Published in final edited form as:
J Neurolinguistics. 2011 July 1; 24(4): 476–496. doi:10.1016/j.jneuroling.2011.03.001.

The Relationship between Syntactic Development and Theory of Mind: Evidence from a Small-Population Study of a Developmental Language Disorder

Natalia Rakhlin[1], **Sergey A. Kornilov**[1,2,3], **Jodi Reich**[1], **Maria Babyonyshev**[1], **Roman A. Koposov**[4,5], and **Elena L. Grigorenko**[1,3,6]

Natalia Rakhlin: natalia.rakhlin@yale.edu; Sergey A. Kornilov: sa.kornilov@gmail.com; Jodi Reich: jodi.reich@yale.edu; Maria Babyonyshev: maria.babyonyshev@yale.edu; Roman A. Koposov: roman.koposov@uit.no; Elena L. Grigorenko: elena.grigorenko@yale.edu

[1] Yale University, New Haven, Connecticut

[2] University of Connecticut, Storrs, Connecticut

[3] Moscow State University, Moscow, Russian Federation

[4] Northern State Medical University, Arkhangelsk, Russian Federation

[5] University of Tromso, Norway

[6] Columbia University, New York, New York

The relationship between the development of language and Theory of Mind in children of preschool age has been the subject of much debate. Theory of Mind (henceforth, ToM) is conceptualized as a capacity to perceive others in terms of their mental states (i.e., desires, emotions, intentions and beliefs), and to understand that the actions of others are not merely automatic responses to outside stimuli, but are motivated by their mental states (Premack & Woodruff, 1978). The main challenge is to explain why children under 4–5 years of age find it difficult to deal with certain ToM-related tasks but not with others, and whether the reason for the late acquisition of some ToM-related skills is that they require a certain, relatively advanced, stage of language development. The issue that is subject to a particularly vigorous debate involves the question of what makes a task of predicting the behavior of another person based on that person's false belief difficult for young children and the role of syntactic development in children's mastery of this skill.

In the current study, we investigated the relationship between language development and false-belief reasoning and sought evidence of this relationship from a small Russian-speaking population characterized by a high rate of a developmental language disorder (henceforth, DLD). The main goal of the study was, first, to understand the respective roles of general cognitive functioning and general language development in children's performance on a standard task of false belief and, second, to see which aspects of language development contribute uniquely to the likelihood of success on such a task. In the remainder of the article, we will 1) talk about the role of false belief reasoning in probing an

Correspondence concerning this article should be addressed to Elena L. Grigorenko, Child Study Center, 230 South Frontage Road, Yale University, New Haven, CT 06519-1124. Phone: 203.737.2316; Fax: 203.785.3002 elena.grigorenko@yale.edu.

individual's ToM capacity, 2) discuss the semantic and syntactic properties of belief ascriptions contributing to the complexity of false belief tasks and explain why acquiring linguistic knowledge of these properties may be necessary for being able to cope with the demands of such tasks, 3) review the existing theories of ToM development and studies of false belief acquisition in typically developing children and clinical populations, and 4) present the results of our study, the goals of which were described above.

1. The Role of False Belief in Identifying Representational Theory of Mind

Recent advances in infant cognition, as well as evidence from brain lesion and neuroimaging studies, indicate that ToM is a complex capacity with multiple basic components, supported by a widely distributed neural system, some of which are present in humans as early as infancy (for a review of behavior studies see Apperly & Butterfill, 2009; Yamaguchi et al., 2009; for a review of the brain literature see Carrington & Bailey, 2009; Segal & Varley, 2002). The capacity to represent a third person's false belief (henceforth, FB), generally considered to be indicative of the successful acquisition of representational ToM (e.g., Dennett, 1978; Wellman, 2002; but see Bloom & German, 2000 for a counterargument), is not typically expected in children until they reach the age of 4 or 5 (e.g., Gopnik & Astington, 1988; Flavel, Green, & Flavell, 1986; Perner, Leekam, & Wimmer, 1987; Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). However, there has been some evidence that FB understanding can be facilitated in children as young as under 3 years of age by lowering the experimental demands suggesting that even younger children possess certain ToM related capabilities (e.g., Chandler, Fritz, & Hala, 1989; Surian, Caldi, & Sperber, 2007; Onishi & Baillargeon, 2005). Such conflicting findings have led to a considerable debate on the significance of FB understanding in the human capacity to reason about minds, componential skills involved in full-fledged FB reasoning, and the role of language in its development.

The importance of FB reasoning in establishing a developed ToM was first addressed in the discussion generated by the seminal paper by Premack and Woodruff (1978), which investigated the ToM capacity of chimpanzees. In a response to their article, Dennett (1978), and others since then, pointed out that in order to demonstrate that individual *A* possesses the ability to represent the content of others' minds, it is not enough to show that *A* can predict the actions of individual *B*. In many cases, correct behavior predictions can be made without understanding mental states, but by simply observing the actual state of the world. Therefore, in order to verify whether an individual is able to reason about others' minds, it is necessary that the measure probes genuine belief attribution and not simply an ability to read situational clues. Since predicting behavior based on a mistaken belief meets this requirement, it has become accepted as a litmus test of a developed ToM.

Even though children under 4–5 years of age cannot cope with standard FB tasks, research has shown that by the same age they reach other milestones of psychological reasoning. Thus, we know that 3-year-old children are successful with some experimental tasks involving pretense and deception (Chandler et al., 1989; Leslie, 1987), and that 2 and 3 year old children can correctly attribute ignorance (Hogrefe, Wimmer, & Perner, 1986), as well as emotions and beliefs in spontaneous speech (Bartsch & Wellman, 1995), long before they are able to explicitly reason about mistaken belief.

This discrepancy led to questioning whether children at the pre-FB stage are truly capable of constructing representations of mental states or instead use some form of quasi-mental-state attributions. For example, under one proposal, at the initial stage of belief attribution, children use a "realist" strategy, reporting current reality instead of belief (Mitchell, 1996).

Under this and other similar proposals, the early "belief" attributions would not have the semantics of belief reports, but would possess some simpler semantic form.

The complexity of genuine belief reports stems from their semantic as well as syntactic properties. Verbs like *think*, *know*, *expect*, *regret*, and so forth express propositional attitudes and denote a relation between a person (the subject) and a proposition (the embedded clause). Thus, a sentence like "Peter believes that Susan is a spy" expresses Peter's attitude towards the proposition "Susan is a spy", namely his belief that this proposition is true. For the attitude ascription to be true, the embedded proposition ("Susan is a spy") doesn't have to be true, but only the subject's attitude towards the proposition expressed by the matrix verb (i.e., it must be true that Peter indeed holds such a belief). This introduces a degree of extra semantic complexity: in order to interpret belief ascriptions, one has to evaluate the truth of the matrix clause ("Peter believes …") in the actual world and the embedded proposition ("Susan is a spy") in a domain distinct from the actual world, commonly conceptualized as the set of all possible worlds compatible with the subject's beliefs (e.g., Hintikka, 1969).[1]

Given children's difficulty with FB ascription, it is quite plausible that until a certain point in language development, children may not be able to construct such a complex semantic form and instead use a simpler semantic representation with the actual world being the evaluation world for both the embedded and matrix propositions.

