD = 11.02$). However, this difference was not statistically significant, $t(40) = -.80, p = .427$ with a small effect size ($d = .25$).

### 3.4 Relationship between Theory of Mind and Pretend Play

The last research hypothesis was that children with lower scores on ToM tasks would also display poorer forms of pretend play. There are some inconsistencies in the literature in terms of whether pretend play in general (i.e., object substitution, representing absent objects and properties) or only role play in particular (i.e., assigning roles and directing play) is related to ToM (e.g., Astington & Jenkins, 1995; Nielsen & Dissanayake, 2000). Therefore, the combined standardized PEPA score from the ChIPPA was used to capture aspects of general pretend play. In addition, the coding category of Direct Role Play (log-ratios) from the role play activity was chosen because this category was significantly correlated with false belief understanding in preschoolers with TD in previous studies (Astington & Jenkins, 1995). Further, it was important to test if, consistent with previous studies, correlations between ToM, language, and SES could be replicated (e.g., Milligan et al., 2007; Jenkins et al., 2003). Therefore, the language score from the SETK 3-5 and the SES index were also included. For all correlational analyses Pearson product moment correlations were used.

As can be seen in the top panel of Table 14, the ToM composite score was significantly associated with the standardized PEPA score as well as with Direct Role Play from the role play activity, $r = .424, p = .004$ and $r = .397, p = .009$, respectively. Thus, consistent with expectations, children who had a higher ToM composite score were more likely to engage in elaborate forms of pretend play and to inform the playmate about play ideas than children with lower ToM scores. The analyses also showed significant correlations between the ToM composite score and language ability, $r = .634, p < .001$ as well as between the ToM composite score and the SES index, $r = .491, p = .001$. Further, there was a significant association between language ability and SES index, $r = .663, p < .001$. Therefore, it was important to determine whether the correlations of ToM with PEPA and Direct Role Play remained significant, when the effects of language ability and SES index were controlled.
The partial correlation analysis (bottom panel of Table 14) showed that the associations between the ToM composite score and both measures of pretend play were no longer significant after controlling for language and SES, $r = .041, p = .805$ and $r = .247, p = .130$ for PEPA and Direct Role Play, respectively. Thus, inconsistent with the main hypothesis, the association between ToM and role play abilities was not independent of language abilities and SES. However, the relationship between PEPA and Direct Role Play remained significant, $r = .501, p = .001$. It might be that PEPA and Direct Role Play remained associated because they both captured children’s ability to put pretend play behaviors into a logical sequence (e.g., load the truck, place the doll into the truck, and drive to the farm to feed the animals).
Table 14

Correlations and Partial Correlations for Theory of Mind (ToM), PEPA, Direct Role Play, Language, and SES for Children with and without SLI

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ToM composite&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. PEPA&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.424**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Direct Role Play&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.397**</td>
<td>.610***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. SETK mean&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.634***</td>
<td>.600***</td>
<td>.403**</td>
<td></td>
</tr>
<tr>
<td>5. SES index&lt;sup&gt;c&lt;/sup&gt;</td>
<td>.491**</td>
<td>.349*</td>
<td>.223&lt;sup&gt;d&lt;/sup&gt;</td>
<td>.663***</td>
</tr>
</tbody>
</table>

Partial correlations with SETK and SES controlled

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ToM composite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. PEPA</td>
<td>.041</td>
<td></td>
</tr>
<tr>
<td>3. Direct Role Play</td>
<td>.247</td>
<td>.501**</td>
</tr>
</tbody>
</table>

Note. PEPA = percentage of elaborate pretend play actions from the ChIPPA. SETK = German test of language development. SES = socio-economic status.

Values of Direct Role Play are represented as log-ratios.

<sup>a</sup>n = 44, <sup>b</sup>n = 42, <sup>c</sup>n = 43, <sup>d</sup>n = 41.

* p < .05, ** p < .01, *** p < .001.
Given that previous studies investigated the relationship between ToM and pretend play in children with TD and not in children with SLI (Astington & Jenkins, 1995; Nielsen & Dissanayake, 2000), the group was split and associations were further analyzed for each language group individually. As can be seen in the top panel of Table 15, in children with TD the ToM composite score was significantly associated with both pretend play measures, \( r = .595, p = .003 \) and \( r = .577, p = .006 \) for PEPA and Direct Role Play, respectively. Further, the ToM composite score was significantly correlated with language abilities and familial SES, \( r = .627, p = .002 \) and \( r = .617, p = .003 \), respectively. The partial correlation analysis (bottom panel of Table 15) showed that the ToM composite score and both pretend play measures remained significantly associated after controlling for language and SES, \( r = .615, p = .005 \) and \( r = .797, p < .001 \) for PEPA and Direct Role Play, respectively.

In contrast, in children with SLI the ToM composite score was negatively correlated with PEPA, \( r = -.352, p = .108 \) and significantly but negatively correlated with Direct Role Play, \( r = -.451, p = .040 \). Moreover, ToM was weakly associated with language, \( r = .061, p = .789 \) and slightly negatively associated with familial SES, \( r = -.144, p = .546 \), which might be due to decreased variability. After controlling for language abilities and SES, the ToM composite was now significantly and negatively associated with PEPA, \( r = -.479, p = .045 \) but was no longer significantly associated with Direct Role Play, \( r = -.382, p = .118 \). In summary, when both children with and without SLI were considered, the ToM composite score was not associated with sophisticated forms of pretend play once language and SES were taken into account. When the two language groups were considered individually in similar analyses, ToM was positively correlated with sophisticated forms of pretend play in children with TD but was negatively associated with such play in children with SLI.
Table 15

*Correlations and Partial Correlations for Theory of Mind (ToM), PEPA, Direct Role Play, Language, and SES for Children with TD*

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ToM composite</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. PEPA</td>
<td>.595**</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Direct Role Play</td>
<td>.577**</td>
<td>.486*</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>4. SETK mean</td>
<td>.627**</td>
<td>.137</td>
<td>.077</td>
<td>--</td>
</tr>
<tr>
<td>5. SES index</td>
<td>.617**</td>
<td>.271</td>
<td>-.029</td>
<td>.540*</td>
</tr>
</tbody>
</table>

Partial correlations with SETK and SES controlled

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ToM composite</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>2. PEPA</td>
<td>.615**</td>
<td>--</td>
</tr>
<tr>
<td>3. Direct Role Play</td>
<td>.797***</td>
<td>.517*</td>
</tr>
</tbody>
</table>

*Note. PEPA = percentage of elaborate pretend play actions from the ChIPPA. SETK = German test of language development. SES = socio-economic status.*

Values of Direct Role Play are represented as log-ratios.

*< .05, **< .01, ***< .001.
Table 16

*Correlations and Partial Correlations for Theory of Mind (ToM), PEPA, Direct Role Play, Language, and SES for Children with SLI*

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ToM composite</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. PEPA</td>
<td>-.352</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Direct Role Play</td>
<td>-.451*</td>
<td>.629**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. SETK mean</td>
<td>.061</td>
<td>.326</td>
<td>.359</td>
<td></td>
</tr>
<tr>
<td>5. SES index</td>
<td>-.144</td>
<td>-.148</td>
<td>.105</td>
<td>.444*</td>
</tr>
</tbody>
</table>

Partial correlations with SETK and SES controlled

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ToM composite</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>2. PEPA</td>
<td>-.479*</td>
<td></td>
</tr>
<tr>
<td>3. Direct Role Play</td>
<td>-.382</td>
<td>.609**</td>
</tr>
</tbody>
</table>

*Note.* PEPA = percentage of elaborate pretend play actions from the ChIPPA. SETK = German test of language development. SES = socio-economic status.

Values of Direct Role Play are represented as log-ratios.

* *p ≤ .05, ** *p < .01.
Furthermore, previous studies used only false belief tasks when assessing the relationship between ToM and pretend play in children with TD (Astington & Jenkins, 1995; Nielsen & Dissanayake, 2000; Youngblade & Dunn, 1995). Therefore, supplementary analyses were conducted to see if the false belief composite score was correlated with the different measures of pretend play in children with and without SLI. Partial correlations (controlling for language and SES) that included both language groups revealed that the false belief composite score was not significantly correlated with elaborate pretend play, \( r = -.104, \ p = .527 \) or Direct Role Play, \( r = .010, \ p = .953 \). When the group was split by language status, the false belief composite score was significantly and positively associated with both measures of pretend play in children with TD, \( r = .467, \ p = .044 \) and \( r = .473, \ p = .041 \) for PEPA and Direct Role Play, respectively. However, for children with SLI the false belief composite score was significantly and negatively associated with PEPA, \( r = -.504, \ p = .033 \) and with Direct Role Play, \( r = -.503, \ p = .034 \). Thus, using the false belief composite score yielded a similar overall pattern of partial correlations as the ToM composite score did.

Finally, further investigation focused on the different patterns of associations in children with and without SLI (i.e., negative and positive relations). The ToM composite score was strongly correlated with PEPA and Direct Role Play in children with TD. The performance of a few children with TD likely contributed to the large effect sizes. That is, three children received PEPA scores of 115 or higher on the ChIPPA and these children also performed well on the different ToM tasks. Moreover, these children also tended to frequently direct the unfolding play events during the role play activity. For instance, one child performed at ceiling in the ToM tasks (i.e., scored 8 out of 8) and spent 54% of coded segments in Direct Role Play, more than any other child in this group. This might explain why the magnitude of the effect here \( (r = .797) \) is considerably stronger than previously reported \( (r = .49) \) for the relation between ToM and Direct Role Play (Astington & Jenkins, 1995).

On the other hand, the performance of three children with SLI might have contributed to the negative associations between ToM and sophisticated forms of pretend play. The two children who had the highest composite ToM score (i.e., 5) relative to the other children with SLI barely engaged in elaborate pretend play and did not take an active role in the role play activity. In addition, one child whose elaborate pretend play score was within the normal range and whose
frequency of directed play events was above the group average had a composite ToM score of 0 It is conceivable that the performances of these three children contributed to the negative correlation between ToM and the different pretend play measures.