In addition to their semantic complexity, propositional attitude reports, such as belief attributions, have an extra degree of syntactic complexity because they contain a full tensed clause embedded within the matrix clause (Pinker & Bloom, 1990; Tager-Flusberg, 1997). Consequently, we can suppose that the acquisition of the complex syntax of sentence embedding is a necessary prerequisite for being able to report propositional attitudes. Although we cannot, without further research, separate the role of semantic as opposed to syntactic complexity in children's difficulty with FB, we can assume that on the one hand, the formal properties of propositional attitude reports present a high degree of complexity, and on the other hand, provide representational means for explicit ToM reasoning.[2]

2. Theories of False-Belief Development

Despite some empirical inconsistencies, there is converging evidence that FB performance shows a uniform developmental pattern across different languages and various task manipulations. Thus, a meta-analysis of 178 studies by Wellman and colleagues (2001) revealed that children undergo a systematic change from below-chance performance to above-chance performance on tasks of FB reasoning between the ages of 3 and 5. Namely, they found that at 44 months of age, children were only about 50% correct in their responses, while at 56-months, their performance improved to 75% correct (Wellman et al., 2001). This finding indicates that FB reasoning undergoes a genuine conceptual change in the preschool years, and cannot be explained as an experimental effect.

[1]Possible world (or intensional) semantics is a formal semantic model developed to account for how language represents displacement (i.e., its ability to express information removed from the present situation). It has been used to analyze a wide variety of linguistic phenomena, including modality, conditionals, tense and temporal adverbs, obligation, and reports of informational and cognitive content, such as belief reports.

[2]There have been several proposals about the exact nature of the connection between language (namely its formal aspects) and ToM. One possibility is that language assists in the representational re-description of knowledge (Karmiloff-Smith, 1992); i.e., implicit knowledge gets re-described using linguistic representations for explicit reasoning. Another suggested alternative is that language serves as a medium for combining and integrating outputs from distinct mental modules (Spelke, 2003). See de Villers (2007) for a review of the various views.

Given this developmental change, it is important to explain how the child achieves the adult-like stage of FB reasoning. Until fairly recently, the major approaches to this issue formed two broadly defined views, which can be categorized as the constructivist and modularity views respectively. Such approaches as the so-called Theory theory (e.g., Gopnik & Wellman, 1994) or Simulation theory (e.g., Harris, 1996, 2009) emphasize the role of social experience and a process of either theory construction or analogy with one's own mental life in coming to understand other minds. On the other hand, according to the modularity approach, ToM capacity constitutes an innate module, which gets turned on by a genetically predetermined time-switch (e.g., Baron-Cohen, 1995; Fodor, 1992). In addition, there are those who propose that the change in children's ability to engage in complex reasoning about the mental states of others is due to the maturation of their information processing abilities, namely their short-term memory (WM) (Case, 1989; Fodor, 1992; Frye, Zelazo, & Palfai, 1995; Keenan, Olson, & Marini, 1998; Olson, 1989, 1993).

The main challenge to these approaches comes from studies of special populations, in particular deaf children. Thus, research showed that deaf children's performance on FB tasks is predicted by their language development, and that on tasks of FB native signers perform better than late signers. Moreover, deaf children showed comparable levels of performance on FB tasks regardless of whether the tasks were verbal, less verbal or nearly non-verbal (e.g., Peterson & Siegal, 1999; Pyers & Senghas, 2009; Schick et al., 2007; de Villiers, 2005; Woolfe, Want, & Siegal, 2002). Similar observations of a connection between language development and FB were reported for other special populations, such as children with autism spectrum disorder (Happe, 1995; Tager-Flusberg & Joseph, 2005), autism as compared with Asperger's syndrome (Paynter & Peterson, 2010), autism and mental retardation (e.g., Yirmiya, Erel, Shaked, & Solomonica-Levi, 1998), and children with typical development (e.g., Astington & Jenkins, 1999; Hughes & Dunn, 1998; Jenkins & Astington 1996; Lohmann & Tomasello, 2003; Ruffman et al., 2003).

The evidence for the connection between FB and language acquisition has been mounting. Thus, the meta-analysis of the relationship between language and FB in children under 7 years of age by Milligan and colleagues (2007) included 104 studies (N=8,891). They reported a moderate to large effect size for this relationship, which remained at this magnitude even when controlled for age. Furthermore, they found no significant differences between types of FB tasks. These findings point to a genuine developmental change in FB reasoning tied to language development. These findings present a challenge to both the constructivist and modularity approaches, which do not predict any deep association between FB reasoning and language. The short-term memory approach, although compatible with an observed link between language and FB (since the acquisition of both may rely on the availability of sufficient memory resources) does not make a prediction about a direct relationship between language and FB.

In light of such evidence, a new, "language-first" approach to the ontogenesis of FB was proposed, according to which the ability to reason about belief is expected to be isomorphic to the syntax of sentential complementation (e.g., Astington & Jenkins 1999; de Villiers, 1998, 2000, 2005; de Villiers & Pyers, 1997) and, presumably, semantics of propositional attitude reports related to it. Populations with disorders affecting language present a particularly valuable avenue for testing the hypothesized relationship between syntax of sentence embedding and ToM because of a grammatical impairment in such populations.

3. Previous Research on False Belief in Populations with Language Impairments

There have been relatively few studies of ToM development in language-impaired children. In a study by Gillott, Furniss, and Walter (2004), children with Specific Language Impairment (SLI) were contrasted with typically developing (TD) children and children with autism. They found that both clinical groups produced fewer responses referring to mental rather than physical states when asked to explain the actions of the characters in the stories used as a measure of ToM. Unlike children with autism, however, children with SLI did not differ from TD groups with respect to the appropriateness of their responses, suggesting that they possess a basic understanding of others' mental states, but are precluded by their language impairment from using the linguistic expressions necessary for describing them.

Although previous studies of children with SLI show a relationship between language and FB, they often point to a seemingly superficial connection between the two. Thus, Farrar and colleagues (2009) examined the relative contributions of various aspects of language to FB and found that while general grammatical development and vocabulary each contributed uniquely to FB reasoning, sentential complementation did not make a unique contribution. Thus, they concluded that the facilitation effect of language on FB is not linked to the syntax of embedded complementation, but to the level of general language development.

Miller (2001) tested children with SLI on the Unexpected Transfer task (Wimmer & Perner, 1983), but manipulated the linguistic complexity of the test question: it either included an embedded clausal complement ("Where does the puppet think the toy is?") or it did not ("Where will the puppet look for the toy?"). The results showed that the SLI group performed similarly to the younger language-matched TD controls when the syntactic complexity of the test question was high, but similarly to age-matched TD controls when the complexity was low. These results were interpreted as an indication that linguistic competence impacts FB performance, but that this connection was not at the deeper conceptual level, but rather at the level of performance. Children with SLI seemed capable of FB reasoning, but the more linguistically demanding task made it harder for this group to demonstrate their competence in this domain. In a later study, however, Miller (2004) added a production measure of sentence complements and reported a relationship between sentence complementation and FB performance even in the linguistically low-demand condition.

In sum, some studies of ToM in language-impaired samples seem to support a weak view of the "language first" hypothesis (see also Farrant, Fletcher, & Maybery, 2006, for evidence of the association between language impairment, visual perspective taking, and ToM). Because both lexical and grammatical levels of development were reported to play an important role in ToM facilitation, this has been interpreted as an indication that children's language development facilitates their success on FB because such tests rely on language. According to this view, in order to comprehend the narrative involved in the task, the child must reach a certain level of general language development, and deficient language leads to deficient performance on such tests.

4. Current Study

In the present study we address the question of the relationship between language and ToM by looking at FB reasoning in a small population characterized by an atypically high rate of a DLD. This population presents an interesting opportunity for testing differential predictions made by the alternative theories reviewed above, which can be summarized as follows.

If the “language-first” theory is correct, we should expect to find a relationship between children’s language development and ToM. Moreover, if the “strong” version of this theory is correct and the parts of grammar used for the embedding of one propositional argument under another is a prerequisite for representational ToM, then we should see a relationship between FB reasoning and children’s syntactic development, specifically their competence with respect to the syntax of sentence embedding (P_1). Of note is that the “strong” version of the “language-first” hypothesis does not predict an equally close association between FB and lexical development (P_2). Moreover, if ToM is a cognitive capacity somewhat independent from general cognitive ability, then we should expect to find a dissociation between ToM reasoning and cognitive ability, and children with DLD should display poor performance on the FB task regardless of their IQ and/or short-term memory capacity (P_3).

In contrast, if language plays no role in the development of FB, and if the child relies solely on his or her social experience and general reasoning abilities to develop an understanding of others’ mental states, as proposed by the constructivist accounts, typical cognitive development coupled with typical experience in observing human behavior is sufficient for the successful acquisition of FB reasoning. Then, as long as the child lives in an environment comparable to what is typical in the level of social and cognitive stimulation, ToM should unfold on a normal schedule even when the child exhibits language impairments. Furthermore, if normal cognitive development is required for the acquisition of FB reasoning, then we should expect to find a difference in IQ and/or short-term memory scores between the high and low FB performing children in the population in question, and, conversely, children with low IQ should underperform on the FB task compared to children with IQs in the normal range (P_4).

5. Method

5.1. Population

The present study is part of a large-scale genetic-epidemiological study qualifying and quantifying DLD (i.e., establishing phenotypic manifestations, prevalence, patterns of transmission, and etiology) in a small Russian-speaking population (henceforth, AZ) with an unusually high prevalence of DLD, i.e. atypical language development despite the absence of apparent neurobiological or sensory abnormality (as ascertained by the medical records and a screening for dysmorphology by a certified clinical geneticist). The term most frequently used to refer to this type of phenomenon is Specific Language Impairment (SLI) (Leonard, 1998).3

The population resides in a cluster of villages in Russia’s European North and is genetically homogeneous, yet characterized by a high degree of genetic variation across the genome and minimal levels of consanguinity. Currently, the total population consists of 861 individuals, of whom 138 are children between the ages of 3 and 18 (Rakhlin et al., submitted). Environmental characteristics and social-economic status (SES) of the population are highly uniform, making family, educational and other background attributes comparable for children growing up in these villages. Thus, all children (regardless of special needs or address) attend one preschool, elementary and secondary school and socialize closely outside of school. As another indicator of environmental uniformity, as part of a larger

[3]Although studies of SLI use an IQ cutoff of 85 to ensure that children included in the study have normal non-verbal cognitive functioning, we did not restrict the IQ range in the sample for two reasons: 1) since we are investigating the respective roles of IQ and language development, restricting IQ variation would be disadvantageous to the goals and validity of the study; 2) there are no reasons to believe that children with language impairment on both sides of the IQ cutoff form two qualitatively different groups with respect to their language development; moreover, low-average IQ seems to be a component of the SLI profile (see Plante, 1998, for a thorough discussion). In order to bypass this issue, we use the term DLD with an understanding that the language profiles of our participants with language impairments are akin to those of children in the SLI literature.

study, we have obtained information about the educational level of over a hundred sets of parents of the school-aged children using the Teacher's Report Form (Achenbach, 1991). According to these data, 8% of mothers and 4% of fathers hold college degrees, while 55% of mothers and 63% of fathers have completed vocational/technical training or had some college education, 21% of mothers and 14% of father have high school diplomas, 16% of mothers and 19% of fathers did not finish high school. Thus, the majority of the adult population has high school or vocational training, illustrating the uniformity of the SES background of the participants.

Our investigation has revealed that about 30% of school-aged children have phonological and grammatical deficits in elicited production, with the number reaching 40% in preschoolers and 30% in adults. Our research also showed that AZ children performed significantly worse than a sample from an SES- and culturally-similar population from the same region on measures of expressive phonology, syntax, and semantics/pragmatics.

The complex phenotype observed in the population exhibits vertical familial transmission, suggesting common genetic etiology for the various types of language impairments. The detailed description of the disorder phenotype is described elsewhere (Rakhlin et al., submitted). In brief, the disorder is characterized by heterogeneity of deficits, namely different areas of deficits and their combinations are found across individuals, including deficits in articulation and prosody, grammatical wellformedness, syntactic complexity, lexical development, and semantic/pragmatic characteristics. Of note is that although phonological short-term memory, as indexed by non-word repetition tasks (Gathercole & Baddeley, 1996), is implicated in the disorder in this population, the magnitude of the effect was found to be small in a study of 199 individuals (Rakhlin et al., submitted).

5.2. Participants

For the current study we recruited 54 elementary school children from this small population (46% girls; age range: 5;0 to 12;9; *M* = 8;5, *SD* = 2;3). The recruitment was carried out through the local kindergarten and elementary school. The sample represents all of the available primary school children at or above the age of FB acquisition generally reported in the literature (e.g., Wellman, Cross, & Watson, 2001). Thus, the sampling was complete, with 4 total refusals to participate (i.e., ~7% of the total number of children within this age range). Informed consents were collected from the children's parents and oral assents from the children at the time of each assessment. None of the children were diagnosed with mental retardation, autism, Down's Syndrome, or hearing impairment, as indicated by their medical records.

The participants were divided into 4 impairment status groups: children with typical development (TD, n=22), children with DLD who had a total score on a standardized assessment of language development (ORRIA) below 85 and whose non-verbal IQ score was 85 or above (DLD, n=21), children with DLD whose IQ was under 85 (DLD+LowIQ, n=4), and, finally, children whose IQ was under 85 and who did not have DLD (LowIQ, n=5).