4 Discussion
Language, ToM, and pretend play develop in concert in children with TD. However, language abilities are impaired in children with SLI, which may be related to the development of ToM and pretend play. Relative to their age-matched peers, children with SLI performed more poorly on ToM tasks and showed less sophisticated pretend play on the ChIPPA. These findings were consistent with the hypotheses stated earlier. However, there were also results that were unexpected. There was a trend for children with SLI to spend more time in lower levels of pretend play than in directing role play in the role play activity. However, the overall distribution of pretend play codes was not significantly different in the two groups. Moreover, the relationship between ToM performance and children’s ability to engage in sophisticated pretend play was not independent of language abilities and SES. These results will be discussed below.

4.1 Theory of Mind
Children participated in a number of ToM tasks that tapped different aspects of ToM understanding. There were similarities as well as differences between children with and without SLI in their ToM performance. Both groups of children showed a decrease in ToM performance across the different tasks that were organized in a developmental order. That is, both groups were more likely to pass desire or own-belief tasks than false belief tasks. Children with SLI did not consistently fail all tasks but demonstrated better performance on tasks that tested earlier ToM concepts, such as desire, than on later ones, such as false belief. This result is consistent with previous studies (Farrant et al., 2006, Wellman & Liu, 2004) and, furthermore, supports Wellman’s (1990) theory that there may be a developmental progression in ToM understanding in which an understanding of desires precedes an understanding of false beliefs.

Performance on individual tasks of the ToM scale tended to be lower for children with SLI than for children with TD. However, it seemed that, overall, children with SLI showed a sequence of ToM understanding that is comparable to that of children with TD. In order to establish the
sequence and consistent pattern of ToM development in children with SLI, the inclusion of a broader age range (e.g., three to seven years) and a larger sample size would be necessary. It is, for instance, unclear how young three-year-olds with SLI would perform on the desire task relative to their peers. Moreover, by including older children with SLI (e.g., six or seven years of age) it might be possible to find out when they consistently pass false belief tasks. Previous studies reported that preschool children with SLI were likely to fail false belief tasks (Farrant et al., 2006; Farrar et al., 2009; Miller, 2001), whereas school-age children with language impairments passed them (Perner et al., 1989; Ziatas et al., 1998). This may suggest that ToM development in children with SLI is delayed but follows a similar developmental trajectory to that for children with TD.

The two groups also displayed marked differences. That is, children with SLI performed more poorly on the false belief tasks than their age-matched peers with TD. This result is consistent with findings of previous studies (Farrant et al., 2006; Miller, 2001, 2004). One possible explanation for the poor false belief performance might be that children with SLI may have had difficulty following the story task narrative. However, children with SLI performed relatively well on the control questions, suggesting that they paid attention and were able to follow the story task narrative. Further, SLI group performance on the language comprehension subtest showed a mean T-score of 40, which indicated that their language comprehension may have been relatively intact. However, caution is warranted as only one subtest assessed children’s comprehension abilities. Thus, it seems more likely that children with SLI in this study were still in an earlier stage of desire-belief reasoning and had not yet gained the conceptual understanding of false belief (Bartsch & Wellman, 1995; Wellman, 1990).

Interestingly, children in both groups performed significantly better on the change-of-location task than on the unexpected contents task. The question of whether false belief performance varies due to different types of false belief tasks (e.g., change-of-location or unexpected contents) or types of questions asked (e.g., phrasing the target question using “think”, “look”, “believe”, “say”, or “know”), has been the focus of numerous studies (e.g., Gopnik & Astington, 1988; Siegal & Beattie, 1991). Although individual studies have reported differences in performance based on question type (e.g., Yazdi, German, Defeyter, & Siegal, 2006), a meta-analysis (Wellman et al., 2001) concluded that different tasks or question types did not facilitate false belief performance in preschool children with TD, i.e., three-year-olds performed at chance

95
levels and four-year-olds performed above chance. The only linguistic exception was that older preschoolers’ false belief performance might be enhanced by including a temporal marker in the target question (e.g., “Where will Susie look first for the chocolate?” vs. “Where will Susie look for the chocolate?”) (Wellman et al., 2001). However, the current study did not use a temporal marker in the change-of-location task.

Nevertheless, children with and without SLI demonstrated an inconsistent performance pattern across the two false belief tasks. Again, it seems unlikely that false belief performance suffered from poor story comprehension abilities. The change-of-location task involved a brief narrative, whereas the unexpected contents task did not and, yet, performance was better on the change-of-location task. A possible explanation for the differences in false belief performance might be linguistic demands of the target question that varied across the two false belief tasks. That is, the unexpected contents task involved a test question using a sentential complement (“What does Susie think is in the box?”), which might have led to a decline in false belief performance. In contrast, the change-of-location task asked a very simple target question (“Where will Susie look for her chocolate?”). Further, a child could have pointed to the correct location, but needed to provide a verbal answer in the unexpected contents task. Thus, linguistic complexity of the target question and/or response may have influenced false belief performance in children with and without SLI.

Miller (2001, 2004) investigated the impact of linguistic complexity on false belief performance in children with SLI. She reported that children with SLI performed comparably to age-matched peers with TD when linguistic complexity of the false belief tasks was low and comparably to language-matched peers when complexity was high. However, the puzzling finding in Miller (2004) was that children with SLI showed inconsistent false belief task performance even when the same target question was used (i.e., “Where does Susie think the toy is?”) compared to the earlier study. It is not clear how to explain the differences between the studies. It might be that individual differences in SES or the presence of siblings could account for some of the differences. However, this remains speculative because information on these variables was not provided.

This study did not attempt to address the effects of linguistic complexity on false belief performance. However, there was an attempt to minimize the linguistic input of ToM tasks, in
order to provide children with SLI with the best opportunity to answer them. It seems that language demands affected false belief performance in both groups, as they demonstrated an improved performance across the two false belief tasks. The effect of linguistic demands has been extensively studied with children with TD. However, a systematic analysis of varying linguistic demands on ToM performance in children with SLI might provide additional information. Linguistic demands might be one possibility to explain the increase in false belief performance across the two tasks. An alternative explanation might come from the ordering of the tasks. That is, the change-of-location task was always administered second, which might have scaffolded children’s false belief performance. The order of the two false belief tasks was not counterbalanced in this study because both tasks were regarded as equally difficult for preschoolers (Wellman et al., 2001). Thus, when analyzing the varying linguistic demands in children with SLI with a number of different false belief tasks, the ordering of the tasks has to be taken into consideration.

Finally, children’s overall ToM performance might also have been influenced by individual differences. Previous studies reported significant associations between children’s false belief understanding and their language abilities (e.g., Milligan et al., 2007) as well as for the presence of siblings (Jenkins & Astington, 1996; Perner et al., 1994). This study applied a broader measure of ToM (i.e., the ToM scale), which was developed relatively recently (Wellman & Liu, 2004). A composite ToM score was used in correlational analyses and a significant association between the ToM composite and language abilities was found. Remmel and Peters (2008) reported a significant association between language skills of children with cochlear implants and the composite score of the ToM scale, which was comparable to that reported here in its effect size. Moreover, in the current study, the relationship between language skills and the ToM composite was even slightly stronger ($r = .634$) than for the false belief composite ($r = .605$). This might be due to increased variability in the ToM scale and the additional tasks or due to the higher internal consistency of the composite ToM score ($\alpha = .70$) than of the composite false belief score ($\alpha = .58$).

Perner et al. (1994) demonstrated a significant relationship between children’s false belief performance and the presence of siblings, whereas Jenkins and Astington (1996) found a stronger association between false belief and the presence of siblings for children who had lower language skills than for children with higher language skills. This finding was not replicated in
this study, as the false belief composite was not associated with number of siblings present in the family for children either with or without SLI. However, there was a trend in the expected direction for the composite ToM score and number of siblings in the home ($r = .412, p = .057$) for children with SLI, who might benefit in their ToM development from having siblings. Having siblings might increase the frequency of interactions and conversations and thereby lead to more exposure for children with SLI to conflicts, teasing, and overall use of mental state verbs, which might help them to compensate for their deficient language abilities (Jenkins & Astington, 1996).

### 4.2 Pretend Play Measures

Previous studies have shown that children with language impairments performed more poorly on measures of play than their age-matched peers (e.g., Leonard, 1998). However, some of these studies have been criticized for measuring functional play instead of pretend play and for being confounded with language (Casby, 1997). This study contributed to the literature by using age-appropriate toys to investigate elaborate pretend play and role play abilities in preschool children with and without SLI. Furthermore, scoring of children’s elaborate pretend play abilities did not rely on their language skills. Three different measures of pretend play were used: a standardized assessment of pretend play abilities, a role play activity, and a pretend play questionnaire. Consistent with previous literature, it was hypothesized that children with SLI would show less sophisticated forms of pretend play than their age-matched peers with TD.

#### 4.2.1 Child-Initiated Pretend Play Assessment

Pretend play is a cognitive skill that includes the ability to substitute objects and/or represent absent objects or properties. Moreover, pretend play captures the child’s ability to organize play actions logically and sequentially (Stagnitti, 2007). In this study, children with SLI did not differ from their peers with TD in their number of object substitutions. However, children with SLI were more likely than children with TD to remain in functional play and demonstrate problems in engaging in elaborate pretend play sequences during the ChIPPA assessment. For instance, children with SLI tended to reorganize the fence pieces multiple times without using the fence area later on for the animals, whereas their peers with TD quickly set up the fence pieces and continued to use this area as “farm” or “shelter” for the animals.
The ability of children with SLI to perform relatively well on object substitutions and to perform more poorly in elaborate pretend play might be somewhat surprising. However, it is consistent with previous studies that reported minor or no differences in object substitution abilities in children with and without SLI (e.g. Casby, 1997; Lombardino et al., 1986). Given that object substitution is one aspect that defines pretend play, this result might indicate that problems in pretend play may not be due to limited object substitution skills. It has been theorized that children with language impairments might be immature relative to their age-matched peers with TD, which could explain the observed differences in pretend play (Locke, 1994; Lombardino et al., 1986; Rescorla & Goossens, 1992). Rescorla and Goossens (1992) suggested that children with language impairment might experience a maturational delay in developing a complex system for symbol use. This slower maturational rate could then cause a delay in the achievement of other cognitive milestones, such as language or pretend play.