The groups differed in age, $F(3,48) = 19.15$, $p < .001$, $\eta_p^2 = .55$. The post-hoc analysis revealed that the grous unaffected with DLD, namely the TD group (*M* = 10, *SD* = 1;11) and the LowIQ group (*M* = 10;8, *SD* = 1;5), were on average older than the DLD (*M* = 6*;11*, *SD* = 1;2) and DLD+LowIQ (*M* = 7;1, *SD* = 1;4) groups. The relationship between age and language disorder status is consistent with the general pattern of the prevalence of DLD in this population, which is highest in the youngest group. The correlation between age and standardized language development score is reported below.

5.3. Assessments

To ascertain the participants' level of language and cognitive development, all of the participants were given a standardized age-appropriate measure of non-verbal cognitive functioning, a standardized Russian language development assessment, a narrative task, and a series of short-term memory tasks.

5.3.1. Measures of cognitive functioning—Most children were given the extended version of the Universal Non-Verbal Intelligence Test (UNIT; Bracken & McCalum, 1998), a detailed, individually administered assessment. As part of a larger study, all of the children in the population, with a few exceptions due to their unavailability at the time of testing, were given the group-administered Culture-Fair Intelligence Test, Scale 2 (CFIT; Cattell & Cattell, 1963). We used UNIT scores for the analysis whenever available (n=43); otherwise we used CFIT scores (n=11). For both tests, standardized general IQ scores were available (M = 100, SD = 15).

The UNIT (Bracken & McCalum, 1998) is a non-verbal test battery for ages 5–18 that is designed to be a fair assessment of general cognitive functioning, especially in individuals with speech, language and hearing impairments or who are from different cultural and linguistic backgrounds, since the administration procedure is fully non-verbal. It requires multiple response modes, including the use of manipulatives, paper-and-pencil and pointing. The extended battery includes six subtests: Object Memory, Spatial Memory, Symbolic Memory, Cube Design, Analogic Reasoning, and Mazes, with the first three designed to assess memory, and the last three to assess reasoning. The test also provides scores for the development of symbolic versus non-symbolic components in general cognitive functioning. In our sample, the internal consistency coefficient (Cronbach's α) for the subtests ranged from .74 to .89 with M = .85. We used both standardized Full-Scale scores (FSIQ) and standardized Memory Quotient scores (MQ). For 9 children who were not administered the whole battery, MQ scores were obtained by administering the corresponding subtests.

The CFIT (Cattell & Cattell, 1963) is a group administered paper-and-pencil test for individuals ages 8 and above. It is a measure of non-verbal fluid intelligence, which is thought to be relatively independent of verbal fluency, cultural background and educational level. The battery consists of four subtests. We used the standardized general IQ score. The sample internal consistency coefficient α for the subtests ranged from .89 to .95 with M = .92.

5.3.2. Additional measures of short-term memory—In order to assess the children's span of short-term memory, they were administered a series of tests measuring their short-term memory for words (SM Words; unrelated words repetition; n=51), digit span (SM Numbers; forward and backward digit sequence repetition; n=50), as well as memory for spatial structure recognition (SM Spatial; n=47). The scores were adjusted for age.

5.3.3. The Assessment of the Development of Russian Language (ORRIA)—Since there are as of yet no published standardized assessments of language development of Russian, we had to rely on an unpublished assessment, namely, Assessment of Russian Language Development (ORRIA, Babyonyshev et al., 2007). This test, comparable to the Clinical Evaluation of Language Fundamentals (CELF; Semel, Wiig, & Secord, 1995), the Test of Language Development (TOLD; Hammill & Newcomer, 1982), and the Comprehensive Assessment of Spoken Language (CASL; Carrow-Woolfolk, 1999), is aimed at comprehensively assessing the range of acquisition of Russian in the areas of phonology, morphology, syntax, logical form, and the lexicon, tapping into both active and passive linguistic knowledge. The assessment is comprised of the following 7 subtests:

"passive vocabulary" (PV; lexical comprehension); "active vocabulary" (AV; lexical production); "linguistic operators" (LO; assessment of semantic knowledge of temporal and logical operators in comprehension); "sentence structure" (SS; assessment of syntactic knowledge in comprehension); "word structure" (WS; assessment of morphological competence in production); "phonological concepts" (PH; assessment of both active and passive knowledge of phonological processes); and "sentence repetition in context" (SR; assessment of both passive and active knowledge of syntactic structures and categories).

We administered the ORRIA to 52 of 54 children. We obtained standardized age-adjusted scores for overall language development (ORRIA; M = 100, SD = 15) and specific subtests (M = 10, SD = 2) using additional samples representative of the general population of Russian children4. The internal consistency of the subtests within our sample ranged from .83 to .97 with M = .87.

5.3.4. Narrative task—Story narratives were collected from the children participating in the study using wordless storybooks: "Frog, Where Are You?" and "One Frog Too Many" (Mayer, 1969). Each child was given a chance to look through the book and was asked to tell the story. We were able to obtain narratives from 52 children out of 54. The transcriptions of the recorded narratives were analyzed by two native Russian-speaking linguists for a number of linguistic traits, such as phonological characteristics, grammatical and lexical errors, mean length of utterance in words (MLU_w), complex syntax, lexical richness, and semantic/pragmatic characteristics. In each of these categories, both raw scores and age-adjusted performance scores based on the data from the comparison population (see Rakhlin et al., submitted) were derived for the purposes of the larger study.

5.3.4.1. Narrative Scoring: For the current study, we used two of the measures, namely Syntactic Complexity and Mental Lexicon, both adjusted for the length of the narrative measured as the total number of words. The measure of Syntactic Complexity was calculated in two ways, both of which were used in the final analysis. The first was a general measure of complex syntax, in which we calculated the frequency of complex constructions of various types (henceforth, Syntactic Complexity). These included verb-complement clauses (both finite and non-finite), relative clauses, coordinated clauses, non-formulaic *wh*-questions, participial constructions, and passives. We have included all of these structures in the measure of Syntactic Complexity to ensure adequate sampling of children's syntactic development and because these structures have been used previously to assess syntactic development in children with DLD (e.g., Reilly et al., 2004; also see Smith, Apperly, & White, 2003 for the relationship between FB and relative clause acquisition). In addition, we separately calculated each narrative's score for complement clauses only (henceforth, Embedding).

As a measure of the children's lexical development with respect to mental and psychological terms, we calculated the frequency of mental, psychological and perception verbs, as well as adjectival predicates (henceforth, Mental Lexicon) in children's narratives. These included Russian equivalents of the verbs *think, know, want, recognize, see, hear*, and adjectival predicates *happy, sad, mad, scared, surprised,* and so on. Importantly, in order to differentiate children's knowledge of lexical semantics from that of syntax, each instance of these lexical items was counted regardless of the correctness of the syntactic frame in which it was used or whether it was used with an embedded proposition or not. Thus, all of the

4The ORRIA, in preparation for publication by a Russian publisher, is currently being standardized, using representative samples of typically and atypically developing children. At this point, we have a substantial amount of normative data to generate standardized scores for this work.

occurrences of the relevant lexical item equivalent to the examples in (1)–(6) would be included in the total count (* indicates ungrammaticality):

1. The boy saw how the big frog pushed the little frog.
2. The boy saw that.
3. *The boy saw.
4. The boy was glad that he found the frog.
5. The boy was glad.
6. *The boy glad.

Psychological, mental, and perception verbs were all included in the total score in order to ensure an adequate sampling of the Mental Lexicon, as an index of the ability to refer to actors' mental states, rather than merely their actions. The analyses were carried out using the total number of both mental/psychological, and perception verbs, as well as using mental/psychological verbs only and excluding perception verbs.

Thus, for the purpose of comparing the FB pass and fail groups, four variables were derived: Syntactic Complexity, Embedding, and Mental Lexicon with and without perception verbs. The mean inter-rater reliability was $r = .96$, $p < .001$, indicating excellent convergence of ratings.

5.4. FB Procedure

The children were tested on a standard "Unexpected Transfer" FB task (Wimmer & Perner, 1983), which consisted of eight items. The scenarios were presented as mini-stories acted out with different toys and locations in each. In the training phase, the child was given a scenario with one character leaving an object in a location chosen from three possible ones and leaving the scene. The child was asked to indicate where the character would look for the object when he returns and wants to retrieve it. If the child indicated the correct location, the test phase began. The children were allowed up to two opportunities to give a correct response to pass training, but all of the children passed after one trial.

In every test phase scenario, there were two characters (pirates *Vasia* and *Petia*), who saw an object being placed in a certain location, after which one of them left the scene and the object was moved to a different location by the remaining character. When the first character came back and wanted to find the object, the child was presented with the test question, "Where will *Vasia* [character 1] be looking for the …?" In order to make sure that the child understood the scenario and that the object has been moved, a control question (e.g., "Where is the ball hidden now?") was asked during the presentation of each scenario (there were no incorrect responses). The child's task was to indicate the location without providing justification; the responses were counted as correct or incorrect regardless of whether they were verbal or pointing.

In the pattern of their responses, over 83% of the children fell into two groups: those who got all items correctly or failed on all items. For the remaining children, those who performed above chance (namely, gave 6 out 8 correct responses) were classified as passing.

6. Results

6.1. Descriptive Statistics and Correlations

Descriptive statistics for the measures used in the present study are displayed in Table 1. Overall, the assessments showed high internal consistency. The general language development scores on the ORRIA followed normal distribution ($Z = .71$, $p = .69$), $M = 87.95$, $SD = 19.42$. The distribution of nonverbal IQ scores was also normal ($Z = .92$, $p = .37$), $M = 94.11$, $SD = 13.75$.

Table 2 shows the correlations between study measures. In sum, language development scores for the ORRIA were highly interrelated and were also moderately related to measures of syntactic complexity, general cognitive functioning, and age. It is worth noting that although language development was related to age, demonstrating a developmental profile of the language disorder in the population; such a pattern was not obtained for the measures of general cognitive functioning. Short-term memory measures also correlated with each other.

6.2. False-Belief Performance and Age and Individual Differences in Language Functioning

First, we checked whether there is a relationship between age and success on the FB task. Of the 54 children who took the FB test, 32 (59.3%) failed and 22 (40.7%) passed. The pass and fail FB groups showed a significant difference in age, $t(52) = -7.19$, $p < .001$, $d = -1.99$, with the older children outperforming the younger ones. The mean age of the fail-group was 7;2 (*SD* = 1;5), while that of the pass-group was 10;5 (*SD* = 1;9), which greatly exceeds the age of acquisition reported previously (Wellman et al., 2001).

Next, since an analysis of the correlations between the study measures revealed strong significant positive associations between different aspects of language development as measured by the ORRIA (Table 2), we further investigated the relationship between language development and FB acquisition. We conducted a one between-subjects factor (FB group) MANCOVA analysis aimed at revealing individual differences in language development scores among the pass- and fail-FB groups when controlled for nonverbal IQ and short-term memory. Since the obtained pattern of results was nearly identical for all four short-term memory measures, only the results for the most comprehensive measure, UNIT MQ, are reported. The results of MANCOVA revealed significant main effects of FB group status on composite language development, Hotteling's $T = .99$, $F(7,34) = 4.78$, $p = .001$, $\eta_p^2 = .50$, demonstrating a statistically significant finding with a large effect size. As can be seen in Figures 1 and 2, children who were successful at FB task scored much higher than those who failed the task on both the total ORRIA score and on all of its subtests.

Next, we investigated whether the children's performance on the FB task was related to the measures of complex syntax and the mental lexicon as assessed by the narrative task. A set of performed ANCOVAs showed that when controlled for IQ, short-term memory, and general language development, the FB pass- and fail-groups differed significantly on Syntactic Complexity, $F(1,37) = 11.48$, $p = .002$, $\eta_p^2 = .24$, and Embedding, $F(1,37) = 4.40$, $p = .04$, $\eta_p^2 = .11$, but not Mental Lexicon, $F(1,37) = .71$, $p = .41$, $\eta_p^2 = .02$.[5] As Figure 3 illustrates, children who were successful on the FB task demonstrated higher scores on both measures of syntactic development, but not a lexical measure of mental verbs and adjectival predicates.

6.3. False-Belief Performance and General Cognitive Development

Our next step was to investigate the relationship between the children's general cognitive ability and FB success. Our results indicated that there was no significant difference in nonverbal IQ scores between the FB-pass ($M = 91.97$; $SD = 12.88$) and FB-fail groups ($M =$

[5]Two of the least syntactically-loaded subtests (passive vocabulary and active vocabulary) were used to control for general language development. All of the results hold when not controlled for the two vocabulary measures. Also, the results were identical for each of the two analyses including and excluding perception verbs from the Mental Lexicon measure.

97.23; *SD* = 14.66), $t(52) = -1.39$, $p = .17$, $d = -.38$ (see Figure 2). There were also no significant differences in short-term memory as measured by the UNIT MQ, $t(50) = -.48$, $p = .63$, $d = -.14$, SM Words task, $t(49) = -.36$, $p = .72$, $d = -.10$, the SM Numbers task, $t(48) = -.10$, $p = .92$, $d = -.03$, and the SM Spatial task, $t(45) = -.48$, $p = .64$, $d = -.14$. The results held when controlled for age. Thus, in our sample of children with delayed FB acquisition, general cognitive development was not significantly related to their performance on the FB task.

6.4. False-Belief Performance in Typically Developing, Language- and Cognitive-Impaired Groups

Our final analysis was aimed at corroborating the results described in the previous sections by examining contingencies between children's FB performance and their impairment status as determined by their scores on standardized measures of language and cognitive functioning. For that purpose, we compared the FB success and failure rates across four groups described in section 5.2., namely, children with DLD, DLD+LowIQ, LowIQ, and TD. Table 3 provides summary rates for passing and failing the FB task in the four groups. In sum, we found a very strong and significant relationship between group and performance on FB tasks, $\chi^2(3) = 13.92$, $p = .003$, Cramer's $V = .52$, with the highest rate of failure in the groups with language disorder (DLD and DLD+LowIQ) compared to the two groups without language disorder (TD and LowIQ)[6].