An alternative explanation might be that children with SLI have problems in accessing or retrieving play scripts (Leonard, 1998; Rescorla & Goossens, 1992). It has been proposed that children with SLI experience a general limitation in their processing capacities. That is, children with SLI have problems with encoding of information or with rapid and spontaneous retrieval of information due to limitations in memory, energy, or time (Leonard, 1998). Rescorla and Goossens (1992) raised the possibility that spontaneous retrieval of play scripts from memory might be more effortful for children with SLI than for children with TD. This could result in less sophisticated pretend play and less motivation to engage in pretend play. In turn, this may lead to less play experience and thus, play may fail to develop in complexity and flexibility. Play is viewed as being “the leading source of development in preschool years” (Vygotsky 1933/1966, p. 537); thus, limited participation in play might therefore have a negative influence on children’s development. In this view of limited information processing, language deficits are not the primary reason for problems in pretend play but deficient language skills may exacerbate the differences in pretend play abilities between children with and without SLI over time.

This study did not attempt to address the question of whether maturational delay or limitations in processing skills could explain the pretend play deficits seen in children with SLI. However, the language test employed (i.e., the SETK 3-5) included subtests that measured children’s ability to repeat nonwords and sentences. These subtests are thought to tap into children’s information-processing skills (Grimm et al., 2001) and may therefore distinguish children with
language impairments from children with TD (Conti-Ramsden & Hesketh, 2003; Graf, Evans, & Else-Quest, 2007). Consistent with previous literature (e.g., Graf et al., 2007) children with SLI in this study performed considerably poorer on these two subtests than their peers with TD. Poor test performance may indicate limited processing skills in children with SLI. However, poor test performance could also be attributed to a maturational delay as these two explanations are not mutually exclusive.

On the other hand, the observed differences in elaborate pretend play might be explained by limitations in narrative competence. Narrative skills emerge during the preschool period and refer to the ability to comprehend and tell stories (Engel, 2005; Pellegrini, 1985). Pretend play may facilitate the development of narrative skills, which, in turn, may support children’s ability to sequence play actions logically (Pellegrini, 1985; Stagnitti, 2007). This does not mean that young children tell complex narratives as is later expected from them in a school setting but rather that language during pretend play can reveal how events are related and how the actions of a character or toy figurine are motivated.

According to Lyytinen (1990), three-year-old children with TD relied heavily on language to describe objects and events during pretend play. Play might therefore provide a context for children to use language. Describing actions and events during play may help children to recognize that play events are related causally or temporally, which in turn might also help them to structure and organize their pretend play. On the other hand, toddlers with SLI displayed more functional play and less sophisticated pretend play behaviors than their age-matched peers with TD (Rescorla & Goossens, 1992). The differences in play may result in different needs to use language during play because functional play may require few clarifications. Not commenting on play or transformations in play, however, may limit the opportunity to practice narrative skills. Thus, children with SLI may not realize how play events can be linked, which in turn, may limit their ideas for pretend play scenarios (what happens and why).

Narrative skills were not assessed in this study nor were verbal pretend play explanations necessary in order to score within the normal range of the ChIPPA. However, the main outcome measure of the ChIPPA (i.e., the percentage of elaborate pretend play actions, PEPA) is based on the assumption that elaborate pretend play is characterized by detailed sequences of play actions that are organized logically (Stagnitti, 2007); therefore, PEPA might be an indicator of a
child’s underlying narrative-like competencies. The PEPA score indicates that a child knows how to link different play actions or events. In this study, children with SLI had a significantly lower PEPA score than their age-matched peers with TD. This shows that during solitary pretend play children with SLI were less able to connect different play events, which might also indicate limited narrative-like skills.

It is still unclear what could account for the differences in pretend play between children with and without SLI. One important aspect that might be of concern is the development of narrative skills. The ability to understand and tell narratives is crucial for academic success (Nicolopoulou, 2005; Stagnitti et al., 2000). Experimental studies have shown that pretend play may facilitate narrative skills in young school-age children (i.e., between six and eight years of age) (Baumer et al., 2005; Pellegrini, 1980). Future studies may extend these findings by including preschool children whose pretend play and narrative abilities are also assessed prior to intervention to investigate the direction of effect.

4.2.2 Role Play Activity

Given that children with SLI scored approximately 2 standard deviations below the mean on the ChIPPA and that the role play activity was a highly verbal activity, it was surprising that there were no significant group differences between children with and without SLI in the role play activity. However, there were trends that suggested possible differences in role play abilities across the two groups. Based on the number of coded segments, children with SLI spent significantly less time in the role play activity than children with TD. Further, they were somewhat more likely to spend time in Basic Pretend Play and Role Enactment and less time in Direct Role Play while the reverse was true for children with TD. This may suggest that there were group differences but that the chosen coding procedures may not have been sensitive enough to capture them.

For practical reasons, the videos files of the role play activity were segmented into time intervals. In each time interval, only the highest level of observed pretend play behavior was coded. In addition, it was also coded if the child was prompted by the experimenter. It is possible that verbal transcripts of the role play activity in combination with a more elaborate coding scheme would have been more sensitive to detect differences in children’s role play behaviors. Furthermore, it is possible that the involvement of the experimenter had an influence
on children’s pretend play (Pellegrini & Galda, 1993). The children were not prompted during pretend play assessment with the ChIPPA and significant group differences in elaborate pretend play between children with and without SLI were found. On the other hand, the experimenter modeled a pretend play story and continued to prompt children during the role play activity when necessary. This might have scaffolded children’s play and thereby minimized play differences across the two groups. However, children with and without SLI did not differ in the amount of prompts that they received. Analyzing verbal transcripts with a more elaborate definition of prompts (e.g., directive vs. conversational) might reveal differences in the quality of prompts that children received.

Finally, it is possible that observing child-child dyads during social pretend play would have yielded different results than adult-child dyads. However, it was not feasible in this study to observe child-child dyads because most children with SLI were recruited from a clinic with a broad catchment area. Therefore, familiar peers were not readily available. Future studies however, could compare role play performances in child-child dyads vs. adult-child dyads to investigate the effect that an adult might have on children’s play performances. However, when different partners are involved, they may introduce variations to the pretend play interaction, which could potentially become a confounding factor (DeKroon et al., 2002), whereas in this study, each child received a standard introduction to the role play activity. Therefore, each child should have had a similar opportunity to show his or her pretend play abilities.

On the other hand, this study was one of the first to have investigated different forms of pretend play in preschool children with SLI. If future studies replicate the findings that preschoolers with SLI show less sophisticated forms of pretend play when playing alone but perform comparably to their peers when interacting with a child or adult, this would provide support for the assumption that more competent partners may facilitate pretend play performance in children with SLI (Vygotsky, 1933/1966). This may then further support arguments for interventions in integrative settings in which peers with TD are taught how to scaffold play and interactions with less competent children with disabilities (e.g., Goldstein, English, Shafer, & Kaczmarek, 1997; Goldstein, Schneider, & Thiemann, 2007).

Alternatively, if children with SLI in this study had performed significantly different from their peers with TD, this would have improved our understanding of SLI. If children with SLI had
consistently displayed a significantly higher number of Basic Pretend Play or Role Enactment codes, it would have been evident that children with SLI are delayed in their development of role play abilities. This could have provided an explanation for why children with SLI might have difficulties engaging in social pretend play with their peers. That is, it might be that children with SLI are more likely to lack the ability than the motivation to engage in more sophisticated forms of pretend play. Moreover, distinct pretend play patterns in children with SLI might have provided a strong rationale for research (e.g., intervention studies) as well as for clinicians to provide targeted treatment to improve children’s pretend play abilities. Finally, if group play performances had been distinctly different and only a few children with SLI had performed well on the role play activity, it might have been possible to investigate their specific compensatory strategies that allowed them to perform well despite their deficient language skills. However, great variability in linguistic as well as nonlinguistic performance has been reported for children with SLI (e.g., Graf et al., 2007; Rice, Tomblin, Hoffman, Richman, & Marquis, 2004; Windsor, Milbrath, Carney, & Rakowski, 2001). Thus, the findings of variability in children’s role play performance are consistent with the heterogeneity within the population of children with SLI.

4.2.3 Pretend Play Questionnaire
The pretend play questionnaire was completed by the parents and was thought to provide information that might support the direct observations of their child’s play. Parents were asked to check the most appropriate answer from four response choices regarding different aspects of play. Response patterns were equally distributed; that is, parents did not always check the first response choice or always choose an answer that could be viewed as socially more desirable. However, the questionnaire did not reveal significant differences between children with and without SLI, even though children with SLI were rated somewhat lower by their parents than were children with TD. This might suggest that there were no distinct group differences in pretend play. On the other hand, it is possible that the response choices were not well explained or not distinct enough to differentiate the different levels of pretend play. It is also conceivable that not all aspects of pretend play are as easy to observe as object substitutions. Moreover, parents might be less familiar with evaluating their children’s play behaviors than, for example, evaluating their language or motor skills. The development of the latter skills is regularly
assessed by a pediatrician in a check-up that is mandatory for German parents and their children at certain ages. Play or pretend play abilities may not be assessed and parents may, therefore, pay less attention to this area of development.

4.2.4 Relations of Gender and SES to Pretend Play

Previous studies have indicated that children’s pretend play might differ depending on the child’s gender or familial SES status (Ayres & Levé, 2006; Ramsey, 2006). In this study, children’s pretend play did not differ as a function of gender. Toys in the ChIPPA were selected as being gender-neutral (Stagnitti, Rodger, & Clarke, 1997) and a similar attempt was made for the toys in the role play activity. Thus, it might be that toy selection was successful and that the toys were not more attractive for either gender, which otherwise could have influenced their play performance. Another explanation might be that the presence of a female experimenter could have influenced the child’s play behaviors. Although the experimenter did not instruct the children on how to play, it is possible that children behaved differently in this test situation (Pellegrini & Galda, 1993). Thus, it is conceivable that gender differences would have been evident if children had chosen the toys themselves and had been observed during free-play in their own environment.

Further, it has been proposed that SES may be related to children’s pretend play abilities because children of lower-income classes might have different access to toys or limited space or time to engage in play (Ramsey, 2006; Rubin et al., 1976). Although familial SES in this study was significantly different for children with and without SLI, SES was not significantly correlated with the dependent measures of pretend play. A possible explanation might be that although the familial SES index for children with SLI was lower than the SES for children with TD, children with SLI did not have a lower-class background. That is, the SES index of children with SLI was indicative of the middle class (Winkler, 1998). The majority of previous research focused on differences between lower or working class and middle class families (Fein & Stork, 1981; McLoyd, 1982; Rubin et al., 1976). Thus, the fact that children with SLI were part of the middle class might explain why there was no strong relation between pretend play and SES in this study.
4.3  Theory of Mind and Pretend Play

The final hypothesis was that children with a lower ToM composite score would also display less sophisticated forms of pretend play than children with a higher ToM score. In the zero-order correlations, PEPA and Direct Role Play were associated with the composite ToM score. However, partial correlations controlling for children’s language abilities and familial SES revealed that the composite ToM score was no longer associated with elaborated pretend play on the ChIPPA or the coding category Direct Role Play of the role play activity. This result was inconsistent with the last hypothesis and with previous studies that demonstrated an association between role play during social pretend play and ToM, independent of children’s language abilities (Astington & Jenkins, 1995; Jenkins & Astington, 2000; Nielsen & Dissanayake, 2000; Youngblade & Dunn, 1995). However, in previous studies, only children with TD participated. Therefore, the associations between ToM and pretend play were investigated for each language group separately. This revealed different patterns of associations for children with and without SLI.

In general, there are two possible interpretations of these conflicting findings. On the one hand, it is possible that the full sample partial correlations revealed the true picture of the relationship between ToM and pretend play. That is, there was no significant association between ToM and pretend play once language abilities and SES were taken into account. As proposed earlier, cognitive resources (e.g., processing skills) and/or representational abilities might be underlying skills that are necessary and common to language, ToM, and pretend play (Astington, Harris, & Olson, 1988; Piaget, 1951/1985). Further, it is conceivable that language and/or familial SES play mediating roles in the development of these three abilities. Thus, when language and SES were accounted for, there was no unique relationship between ToM and pretend play.

Previous studies that reported small- to medium-sized correlations between ToM and pretend play controlled for children’s language abilities and/or age (Astington & Jenkins, 1995; Jenkins & Astington, 2000; Nielsen & Dissanayake, 2000; Taylor & Carlson, 1997) but not for familial SES. It is therefore possible that methodological differences, such as controlling for SES and/or including an extended range of language, ToM, and pretend play performances, might explain the different findings. An alternative explanation might be that the two measures of pretend play used in this study (i.e., ChIPPA and role play activity) were not demanding with respect to
children’s ToM understanding. That is, children’s elaborate pretend play was assessed during solitary play and their role play abilities were observed while playing with the experimenter, who provided prompts. These tasks might not have required children to rely extensively on their ToM abilities.

In this case, the different patterns of partial correlations in children with and without SLI might be spurious and reflect the instability of small samples. Moreover, the extreme performances of a few children in either language group might have resulted in the varied patterns of association. Thus, caution is needed when interpreting the findings of the individual language groups given their smaller sample sizes.

However, Jenkins and Astington (2000) reported moderate associations between pretend play and ToM in a sample of only 20 children with TD. Moreover, with respect to the appropriateness of the pretend play measures used in this study, Taylor and Carlson (1997) also observed solitary pretend play in children with TD and found a significant association between play and ToM. Finally, Nielsen and Dissanayake (2000) observed child-adult dyads and parents were asked to play with their children as they usually would. It is conceivable that parents used clarification questions, such as “What is that?” or “What are you doing?” which were coded as prompts in the current study. Thus, parents may have prompted or scaffolded their children’s pretend play as well. Nevertheless, the authors reported significant correlations between ToM and pretend play independent of age. This raises the possibility that the lack of correlation between ToM and pretend play in the full sample is not necessarily attributable to the nature of the pretend play tasks employed here.

A second possibility is that the lack of association between ToM and pretend play in the full sample might be due to the different patterns of association in the two language groups which hide the true picture. In this alternative view, children with TD and with SLI may have approached the tasks differently, which would explain the different patterns of correlations. For instance, children with SLI might rely on different strategies during social pretend play or other social situations than their peers with TD. Rice, Sell, and Hadley (1991) reported that children with SLI were more likely to turn to or interact with adults than with their peers with TD. It is therefore conceivable that children with SLI relied on the social support of the experimenter during pretend play.
This idea is consistent with research suggesting that children with SLI have a greater need for adult support in different social situations than do children with TD. When a scaffolding adult was not present, preschoolers with SLI were found to be more withdrawn and less responsive to social bids than their peers with TD and they tended to display more passive behaviors during free play or recess (Fujiki et al., 2001; Guralnick, Connor, Hammond, Gottman, & Kinnish, 1996; Guralnick, Hammond, & Connor, 2006). These differences in social-emotional behaviors, which might be related to their language disabilities (Redmond & Rice, 1998), may limit the opportunity of children with SLI to engage in social pretend play and, consequently, limit their experiences with peers and play scripts. With less experience or practice, it may be more difficult for them to focus on pretend play activities while simultaneously representing their playmate’s mind. As a result, children with SLI may only focus on pretend play actions without acknowledging the playmate which then might lead to parallel or solitary play (DeKroon et al., 2002).

Alternatively, children with SLI may try to pay attention to the play ideas of their playmate. However, children with SLI are limited in their processing abilities and tend to react more slowly than their peers with TD (Graf et al., 2007; Leonard, 1998). It may therefore be more demanding for children with SLI to process all the play information provided (e.g., keeping track of changed object meanings and relatedness of events). They may not respond in a timely manner or they may not demonstrate play behaviors appropriate to the play context – both behaviors that may make it more likely for the social pretend play episode to falter.

Children with TD, on the other hand, may have had more opportunities to engage in social pretend play. Therefore, they may have had more exposure to a variety of different play scripts and more social experience in dealing with peers. Moreover, children with TD are, in general, not limited in their processing abilities. Thus, processing information concerning the pretend play while simultaneously representing their own and their playmate’s mental states may be less demanding for children with TD than for children with SLI. In this view, the subgroup findings for children with TD were consistent with previous studies (Astington & Jenkins, 1995; Jenkins & Astington, 2000) and provided further support for the hypothesis that children with better ToM understanding are better playmates during social pretend play than children with poorer ToM understanding. That is, a child who is better able to represent what his playmate knows and
informs him about the unfolding play events provides scaffolding, which increases the likelihood of successful social pretend play sequences.

Consequently, the subgroup findings for children with SLI might suggest that children with SLI show patterns of behaviors with respect to ToM and pretend play that might differ from their peers with TD. This finding has not been previously reported and might be a novel contribution. However, prospective studies are needed to replicate and extend these results. Future studies might include observations of social pretend play in children with SLI when playing with other children with and without SLI. In addition, it might be valuable to assess children’s processing abilities to determine if processing limitations are associated with ToM and/or pretend play in children with SLI. Finally, children with SLI were reported to be shy and withdrawn (Fujiki et al., 2001). It might therefore be informative to use behavioral rating scales, such as the Child Behavior Checklist (Achenbach, 1991) in order to take individual behavioral differences (e.g., shyness, passiveness) into consideration when interpreting data.

4.4 General Implications of Findings

4.4.1 Theoretical Implications

In children with TD, language, pretend play, and ToM develop in concert and each aspect seems to be related to the others. That is, language is associated with pretend play and may also facilitate ToM understanding (McCune-Nicolich, 1981; Milligan et al., 2007). Further, ToM and pretend play are related and pretend play may facilitate language abilities (e.g., narrative competence) (Astington & Jenkins, 1995; Pellegrini, 1985). However, in children with SLI, language development is delayed, which may be related to their development of pretend play and ToM. Children with SLI in this study not only performed poorly on ToM tasks, they also sometimes engaged in less sophisticated forms of pretend play than their peers with TD. This suggested that language, pretend play, and ToM do not develop independently of each other as some theoretical accounts suggest (Chomsky, 2005; Leslie, 1987). This may have consequences for the theoretical frameworks of ToM and SLI.

For instance, some researchers have argued that ToM is innate, domain-specific, and matures independently of external influences (e.g., Leslie, 1994; Leslie, Friedman, & German, 2004). In this extreme view, ToM is like a ‘separate module’ that develops independently of language or
other social influences. This account would have been supported if children with SLI in this study had performed well on the ToM tasks despite their language impairments. However, there was only one child with SLI that could be considered as supporting the account because she scored relatively high on ToM tasks despite having the weakest language abilities of the group. Perhaps individual or social differences (e.g., siblings) were influential in her ToM development. Nonetheless, this pattern of high ToM performance despite poor language skills describes only a single case. Finally, this account is not congruent with empirical findings that demonstrated a strong association between language and ToM in children with TD (Milligan et al., 2007), a finding that was replicated here.

On the other hand, one could argue that not only the language but also the ToM module is damaged in SLI and thus, resulted in poor ToM performance. However, this thought raises several issues. If, in some children with SLI, both modules are damaged, the poor performance of individual children could be explained. However, the majority of tested children with SLI experienced problems with ToM and, therefore, should have an impaired ToM module. Leslie (1994) proposed that impairments in ToM may be the possible cause of the symptoms described in children with autism. This would imply that children with SLI should present with more severe or additional social and pragmatic difficulties, which is not necessarily the case (Bishop & Norbury, 2002).