Of note is the finding of the 32% rate of failure on the FB task among the children with TD. To further investigate the reasons for such a high rate, we compared those TD children who passed and those who failed the FB task on their IQ, UNIT MQ, and ORRIA scores. The two TD subgroups did not differ on IQ [$F(1,20) = 1.82$, $p = .19$, $\eta_p^2 = .08$] or short-term memory [$F(1,20) = 2.22$, $p = .15$, $\eta_p^2 = .10$], but significantly differed on their ORRIA scores, demonstrating a large effect size [$F(1,20) = 33.08$, $p < .001$, $\eta_p^2 = .62$]. Although these children were classified as TD, a careful examination of their ORRIA scores revealed that most of them were near the cut-off used to determine the language disorder status. This result further supports the hypothesized association between FB reasoning and language development.

7. Discussion

7.1. Predictions

In the remainder of the paper we will discuss the obtained pattern of results from this study pertaining to the set of predictions introduced in section 4 and summarized below.

The first goal of the present study was to address the following predictions (P) made by the alternative approaches to the issue of the ontogenesis of FB reasoning as a critical milestone in ToM acquisition:

P_1 If the development of FB requires typical language development, then we should expect high and low FB performing children to differ in their level of language development. Moreover, if the parts of grammar used for embedding of one propositional argument under another are necessary for

[6]As pointed out by the anonymous reviewers, the association between failing FB and poor language functioning may be confounded by a lower SES in the children with lower language skills. Low SES may be detrimental for children's social and cognitive environment, thus explaining the lower performance of language-impaired children on FB without attributing to language a direct role in FB development. This explanation is unlikely, however, because in this study we failed to find an association between parental education (one of the SES indicators), and both language status and FB performance (*p*'s > .45). Although we only had information about parent education for the elementary school children from this sample, as we pointed out in the Introduction, the education level and other environmental characteristics for the population at large is highly uniform.

FB, then we should see a relationship between syntactic development, namely, syntax of sentence embedding, and FB reasoning.

P_2 If FB competence is independent from syntactic development, but is solely a result of the child's sufficient understanding of others' minds, then acquiring lexical meaning of mental state verbs may be sufficient for success on FB.

P_3 If FB is a cognitive capacity somewhat independent from general cognitive ability, then we should expect to find dissociation between FB reasoning and IQ or short-term memory, and children with a language disorder should display poor performance on the FB task regardless of their IQ.

P_4 If the development of FB representation is dependent on normal cognitive development, then we should expect to find a difference in IQ or short-term memory scores between the high and low FB performing children in the population in question, which is characterized by variable IQs.

Our results are consistent with P_1 and P_3, but not P_2 and P_4. Whereas those participants who passed the FB task performed significantly better than those who failed on the ORRIA (including all of its subtests), they exhibited no significant difference in their cognitive functioning as measured by standardized non-verbal IQ tests. These findings are in concert with our observation that age was another factor significantly related to success on FB. Since children's standardized performance scores on the ORRIA, but not their standardized IQs, are age-related, it is not surprising that FB performance would be subject to the age-effect. Moreover, as in the population of interest, the rate of language disorder declined across age groups, this is linked to the successful acquisition of FB.

7.2. The Role of Syntactic Complexity and Mental Lexicon in False Belief

A relationship between language and FB has been demonstrated in a large body of experimental literature and at least in some sense is accepted as well established. An important issue that is still unresolved is whether the observed relationship goes beyond trivial. As was discussed above, among those who accept a role of language in FB development, there are theorists who maintain that the role of language in FB is merely superficial and is engendered by the verbal nature of FB-related tasks, successful performance on which requires a certain level of general language development (Chandler, Fritz, & Hala, 1989). On the other end of the theoretical spectrum are those who posit a more fundamental role of language in the ontogenesis of ToM.

Among the latter school of thought, there are some researchers who believe that the crucial aspect of language involved in bringing about a mature understanding of other minds is the lexical semantics of propositional attitude verbs. According to this view (e.g., Bartsch & Wellman, 1995), acquiring the lexical meanings of verbs of perception, emotion and desire as well as verbs used for belief and knowledge reports, such as *see*, *want*, *know*, *think*, *remember*, etc., which encode the concepts involved in reasoning about others' minds, allows the child to crystallize these concepts first in relation to their own mental states and then to those of others' (Astington, 1996). Another approach attributes a crucial facilitative role not to the terms denoting mental states, but to the syntactic structures used for the embedding of one proposition under another, which serve as the syntactic frame for the semantics of propositional attitude reports, such as belief, knowledge and desire attribution (Astington & Jenkins, 1999; de Villiers, 1998, 2000, 2005; de Villiers & Pyers, 1997).

Evidence in favor of the sentential complementation view comes from the following observations. First, it has been reported that children with TD start producing mental verbs in their spontaneous speech as early as 2 years of age (Bartsch & Wellman 1995; Bloom et al., 1989), well before they become able to pass a standard FB task. However, it is well-known that the acquisition of the syntax of complement clauses lags behind the lexical acquisition of mental state verbs and precedes children's development in the ability to reason about FB. This suggests that the former is not sufficient for FB acquisition, and that the latter is necessary for it. Thus, a longitudinal study with 3- and 4-year-olds (de Villiers & Pyers, 2002) found that success with the standard FB tasks was contingent on the acquisition of the structures of communication verbs with complements. In the regression analyses they conducted, sentential complement comprehension was the best predictor of the FB score. The facilitative role of sentential complements for FB understanding was confirmed by the training study by Lohman and Tomasello (2003). They showed that 3-year-old children who were trained on sentential complement syntax improved their FB understanding similarly to the children trained on perspective-shifting discourse about deceptive objects. In contrast, children who were exposed to deceptive objects without accompanying language did not improve.

In our study, we addressed the question of whether the children's mastery of the syntax of sentence embedding and/or the lexicon of mental and psychological verbs is related to their success on FB reasoning. In order to answer this question, we examined the children's spoken narratives coded for the use of complex syntax and the mental lexicon. On the measure of complex syntax we conducted two separate analyses measuring all complex syntactic structures as a proxy for the child's development of complex syntax, and, separately, only the instances of embedded complement clauses. Supporting P_1, we found a relationship between the frequency of complex structures on both measures and FB group membership. Even when controlled for IQ, short-term memory and measures of general language development, syntactic complexity scores were significantly higher in the FB-pass group (see also Miller, 2004). In contrast, the two groups did not differ on the measure of the mental lexicon, contrary to P_2.

The found dissociation between syntactic complexity and the lexicon of psychological and mental terms in the children's performance on FB tasks has interesting implications not only for the study of ToM development, but also for the field of DLD. A question we can ask is whether the lack of contrast between FB-pass and -fail groups on the measure Mental Lexicon is reflective of a lack of such contrast between typically developing and language-impaired children in the population of interest. The answer to this question is yes: even though the language-impaired groups (DLD and DLD+LowIQ) performed significantly worse on the syntactic complexity measure compared to the two language-unimpaired groups (TD and LowIQ; $p = .01$), there were no significant differences in their performance on the Mental Lexicon measure ($p = .43$). This suggests that the lexical acquisition of mental state verbs and adjectival predicates, and hence the ability to refer to the mental states of others using these lexical items, is not compromised in children with DLD in the population under investigation, but is not sufficient for the development of FB competence.

It is important to point out that our measure of Mental Lexicon did not assess the richness of the children's lexicon, the degree to which they acquired full lexical meaning of these terms or their rate of lexical errors. The two lexical subtests of the ORRIA (active and passive vocabulary), which were designed to assess children's general lexical development, showed that the language-impaired groups had significantly lower scores than their unimpaired counterparts. Our measure was meant to probe children's awareness that others possess mental states and that their actions are motivated by their mental states, indicated by the frequency with which the children referred to the protagonists' mental states rather than merely referring to their actions while relating the events taking place in the story.

An additional question we can ask with respect to the dissociation between the measures of mental lexicon and complex syntax in children's performance on FB that we found is what it tells us about the respective roles of lexical knowledge of mental and psychological predicates and the syntax of sentence embedding in the standard FB task. Our results indicate that even though a child may possess the knowledge of concepts associated with reasoning about the mental states of others, as substantiated by their use of mental predicates, such lexical knowledge is not sufficient, and until he or she reaches the point of language development where the syntax of sentence embedding has been fully acquired, the child does not perform well on FB tasks.

It is important to note that although the syntax of sentential complementation is thought by some researchers to play a crucial facilitative role in FB reasoning development (e.g. de Villiers & de Villiers, 2003), some studies failed to detect the unique contribution of sentence complementation to FB success. Moreover, other aspects of language, such as general grammar and lexicon, were shown to correlate with FB performance equally strongly (Farrar et al., 2009; Slade & Ruffman, 2005). Previously, this was taken as evidence against a special role for sentence embedding in ToM development. It is, however, possible to interpret such findings differently. Namely, one can argue that the reason it is difficult to isolate a contribution of sentence embedding to FB success is because, 1) the assessments used to measure various aspects of language development contain multiple redundancies and fail to isolate single parameters of children's performance and, 2) various aspects of language development are highly correlated with each other. As a result, a child at the point of development when he or she has acquired sentence embedding, would be more advanced on most other linguistic measures compared to a child who has not.

Thus, in our study the general measure of complex syntax, which included passives, *wh*-questions and participial constructions in addition to embedded sentences was just as strongly related to FB performance as the more relevant measure of sentence embedding. We also found that all of the subtests of the ORRIA, including phonology, passive vocabulary and active vocabulary, were strongly correlated with FB performance, as well as with each other.

7.3 The Role of Short-Term Memory in False-Belief

Another hypothesis that we addressed in the study was whether short-term (or working) memory plays a role in children's success on the FB task. According to one approach, the core deficit responsible for children's failure to predict how an actor who holds a mistaken belief would behave as a consequence of such a belief is a limitation in their information processing abilities, namely working memory (Case, 1989; Fodor, 1992; Frye, Zelazo, & Palfai, 1995; Keenan et al., 1998; Olson, 1989, 1993). According to this proposal, working memory limitations constrain the expression of children's understanding of other minds and force them to resort in such tasks to heuristics, which often leads to errors.

Due to space limitations, we cannot include a detailed discussion of this literature, which contains various models of working memory and disagrees on the precise mechanism responsible for the apparent increase with age in children's capacity to "hold in mind" information. In brief, under one view, it is the capacity of working memory itself that increases with maturation (e.g., Pascual-Leone, 1970). Under another view, what increases with development is the sophistication of the available control processes and strategies leading to "operational efficiency," consequently requiring less working memory and leaving more capacity to be allocated to other tasks (e.g., Case, Kurland, & Goldberg, 1982).[7]

Regardless of the specific model of short-term or working memory one adopts, the common view that they all share is that the amount of information that can be actively held in mind increases with age and is strongly related to cognitive development (Case, Kurland, & Goldberg, 1982). The growth of working memory has been linked to children's developing understanding of FB (Keenan et al., 1998; Olson, 1989, 1993). According to this hypothesis, the increased capacity of a child's working memory allows him or her to form more complex representations of the social world and not only to represent a person's belief, but to represent it in conjunction with a proposition denoting the state of affairs in the real world, and consequently being able to represent the former as a false description of some fact.

In previous research, there has been some evidence in support of the relationship between FB understanding and children's working memory (Davis & Pratt, 1995; Keenan et al., 1998). Keenan and colleagues have shown that children's performance on the counting span measure predicted their performance on the FB measure, even when controlled for age, accounting for about 7% of the variance in the FB measure. A limitation of Keenan et al.'s study, however, was that it did not control for children's language proficiency, which may be responsible for the relationship between working memory and FB. In a subsequent study by Keenan (1998), language ability was taken into account and a hierarchical multiple regression analysis revealed a significant amount of unique variance in FB scores attributable to the working memory scores in a sample of TD 4- and 5-year-old children.

Results of our study are not consistent with Keenan and colleagues' (1998). In our sample of older participants with a much wider age range, we found that even when controlling for age, there was no significant difference between FB-pass and FB-fail groups on any of the four short-term memory measures we administered, including the general memory scale from the UNIT, and the digit span, a measure of working memory required to manipulate temporarily stored information. Thus, we found that children in the fail-group did not differ significantly on their memory span and that, despite the contrast in FB performance, the memory scores for both groups were above the cutoff for low functioning (see Table 1). The only parameters on which the two groups differed significantly were language ability and age, the latter being a factor related to both FB performance and language ability in our sample.[8]

Based on these results, we can conclude that while children's general cognitive ability and short-term memory (including the working component) are not related to their success on FB[9]; their language ability is.

7.4. Implications for the Understanding of the Structure and Cognitive Bases of Theory of Mind

Our finding supports the idea that ToM is a complex multifaceted area of cognition with certain core components independent from language ability, with respect to which children with impaired language reach typical functioning despite their disorder. On the other hand, this domain contains certain functional components that are shared with other systems, such as language, in the development of which DLD has a delaying effect.

---

[7]A related group of hypotheses attributes the change in success on FB tasks in children between 3 and 5 to the maturation of executive functioning (e.g., Carlson, Moses, & Breton, 2002; Frye, Zelazo, & Palfai, 1995; Hala, Hug, & Henderson, 2003; Russell, 1996).

[8]The age difference between normal language and language-impaired groups is found in standardized scores, initially adjusted for age based on the normative sample data. This difference indicates higher rates of DLD among the 3–6-year-old children and a tapering off of the disorder in the older age group (either due to development or compensatory strategies).

[9]The present study did not investigate one of the posited components of executive functioning, namely, inhibitory control (IC), which has been shown to be implicated in FB acquisition and performance (e.g., Carlson & Moses, 2001; Carlson, Moses, & Breton, 2002). The contribution of this component to FB development in populations with language disorders needs to addressed in further research.