This study seems to be more supportive of theories that assume a cognitive (e.g., Bartsch & Wellman, 1995) or social-cognitive framework (e.g., Nelson, 2007) of ToM. In this view, the child is actively involved in reasoning and making sense of the social world. Language serves an important role in reasoning and making sense of the social world and pretend play might provide an additional context to learn about other’s mental states. Moreover, children do not develop cognitive abilities, such as ToM or pretend play, in isolation but within social contexts. In this study, there were no group differences in children’s role play abilities when playing with an adult. That is, children with SLI demonstrated better role play abilities than anticipated, given their low scores of elaborate pretend play on the ChIPPA. This may support Vygotsky’s assumption (1933/1966) that when children play with other, more competent partners, that children’s play might be more elaborate than during solitary play or when playing with less competent partners (DeKroon et al., 2002).
The poor ToM performance of children with SLI might be explained by having difficulty forming mental representations, which may delay overall ToM development. It has been hypothesized that children with SLI experience a general representational deficit, which could explain the observed problems in pretend play, language, and ToM (Leonard, 1998; Rice & Kemper, 1984). This hypothesis might have been better supported if children with SLI had consistently performed more poorly on all measures of pretend play. Children with SLI are often late in acquiring their first words as toddlers (Paul, 2001). Given that language is associated with play (McCune-Nicolich, 1981), a delay in language development may be associated with a delay in pretend play development. Previous studies found that young children with SLI were more likely to engage in nonplay behaviors or functional play than their age-matched peers with TD (Rescorla & Goossens, 1992). However, not engaging in elaborate pretend play may impede the development of ToM, as pretend play was a significant predictor of false belief performance in young preschoolers with TD (Youngblade & Dunn, 1995).

Furthermore, preschoolers with a better understanding of ToM were also better playmates during social role play (Astington & Jenkins, 1995). Children with SLI are not typically preferred playmates (Guralnick et al., 1996), which might be due to their limited abilities to engage in elaborate pretend play or their limited abilities to represent the mental states of their playmates, which, in turn, may negatively impact on social pretend play. Thus, in this view, a representational deficit would affect the development of the individual linguistic and nonlinguistic skills, which may also exert negative influences on each other.

In this study, however, children with SLI displayed an uneven profile in their linguistic and nonlinguistic performances, which poses a dilemma for the representational deficit hypothesis. Children with SLI performed poorly on ToM and on the ChIPPAt but there were no statistical group differences on the role play activity or the pretend play questionnaire. Moreover, there were no group differences on the nonverbal intelligence test and both the children with and without SLI showed average performance with a mean of 100 on the CPM, which is unusual. As a group, children with SLI have normal nonverbal skills; however, they tend to score in the low average range (e.g., Tomblin et al., 1997) and may show significant differences from age-matched peers.
It is possible that, if the children with SLI in this sample had shown low average nonverbal abilities, this might have resulted in distinct group differences on all pretend play measures. On the other hand, nonverbal intelligence tests have been suspected to assess a relative strength in children with SLI and to overestimate their overall nonlinguistic abilities (e.g., Hick, Botting, & Conti-Ramsden, 2005; Swisher, Plante, & Lowell, 1994). Children in this study were asked to identify a missing pattern to complete a larger visual pattern and did not participate in any other nonverbal task (e.g., visual memory or spatial rotation). Therefore, it is impossible to conclude that some nonlinguistic tasks might have been easier for children with SLI than other tasks. However, their relatively high nonverbal scores may support the hypothesis that the CPM assesses a relative strength in children with SLI.

This study suggested that the development of cognitive skills such as certain aspects of pretend play and ToM are affected in children with SLI. Problems in nonlinguistic or cognitive tasks have been described (e.g., Leonard, 1998); however, they are not considered in the definition of SLI. That is, children are commonly defined as having SLI when they experience severe language problems in the absence of serious hearing, neurological, or cognitive problems (Leonard, 1998). It seems, however, that problems are not specific to language development and that other cognitive aspects are also implicated. Thus, it might be necessary to re-think the construct of SLI.

4.4.2 Research Implications

This study was one of the first to have used the ToM scale as a broader measure of ToM in children with and without SLI, instead of measuring only false belief performance. The ToM scale allows children to demonstrate their knowledge with respect to different ToM concepts, including false beliefs. However, the addition of another false belief task to the ToM scale led to the unexpected finding that children’s false belief performance differed significantly across the two tasks. That is, although a meta-analysis concluded that type of task or type of question did not affect ToM performance (Wellman et al., 2001), children in both groups performed better when the target question was syntactically less complex. This suggests the need for future research that manipulates language complexity when assessing ToM in children with and without SLI. In addition, it is also important to counterbalance tasks when a number of different
false belief tasks are used, as the increase of false belief performance across the two tasks here may have also reflected a practice effect.

Previous studies that investigated play in children with SLI were criticized for being confounded with language and for assessing functional play and not sophisticated forms of pretend play (Casby, 1997). This study employed the ChIPPA, which is a newly developed test that provides the opportunity to assess different forms of play behaviors during solitary pretend play (e.g., behavioral, functional, and elaborate pretend play actions) (Stagnitti, 2007). This test uses minimal verbal instructions and scoring does not rely on children’s language abilities. Therefore, the ChIPPA could provide valuable information on play abilities in children with different forms of disabilities.

Moreover, the ChIPPA might also be an interesting research tool when investigating associations between pretend play and emergent literacy skills (i.e., reading and writing), which are thought to be linked (Harris, Kavanaugh, Wellman, & Hickling, 1993; Nicolopoulou, 2005; Vygotsky, 1933/1966). There is some evidence that pretend play may facilitate narrative skills in children with TD (Pellegrini & Galda, 1993). Narrative competence is part of emergent literacy skills, which are crucial for children’s eventual success in school (Nicolopoulou, 2005; Pellegrini, 1985; Stagnitti et al., 2000). Research has indicated that children with SLI are at risk for academic failure due to problems in reading and writing (Catts, Fey, Tomblin, & Zhang, 2002; Schuele, Spencer, Barako-Arndt, & Guillot, 2007). Thus, the ChIPPA could be used to investigate the relation between pretend play and literacy skills in children with SLI.

Finally, this study was the first to investigate the relationship between ToM and pretend play in children with and without SLI. There were no significant associations between ToM and sophisticated forms of pretend play in children with and without SLI after controlling for children’s language abilities and familial SES. However, when the two language groups were considered individually, different patterns of associations emerged for children with and without SLI. Moreover, performances were more heterogeneous in children with SLI than in children with TD. Some children with SLI did not engage in pretend play despite relatively high ToM scores, whereas others displayed sophisticated forms of pretend play even though their ToM composite scores were low. This warrants further investigation. Future studies might recruit a larger sample of children with SLI that includes older children because it is conceivable that
associations between ToM and pretend play change with age given children’s delay in language and ToM development. Given that children with SLI tend to prefer adults to peers (Rice et al., 1991), it might also be interesting to investigate the relationship between ToM and pretend play when children interact with different play partners. It is possible that the presence of the adult encouraged certain behaviors whereas children with SLI might behave differently when playing with familiar peers. These prospective studies may help to provide further evidence for or against an association between ToM and pretend play in children with SLI.

4.4.3 Practical Implications

Obviously, children with SLI have severe problems with language production and/or comprehension. Therefore, these areas will probably receive most of the attention of a speech-language pathologist (SLP). Moreover, the definition of SLI excludes other concomitant problems. Thus, an SLP may not necessarily expect and test for additional deficits in cognitive areas such as ToM or pretend play. However, this study suggests that children with SLI often experience these additional deficits. This implies that clinicians may have to broaden their assessment procedures to capture developmental deficits in other areas as well. The ChIPPA might be a valuable extension to the assessment procedures because it is easy to administer and online scoring is available.

Children with SLI might benefit from intervention sessions that involve pretend play. Play is often used as a child-appropriate means to target specific deficits (Stagnitti & Unsworth, 2000), for instance, practicing subject-verb agreement. However, pretend play may also provide a valuable context for broadening a child’s complexity and flexibility of play themes or for highlighting the different mental states of the characters involved in pretend play (Miller, 2006).

Moreover, given that pretend play may facilitate narrative skills in children with TD (Pellegrini & Galda, 1993), it is conceivable that pretend play might also facilitate narrative-like language in preschool children with SLI. For instance, enacting invented stories or those from storybooks in pretend play may provide a context for the SLP to highlight how play actions and events are causally or temporally linked. By engaging in pretend play, the child may have the opportunity to realize that some play events have to occur first or that others are a consequence of previous actions. This may also provide a scaffold for the use of narrative-like language (e.g., “First, he
runs and then she catches him”). Thereby, pretend play may improve narrative skills in children with SLI and may further help to reduce children’s risk for academic problems.

4.5 General Limitations

4.5.1 Limitations in Study Design

This study has several limitations. First of all, it is important to point out that this was an exploratory study and that a liberal probability level of \( \alpha = .05 \) was employed. This raises the risk of making Type I errors. Thus, caution is warranted in interpreting these results.

Another general limitation of this study was that the design was correlational in nature. Thus, it was not possible to establish cause and effect relations among the different cognitive variables. That is, in children with TD, the good performance in one cognitive domain (e.g., pretend play) cannot be attributed to the good abilities in another domain (e.g., ToM or language). Given that language abilities are impaired in children with SLI, it is possible that this deficiency affects the other two variables. However, it might also be the case that slow development in play and ToM affects language development or that problems in all three areas are due to another underlying problem, such as the ability to process information or form mental representations.

This study was also limited with respect to matching criteria. There was an attempt to match children with TD to those with SLI for the highest level of parental education as an estimate of family SES. However, parents of children with SLI were significantly more likely to have lower levels of education than parents of children with TD. SES was correlated with many different measures (e.g., language, ToM, and pretend play) and it was therefore important to control for SES when relationships were assessed. On the other hand, statistical analyses showed that when SES was entered as the covariate, familial SES did not account for the observed group differences in ToM or pretend play.

Further, children with SLI were matched to only one control group. This study demonstrated that children with SLI performed more poorly on several measures compared to age-matched peers. This raises the question of how they might perform relative to language-matched peers with TD, which might help to learn more about the relative severity of the delays in various cognitive abilities. In this study, group differences on different measures of ToM and pretend play varied in their effect sizes from small to large (\( r = .18 \) to .50). However, while keeping in
mind that children with SLI were only included when they scored 1.25 standard deviations below the mean on language measures, the group differences on the SETK were considerably larger in magnitude ($d = 3.66$). This suggests that children with SLI might perform better than language-matched peers with TD on measures of play and ToM. Nevertheless, the inclusion of a second, language-matched control group might be beneficial in future work.

4.5.2 Limitations in Generalizability
The generalizability of study results to other populations might also be limited in some respects. Although kindergartens for children with TD were chosen to potentially build a diverse pool of parents from different economic backgrounds (Landeshauptstadt Hannover, 2007), parents with higher levels of education appeared more likely to give their consent for research. As mentioned earlier, parents of children with TD were significantly more likely to have a university degree than the federal average of the adult population, whereas educational levels of parents of children with SLI were more representative of the population. Thus, performances of children with TD here may not be representative of children with TD in general.

Another question is whether the sample of children with SLI was also representative of the SLI population. German longitudinal studies demonstrated that children with developmental problems (e.g., language, cognitive, social-emotional, or motor skills) or other health-related concerns (e.g., obesity) are overrepresented in samples of families with low SES backgrounds (Bundesministerium für Gesundheit und Soziale Sicherung, 2006; Laucht, 2005). However, these studies did not specifically investigate the link between risk factors and specific disorders, such as language impairments. Thus, it remains unclear whether German-speaking children with SLI tend to come from lower SES backgrounds. In this study, the SES index of children with SLI was indicative of the middle class. Thus, it is possible that this sample of children with SLI might not have been typical of the SLI population. It might be that parents with higher levels of education were more likely to allow their children with SLI to participate in this study.

Another concern might be that children’s pretend play was assessed when interacting with an experimenter. The involvement of the experimenter in the role play activity may have influenced children’s play behaviors (Andresen, 2005; Pellegrini & Galda, 1993). Further, most children with SLI were tested in an unfamiliar environment, whereas children with TD were tested in a more familiar setting. Thus, it is conceivable that the observed play behaviors might
not have been typical for the children and might have been different when playing with other children during free play.

4.5.3 Limitations in Measurement

The pretend play questionnaire was developed because it was thought that parents’ perspectives on their children’s play might complement the observed play behaviors during the test sessions. Results indicated that children with and without SLI did not differ on the questionnaire once SES was controlled. A possible explanation might be that the response choices were not distinctive enough to capture the differences in children’s play behaviors. On the other hand, it might have been difficult for parents to report on children’s use of mental state verbs during play.

Another limitation might concern the coding procedures for the ChIPPA. As can be seen in Table 10, the mean of the elaborate pretend play score was below 100 for children with TD. Given the sample size, one might expect the group mean to be closer to 100. It might be that the coding procedures affected the group mean. According to the ChIPPA manual (Stagnitti, 2007), the child’s pretend play behaviors are supposed to be scored online. Then the elaborate pretend play score is calculated as a percentage of all elaborate actions over the sum of all behaviors times 100. However, online scoring was not feasible for this study because of the need for reliability testing. Therefore, video files of the ChIPPA session were analyzed later, as described in the Method section. The videos were segmented into 10-second time intervals and each segment was coded for every completed play behavior. It is possible that by coding every segment somewhat more play behaviors were coded overall than would have been the case during online scoring. This may have artificially inflated the denominator (i.e., the sum of all play behaviors in one play session) and thereby lowered the overall score of elaborated pretend play.

However, both groups of children should have been similarly affected by this procedure. Group means may have been somewhat higher during online scoring. Nevertheless, the group differences in elaborate pretend play would remain and children with SLI would score more than one standard deviation below the mean.
4.6 Contributions
This study applied a number of ToM tasks with varying degrees of linguistic and conceptual difficulty. Results added support to the sparse literature that children with SLI have poorer ToM understanding than age-matched children with TD (Farrant et al., 2006; Miller, 2001, 2004). This may suggest that children with SLI are delayed in their conceptual development of ToM but findings also indicated that these children followed a similar developmental trajectory to that for children with TD.

Further, this study used a variety of different play measures that were appropriate in their scope and use of materials for children between the ages of four and six. For instance, attractive toys were used (e.g., Playmobil) and basic pretend play skills were assessed without relying on children’s use of language. Moreover, this study extended previous research by also assessing the role play abilities of preschool children with SLI.

Results showed that relative to their peers with TD, children with SLI performed a similar number of object substitutions but performed more poorly on elaborate pretend play. This may indicate that differences in pretend play may be not due to limited abilities in object substitutions but may be due to problems in connecting pretend play events to form a logical play sequence. Further, children with and without SLI did not differ in their role play abilities. However, trends were in the expected direction, which might indicate insensitive coding procedures. These findings are important because play has been viewed as an important context for social-cognitive development in preschool children (e.g., Stagnitti, 2007; Vygotsky 1933/1966). Thus, limited participation in play might have a negative influence on children’s development. This study revealed that children with SLI may not only have language problems but also may experience difficulty with other cognitive skills, such as ToM and certain aspects of pretend play, which may have important implications for clinical assessment and intervention.

Finally, this study contributed to the literature by investigating the relationship between ToM and measures of sophisticated pretend play in children with and without SLI. After controlling for children’s language abilities and familial SES, there were no significant associations between ToM and both measures of pretend play in children with and without SLI. Supplementary analyses revealed different patterns of partial correlations for children with and
without SLI, perhaps suggesting that the two language groups behaved differently. ToM was positively associated with sophisticated forms of pretend play in children with TD, whereas negative correlations were found in children with SLI. The performance of a few children with SLI might have contributed to the negative relationship. That is, some children with SLI demonstrated an understanding of ToM but did not engage in pretend play, whereas others demonstrated sophisticated forms of pretend play despite poor ToM understanding. This warrants further investigation. To summarize, this study provided further evidence that children with SLI may not only have deficient language skills but may also experience limitations in their ToM and pretend play abilities. Finally, children with TD and SLI might rely on different strategies when engaging in pretend play, suggesting that it might be more demanding for children with SLI to engage in pretend play while representing their playmate’s mental states than it is for children with TD.
References


Appendices

Appendix A: English Translation of Questionnaire on Demographic Information

Dear Parents,
This questionnaire asks for some information about your child, who is participating in this study, and about you, as a parent. The information will be kept confidentially. Thank you for your cooperation.

Name of the child: ___________________________________________

Date of birth: ____________________

Name of person who completed this questionnaire _______________________

Relationship to the child ____________________________________________

1. What languages (if any) other than German does your child speak? ________________

2. Has your child ever been diagnosed as having:
   - General developmental delay (e.g. autism) No Yes
   - Neurological disorders (e.g., epilepsy) No Yes
   - Hearing impairments No Yes
   - Vision impairments No Yes

4. Did you experience complications during pregnancy or childbirth (e.g., preterm birth)? If yes, please specify ____________________________________________
5. Do you or did you have concerns about your child’s development (other than language) (e.g. attention, social behaviors)? If yes, please specify______________________________

6. Does your child have a “best friend”? No, not right now Yes

7. Does your child have siblings? No Yes

(If YES, please answer the next questions; if NO, please continue with question 11)

8. How old are the other siblings? ____________________________ years

9. If your child has more than one sibling, how old is the sibling your child prefers to play with? ____________________________ years

10. How many hours per week does your child play with this sibling?

__________________________ hours per week

11. How much opportunity does your child have with Playmobil toys?

My child has the opportunity to play with Playmobil at home or in kindergarten and likes to play with it

My child has the opportunity to play with Playmobil but prefers currently different toys

My child did not have the opportunity to play with Playmobil

12. Is there any problem that prevents your child from eating “Smarties” or “Haribo Gummibärchen” (e.g. allergies, diabetes)? No Yes

13. What is your marital status?

Married / living with a partner

Single parent
14. What is your highest level of education?

Mother:  
no school diploma + no formal job training  
graduating after 9th grade + no formal job training  
graduating after 9th grade + formal job training  
graduating after 10th grade + no formal job training  
graduating after 10th grade + formal job training  
graduating after 12th or 13th grade + no formal job training  
graduating after 12th or 13th grade + formal job training  
graduating after 12th or 13th grade + completed University degree  
Other, please specify ________________________________

Father:  
no school diploma + no formal job training  
graduating after 9th grade + no formal job training  
graduating after 9th grade + formal job training  
graduating after 10th grade + no formal job training  
graduating after 10th grade + formal job training  
graduating after 12th or 13th grade + no formal job training  
graduating after 12th or 13th grade + formal job training  
graduating after 12th or 13th grade + completed University degree  
Other, please specify ________________________________

14. What is your current occupation?

Mother: ________________________________

Father: ________________________________
Appendix B: English Translation of Pretend Play Questionnaire

Name of the child: ___________________________________________

Date of birth: ____________________

Dear Parents,

This questionnaire is about your child’s play behaviors that you observe at home. Please think back about the types of play behaviors you observed in the last 4-6 weeks. Think about what kinds of toys or with whom your child enjoyed playing and what happened in the play sessions. Children can show very different interests in play. This questionnaire tries to capture this variety – none of the answers provided here is either ‘correct’ or ‘incorrect’. Please, check only the one answer for each question that is most appropriate for your child.

1. What category of books does your child like best?
   - Picture books [score: 2]
   - Factual books (e.g., books describing ‘how things work’) [score: 3]
   - Fairy tales and imaginary stories [score: 4]
   - My child shows little interest in books [score: 1]

2. With whom does your child prefer to play?
   - Alone [score: 1]
   - 2-3 other children [score: 4]
   - Only with his/her best friend [score: 3]
   - Parents, other adults [score: 2]
3. **How often does your child play with other children after kindergarten?**
   - Seldom [score: 1]
   - Once a week [score: 2]
   - 2-3 times a week [score: 3]
   - Every day [score: 4]

4. **When your child is playing, how would you describe his/her verbal activity?**
   - My child uses an acting voice (high/low voice) to verbalize different characters [score: 3]
   - My child uses short sentences to describe a character, event, or action [score: 4]
   - My child uses a combination of sounds (e.g., toy car engine, crashing objects, sounds of eating) and single words [score: 2]
   - My child plays quietly [score: 1]

5. **What stories or events does your child act out in play?**
   - Playing everyday activities (e.g., fixing things, having a birthday party) [score: 2]
   - Playing home-like scenes (e.g., mother and child) [score: 1]
   - Playing fantasy characters (e.g., princess, pirates, superman) [score: 4]
   - Playing events in the environment (e.g., police, fighting, accidents, storms) [score: 3]

6. **How would you describe your child’s creativity?**
   - My child sometimes has new ideas and completes the new play sequence [score: 3]
   - My child frequently has novel ideas, which are completed in a play sequence [score: 4]
   - My child often engages in similar play activities (e.g., bathing/feeding doll, car race) [score: 1]
   - My child sometimes creates new play activities but does not always complete the new play sequence [score: 2]
7. **How would you describe your child’s role play with other children?**

(Role play means that children pretend to ‘be somebody else’; they create the role of characters e.g., doctor/nurse, teacher/student)

- My child sometimes engages in role play when playing with other children [score: 1]
- Within an imaginative play session, my child plays several roles/characters (e.g., mother, worker, cook) during one play session [score: 3]
- My child role plays with other children, but only for a short time and roles change very quickly [score: 2]
- Within an imaginative play session, my child creates one character in detail and maintains the same role throughout the play [score: 4]

8. **When your child is playing alone with a toy figurine (e.g., a doll, teddy, or Lego/Playmobil figurines), how would you describe her/his role play?**

- My child makes the figurine ‘talk’ and creates a personality for the toy figurine (e.g., toy figurine plays tricks or has certain strengths/magic powers) [score: 3]
- My child makes the toy figurine part of everyday activities (e.g., toy figurine is fed, bathed, treated when ‘sick’) [score: 2]
- My child makes the toy figurine an active part of role play (e.g., figurine fixes car, treats sick animals) and describes what is happening during play [score: 4]
- My child sometimes engages in role play using a toy figurine [score: 1]
Appendix C: English Translation of Informed Consent Form (Clinic)

Dear Parents,

My name is Melanie Stich. I am a German speech-language pathologist and I am currently a Doctoral student in the Department of Speech-Language Pathology at the University of Toronto in Toronto, Canada. I am conducting a research study with German-speaking children. This study will be conducted in cooperation with two agencies: the Petri-Kindergarten, Hannover and the Clinic of Phoniatrics and Pedaudiology, Medizinische Hochschule Hannover. I invite you and your child to participate. It is important that you know that taking part in this study is voluntary. You or your child can withdraw from the study at any time. There will be no disadvantages or penalties if you decide to withdraw. The study is described below. This description tells you about what will happen during the study. It will also address any concerns that you might have about risks or discomfort. If you have any questions that are not answered here, please contact me (Melanie.stich@utoronto.ca). I will be happy to discuss any concerns with you.

Reasons for the Study

I would like to find out how well children with and without language problems can understand what other people know, feel, or believe. Further, I would like to find out how this understanding relates to children’s ability to play. Therefore, 40 preschool children with and without language problems will be invited to participate in this study. Your child has been diagnosed as having language problems. Therefore, I would like to invite you and your child to participate in this study.

What do you have to do?

I invite you and your child to participate in several tasks. There will be two or three sessions of testing, depending on your child specific needs. Overall, testing will take approximately 2 hours. First, your child will be tested to see if she/he can participate in the study. The tests described below are part of the standard procedure of the clinic.

1. Your child will participate in a hearing screening. Your child will receive headphones and have to move wooden blocks into a basket as soon as she/he hears a tone. This task will take about 10 minutes. If your child has problems with hearing, your doctor in the
clinic will initiate a complete hearing evaluation. If your child has problems with hearing, she/he cannot take part in the study.

2. If your child passes the hearing screening, she/he will participate in another task. Your child will recognize objects or patterns that match each other or do not belong together. This task will take approximately 15 minutes. If your child scores within age expectations, she/he can take part in the study.

3. Then, your child will participate in a language test. Your child will label pictures, repeat words and sentences, and act out instructions (e.g., “Give me the red pencil”). This test will take approximately 45 minutes.

If your child is eligible to participate in the study, there are three more tasks that are related only to the study. If possible, the assessment of these tasks will be combined with other appointments that you may have in the clinic.

4. In this task, I will tell your child 8 short stories using Playmobil figurines and pictures. Then your child will answer questions about what the main character in the story wants, knows, or believes. This task will take approximately 20 minutes.

5. Your child will participate in a test that assesses play. Your child will receive two different sets of toys and we will play together with these toys in whatever ways your child chooses. This session will take 30 minutes and will be videotaped. This is necessary to analyze the play later on.

6. Finally, your child and I will play with Playmobil. Your child will receive a box with different toys from Playmobil and I will ask your child to think of a story that we then could play with the toys. This session will also be videotaped and will take about 15 minutes. One copy of this play activity is yours to keep.

I also invite you to participate in this study. You will be asked to fill out two questionnaires. The first questionnaire asks you to provide some information about your child’s development. It also contains questions about your family and about you, as parents. The second questionnaire is about your child’s play behaviors that you can observe at home. This task should approximately take 20 minutes of your time.
What are the risks and benefits?

I will ask your child to agree to participate before any testing takes place. It is possible that your child may feel uncomfortable in the beginning of the test session. I will give your child time to become familiar with the new situation and I will explain to her/him what is happening next. If you or your child wish, you can accompany your child during the test session. If your child keeps feeling uncomfortable, I will not continue with testing.

The possible benefit of this research is that you get a report of your child’s play development. Playing alone and with other children provides an important context for your child to learn. If your child seems to have problems with play, you can receive referral to an occupational therapist, if you wish. It might also be that you do not benefit directly from this study. However, the results might help to gain knowledge that will benefit others.

Compensation

Your child will receive candy during the tasks and stickers at the end of the sessions to compensate her/him for her/his time and help. Your child will also receive a copy of her/his Playmobil play activity. For any additional trip that is due to your participation in the study, you will receive EUR 10,-.

Confidentiality

The results of the study may be published, but your child’s name will not be used. Any information obtained during this study will be kept strictly confidential. Each participant will be assigned a numeric code. It will not be possible to link any information with your or your child’s identity. The test protocols and the code will be stored in a secure and locked location. Only the members of the research team will have access to the data. The protocol and code will be stored for five years after publication and then destroyed with a document shredder. The database without any identifying information will be kept indefinitely.

Your child will be videotaped during the play sessions. The tape will get a numeric code and there will be no identifying information on the tape. The tape will not be published without consent. The tapes will be stored for five years after publication and then destroyed.
**Participation**

Participation in this research is voluntary. You or your child can withdraw from the study at any time, even after giving consent/assent. You will receive a copy of this consent form for your records.

By signing the next page,

- you give permission for your child to participate in the study described above.
- you agree to participate in the study described above

**Questions**

If you have any questions or concerns, you can contact me (Melanie.stich@utoronto.ca) or Professor Ptok at (0511) 532-9104; mailing address Phoniatrie, OE 6510, Carl-Neuberg-Str. 1, 30625 Hannover. If you have questions about your or your child’s rights as a participant in this research, or if you feel that your child has been placed at risk, you can contact (contact person in the Ethics Office, Teaching Hospital).

This information is yours to keep for your records.
Consent Form

I have had the opportunity to read the explanation about this study on the previous pages. I have had the opportunity to discuss any concerns. My questions have been answered to my satisfaction. I give permission for my child to participate in this study. I also agree to participate in this study. I understand that our participation is voluntary. My child and I can withdraw from the study at any time, even after giving consent/assent. If I withdraw, there will be no disadvantages or penalties for me or my child.

Please sign and return the form to the clinic with the envelope provided.

__________________________________________________  _________________
Parent Signature   Printed Name     Date

_____________________________________________________________________________
Print Child’s Name       Date of Birth

By signing below, I do not agree that my child and I participate in this study.

__________________________________________________  _________________
Parent Signature       Date
Appendix D: English Translation of Informed Consent Form (Kindergarten)

Dear Parents,

My name is Melanie Stich. I am a German speech-language pathologist and currently, I am a Doctoral student in the Department of Speech-Language Pathology at the University of Toronto in Toronto, Canada. I am conducting a research study with German-speaking children. This study will be conducted in cooperation with two agencies: the Petri-Kindergarten, Hannover and the Clinic of Phoniatics and Pedaudiology, Medizinische Hochschule Hannover. I invite you and your child to participate. It is important that you know that taking part in this study is voluntary. You or your child can withdraw from the study at any time. There will be no disadvantages or penalties if you decide to withdraw. The study is described below. This description tells you about what will happen during the study. It will also address any concerns that you might have about risks or discomfort. If you have any questions that are not answered here, please contact me (Melanie.stich@utoronto.ca). I will be happy to discuss any concerns with you.

Reasons for the Study

I would like to find out how well children with and without language problems can understand what other people know, feel, or believe. Further, I would like to find out how this understanding relates to children’s ability to play. Therefore, 40 preschool children with and without language problems will be invited to participate in this study. I invite you because I assume that your child will have normal language abilities.

What do you have to do?

I invite you and your child to participate in several tasks. There will be two or three sessions of testing, which will take place in the kindergarten. The number of sessions depends on your child specific needs. Overall, testing will take approximately 2 hours. First, your child will be tested to see if she/he can participate in the study.

1. Your child will participate in a hearing screening. Your child will receive headphones and have to move wooden blocks into a basket as soon as she/he hears a tone. This task
will take about 10 minutes. If your child has problems with hearing, she/he cannot take part in the study.

2. If your child passes the hearing screening, she/he will participate in another task. Your child will recognize objects or patterns that match each other or do not belong together. This task will take approximately 15 minutes. If your child scores within age expectations, she/he can take part in the study.

3. Then, your child will participate in a language test. Your child will label pictures, repeat words and sentences, and act out instructions (e.g., “Give me the red pencil”). This test will take approximately 45 minutes.

If your child scores below age expectations on any of the tests, you receive the results. You and your pediatrician/family doctor can then initiate further assessment. If your child is eligible to participate in the study, there are three more tasks.

4. In this task, I will tell your child 8 short stories using Playmobil figurines and pictures. Then your child will answer questions about what the main character in the story wants, knows, or believes. This task will take approximately 20 minutes.

5. Your child will participate in a test that assesses play. Your child will receive two different sets of toys and we will play together with these toys in whatever ways your child chooses. This session will take 30 minutes and will be videotaped. This is necessary to analyze the play later on.

6. Finally, your child and I will play with Playmobil. Your child will receive a box with different toys from Playmobil and I will ask your child to think of a story that we then could play with the toys. This session will also be videotaped and will take about 15 minutes. One copy of this play activity is yours to keep.

I also invite you to participate in this study. You will be asked to fill out two questionnaires. The first questionnaire asks you to provide some information about your child’s development. It also contains questions about your family and about you, as parents. The second questionnaire is about your child’s play behaviors that you can observe at home. This task should approximately take 20 minutes of your time.
What are the risks and benefits?

I will ask your child to agree to participate before any testing takes place. It is possible that your child may feel uncomfortable in the beginning of the test session. I will give your child time to become familiar with the new situation and I will explain to her/him what is happening next. If you or your child wish, you can accompany your child during the test session. If your child keeps feeling uncomfortable, I will not continue with testing.

The possible benefit of this research is that you get a report of your child’s hearing status and her/his development of language and play. If your child performs below expectations, you will be informed immediately. You receive the results and your pediatrician/family doctor can then initiate further assessment. It might also be that you do not benefit directly from this study. However, the results might help to gain knowledge that will benefit others.

Compensation

Your child will receive candy during the tasks and stickers at the end of the sessions to compensate her/him for her/his time and help. Your child will also receive a copy of her/his Playmobil play activity.

Confidentiality

The results of the study may be published, but your child’s name will not be used. Any information obtained during this study will be kept strictly confidential. Each participant will be assigned a numeric code. It will not be possible to link any information with your or your child’s identity. The test protocols and the code will be stored in a secure and locked location. Only the members of the research team will have access to the data. The protocol and code will be stored for five years after publication and then destroyed with a document shredder. The database without any identifying information will be kept indefinitely.

Your child will be videotaped during the play sessions. The tape will get a numeric code and there will be no identifying information on the tape. The tape will not be published without consent. The tapes will be stored for five years after publication and then destroyed.
Participation

Participation in this research is voluntary. You or your child can withdraw from the study at any time, even after giving consent/assent. You will receive a copy of this consent form for your records.

By signing the next page,

- you give permission for your child to participate in the study described above.
- you agree to participate in the study described above

Questions

If you have any questions or concerns, you can contact me (Melanie.stich@utoronto.ca) or Professor Ptok at (0511) 532-9104; mailing address Phoniatrie, OE 6510, Carl-Neuberg-Str. 1, 30625 Hannover. If you have questions about your or your child’s rights as a participant in this research, or if you feel that your child has been placed at risk, you can contact (contact person in the Ethics Office, Teaching Hospital).

This information is yours to keep for your records.
Consent form

I have had the opportunity to read the explanation about this study on the previous three pages. I have had the opportunity to discuss any concerns. My questions have been answered to my satisfaction. I give permission for my child to participate in this study. I also agree to participate in this study. I understand that our participation is voluntary. My child and I can withdraw from the study at any time, even after giving consent/assent. If I withdraw, there will be no disadvantages or penalties for me or my child.

Please sign and return this form to Mrs/Mr ________ (director of the kindergarten).

________________________ ________________________ ___________
Parent Signature Printed Name Date

_____________________________________________________________________________
Print Child’s Name Date of Birth

By signing below, I do not agree that my child and I participate in this study.

_____________________________________________________________________________
Parent Signature Date
Appendix E: English Translation of Verbal Assent Form (Child)

The investigator will ask the child:

Here, I have some pictures and toys. I need your help to answer some questions and I would like to play with you. Would you be willing to help me?

Did the child assent to the task? □□ □□ □□ □□ YES NO

(A positive answer will be considered as an assent to participate in the task)

________________________________________
Signature of the Investigator Date
Appendix F: English Translation of Theory of Mind Tasks

1.) Not-Own Desire

This is Lars. It is time for a break. Lars wants to eat something. Here are two foods, a carrot and a cookie.

Pretest/Control: Which do you like best? Do you like a carrot or do you like a cookie?

That’s a good choice. But Lars doesn’t like [child’s choice]. He likes [opposite of child’s choice]. He loves to eat [opposite of child’s choice] best of all. So now Lars can choose only one food.

Target: Which will Lars choose?

Forced choice: Will Lars choose a carrot or a cookie?

(Correct answer = food the Lars likes, always opposite to child’s own preference)

2.) Not-Own Belief

This is Julia. Julia wants her cat. The cat is hiding. It could be in the bushes or it could be in the garage.

Pretest/Control: Where do you think the cat is? In the bushes or in the garage?

Well, that’s a good idea. But Julia thinks the cat is in [opposite of child’s choice]

Target: Where will Julia look for her cat?

Forced choice: Will she look in the garage or the bushes?

(Correct answer = opposite place to child’s own belief)

3.) Knowledge Access

Here is a little box. What do you think is in it?

That’s a good guess. Let’s open it... Oh, look! There is a dog in it!
Control: So what is in the box? [dog]

This is Anna. Anna has never seen this box before. She has never opened it.

Target: So does Anna know what is in the box? [no]

Control: Has Anna looked in this box? [no]

4.) Unexpected Contents

Here is a Smarties-box. What do you think is in it? [Smarties]

Let’s look in the box.... Oh! There is a pig in it.

Control: Okay... so what is in the box? [pig]

This is Lukas. Lukas has never looked in this box.

Target: So... what does Lukas think is in the box? [Smarties]

Forced choice: Smarties or pig?

Control: Did he look in the box? [no]

5.) Change-of-location

This is Tina. Tina ate some Smarties. She puts her last Smartie in a box. Then she goes out to play. This is Tina’s mom. The mom takes the smartie and puts it in a tin.

Control: Where did Tina leave the Smartie? [box]

Here comes Tina back and says: “I want to eat Smarties!”

Target: Where will Tina look for her Smartie? [box]

Control: Where is the Smartie really? [tin]

Appearance-Reality Emotion Scale Training, using Smileys
I’m going to tell you a story of a boy. The boy can look happy [point to happy Smiley]. He could also be sad [point]. Or he could be neither happy nor sad, he could be “in between” [point to neutral smiley].

Can you show me the face where the boy is:

sad?

neither happy nor sad?

happy?

6.) Appearance-Reality Emotion

This is the story of Tim. Tim’s aunt is just back from a trip. She promised Tim to bring him a toy car.

But: She brought Tim a book instead. Tim doesn’t like books. Tim really wanted a toy car.

But: Tim cannot show how he feels. If his aunt knew how he felt, she would never bring him anything again.

Control: What did Tim’s aunt bring? [book]

Control: What would Tim’s aunt do, if she knew how he really felt? [don’t bring gifts]

Target 1: So... how does Tim really feel when he gets the book from his aunt? Is he happy, sad or neither happy nor sad? [neutral or sad]

Target 2: What kind of face will Tim try to make when he gets the book from his aunt? A happy, sad, or neither happy nor sad face? [happy or neutral]

(correct answer has to indicate a difference (e.g., feels sad, shows neutral face or feels neutral, shows a happy face)

7.) Faux-pas: Violin

This is Lisa and this is her sister Sara. Sara has just finished her violin lesson, which she really enjoys. Sara really loves playing the violin and she thinks that it’s a really nice instrument.
Anyway, they are walking to the playground when they bump into Lisa’s friend, Dennis. Sara has never met Dennis before.

“Hello”, says Dennis, “I’m just back from football practice.”

Lisa asks him: “Are you going home now?”

“No”, says Dennis, “I’m going to meet my sister. She’s really boring because she plays the violin. Yuck, I think the violin is horrible!”

**Control:** Did someone say something they shouldn’t have?  [Yes]

**Control:** Who was it?  [Dennis]

What did he say?

**Target:** Did Dennis know Sara loves playing violin?  [no]

**Target:** What does Sara think that Dennis thinks about people playing violin?

**Target:** Did Dennis hurt Sara’s feelings on purpose?  [no]

**Justification**: How do you know?

(*question was asked but not scored because no child gave a correct answer.)*

---

8.) **Faux-pas: Rocket**

This is Thomas and this is his friend Sven. Sven has painted a picture of a rocket for a special exhibition in the school. Sven is really proud of his picture and thinks it’s great. Anyway, they are walking home from school when they bump into Thomas’ friend, Annika. Annika has never met Sven before.

“Hi”, says Annika.

Thomas asks her: “Hi Annika, have you been to the exhibition in the school?”

“Yes”, says Annika. “It’s quite nice, except there’s that really horrible painting of a rocket. I think it’s a really bad painting!”

**Control:** Did someone say something they shouldn’t have?  [Yes]

**Control:** Who was it?  [Annika]
What did Annika say?

**Target**: Did Annika know it was Sven who painted the picture?  [no]

**Target**: What does Sven think that Annika thinks about his painting?

**Target**: Did Annika hurt Sven’s feelings on purpose?  [no]

*Justification*: How do you know?
Appendix G: English Translation of Story Description for Role Play Activity

Introducing the play theme and setting up the play scene

Look, I have lot of toys in this box here. Look, there are many animals. I thought we could play zoo, OK? For fun, this could be the fence for the animals. And this could be the ticket box where we buy the tickets for the zoo, OK? We could pretend that this is a tower.

These animals belong to our zoo. Are there any more animals that we have in our zoo?

Who is going to the zoo? Who do you want to play/ do you want to be?

OK, you are xxx. This is the xxx. I’m going to be your xxx and this is our car.

Who is going to sell us the tickets?

Look, for fun this in my baby chick that wants to come with us to the zoo. Can I take my baby chick with me? OK, the baby chick hops onto the car and comes with us.

OK, let’s go to the zoo.

Ticket sale conversation

Hello, we would like to go into the zoo!

How many tickets do you need? – We are xxx people.

OK, that is xxx EURO. Have fun! – Thank you!

Events happening in the zoo

Let’s go to the petting zoo!

While we are at the petting zoo, no one looks after the baby chick and the baby chick walks off. And we don’t see it because we are petting the animals. The baby chick gets trapped in the hole of the tower.

Then we come back from the petting zoo: Oh, my baby chick is gone! I can’t find it! It’s nowhere! Can you please help me to find my baby chick?

Ah, thank you! You’ve found my baby chick! Thank you so much! Now, I will put it here in this box so that it cannot run away again.

OK, what else could happen in the zoo? What are we playing now?