In contrast to the present study, studies with other clinical populations indicate that poor ToM reasoning can be related to abnormalities in the core system of ToM impairing children's basic understanding of other minds, as has been suggested for children with high functioning autism (Baron-Cohen, Leslie, & Frith, 1985; for literature review see Happe, 1995 and Yirmiya et al., 1998). Thus, Baron-Cohen, Leslie and Frith (1986) reported that children with autism provided fewer mental state terms in their narratives compared to typically developing controls. Tager-Flusberg and Sullivan (1995) showed that ToM performance correlated significantly with the use of emotion and cognition terms in adolescents with autism. In addition, in their study the participants with autism gave significantly fewer appropriate explanations for the emotional states of the story characters when prompted with questions (Tager-Flusberg & Sullivan, 1995; see also Gillott, Furniss, and Walter, 2004 for a similar finding). These findings point to a developmental profile somewhat distinct from the one we found in the population with DLD in our study.

In summary, the results of our study provide evidence for the "strong" version of the "language-first" theory of FB acquisition, supporting similar findings with other special populations and children with TD. We found that that the capacity to explicitly reason about FB and predict behavior based on such reasoning is impaired in the children with DLD and, furthermore, is significantly related to their level of performance on syntactic complexity, but not that of mental lexicon.

The unique contribution of this study stems from the fact that the participants are part of one large extended pedigree and therefore share both the environmental and the genetic factors substantiating the high prevalence of DLD observed in the given population. This presents us with a unique opportunity to examine the variable phenotype in all its complexity and look at a sample with a wide range of linguistic and cognitive abilities, yet with shared etiology. In a typical study of FB reasoning in a clinical population, participants are selected based on strict inclusion criteria, such as performance below a certain cut-off on a standardized linguistic scale or a certain (low) level of expressive language, as well as certain level of cognitive functioning. Being able to look at a full range of linguistic and cognitive abilities affords us a new perspective on the relationship between these abilities and FB reasoning.

Another unique feature of this study is that the population of interest is the subject of a large-scale ongoing study, which provides an opportunity for a detailed investigation of the behavioral and linguistic characteristics of the phenotype, providing a more complete picture of the developmental trajectory of the ToM and language relationship in a special population. Further issues that need to be investigated include looking at the precursors of ToM in infants at risk for DLD, comparing the performance of children with DLD on verbal and non-verbal measures of FB, and investigating the reverse relationship between delayed ToM and certain aspects of language development, such as looking at the role of ToM delay in the pragmatic deficits found in the population of interest.

The study has a number of limitations. One stems from the fact that there are no theoretically-based accounts of DLD in Russian and very few empirical investigations on which we could draw. Second, our results showed a significant difference in age between children who passed and failed the FB task. There was also a strong positive relationship between age and the measure of syntactic complexity. Although this is a natural finding and children are expected to increase in their syntactic complexity with age, further research is required to tease apart the effects of age from the effects of syntactic complexity on FB. Third, since there are no standardized normed assessments of both language and cognitive development in Russian, we had to rely on 1) a yet unpublished assessment of language development (ORRIA), and on 2) general ability measures standardized on American samples as proxies for cognitive development. It is worth noting that the present study recruited children from a population characterized by a higher than typical prevalence of DLD and, therefore, the findings from this study must be extended to other populations. However, these findings are in line with a growing body of evidence on the relationship between ToM and language development from other typical (e.g., Milligan et al., 2007; de Villiers & Pyers, 2002) and atypical populations, including children with language impairments (e.g., Miller, 2001, 2004; Farrar et al., 2006). Furthermore, it is the specificity of the disorder in the population, namely, deficient language functioning without apparent sensory, neurological, and general cognitive explanatory factors, which make this population of special interest for the question at hand. Thus, a whole-population study (Rakhlin et al., submitted) with an external demographically and socio-economically similar control sample has revealed that despite the large differences in various indicators of language functioning between this population and the external controls, the difference in nonverbal cognitive functioning was small (i.e., within half a standard deviation, mirroring the difference between the TD and DLD groups in the present study). Such a difference is a typical result of studies of general nonverbal cognitive functioning of children with language impairments, and is even acknowledged by the developers of nonverbal intelligence measures (i.e., for UNIT, see Bracken & McCallum, 1998). Therefore, we believe that our findings represent a magnified rather than distorted view of the relationship between language functioning and the acquisition of FB understanding.

8. Conclusions

Our results indicate that the DLD observed at a high rate in the population in question appears to be associated with a significant delay in FB reasoning acquisition. In contrast to the previously published observations of typically developing children, who undergo a developmental change in the FB reasoning sometime between the ages of 3 and 5, in this population the comparable change occurs sometime between the ages of 8 and 10.

We also found that in addition to age, the only other factor that distinguishes the fail- versus pass-FB groups was their language ability as measured by the standardized scores on the Russian language assessment (ORRIA). Neither the children's general cognitive functioning nor their short-term memory (including the working component) was related to FB reasoning.

Finally, we demonstrated that complex syntax in general, as well as sentence embedding specifically, but not the lexicon of mental and psychological terms (as measured on the narrative task), were related to children's FB success. These results indicate that language plays a facilitative role in FB development and support the "strong" view, which maintains that the syntax of sentential complementation is a necessary prerequisite for the development of FB reasoning, widely considered to be a crucial stage of developing explicit representational ToM.

To conclude, we must offer a disclaimer. The theory that our findings appear to support suggests a causal relationship between the development of complex syntax and the acquisition of FB reasoning. We, however, do not wish to overstate the nature of the relationship between the two observed in our study, but would like to merely suggest that this relationship exists, at least at the correlational level. On the other hand, we have not found the same for general cognitive development, nor for children's abilities to invoke the mental states of actors in their narratives. Therefore, we feel that our findings might provide a foundation for further research in the field, but this research still needs to be done to confirm our observations.

Acknowledgments

The first two authors have contributed equally to this article. This research was supported by DC007665. Grantees undertaking such projects are encouraged to express freely their professional judgment. This article, therefore, does not necessarily reflect the position or policies of the National Institutes of Health, and no official endorsement should be inferred.

References

Achenbach, TM. Manual for the Teacher's Report Form and 1991 profile. Burlington, VT: Department of Psychiatry, University of Vermont; 1991.

Apperly IA, Butterfill SA. Do humans have two systems to track beliefs and belief-like states? Psychological Review. 2009; 116:953–970. [PubMed: 19839692]

Astington, JW. What is theoretical about the child's theory of mind? A Vygotskian view of its development. In: Carruthers, P.; Smith, PK., editors. Theories of theories of mind. Cambridge, MA: Cambridge University Press; 1996. p. 184-199.

Astington J, Jenkins J. A longitudinal study of the relation between language and theory of mind. Developmental Psychology. 1999; 35:1311–1320. [PubMed: 10493656]

Babyonyshev, M.; Hart, L.; Reich, J.; Kuznetsova, J.; Rissman, R.; Grigorenko, EL. Unpublished assessment. 2007. *Оценка Развития Русского Языка* [Assessment of the Development of Russian].

Baron-Cohen, S. Mindblindness: An Essay on Autism and Theory of Mind. Cambridge, MA: MIT Press/Bradford Books; 1995.

Baron-Cohen S, Leslie A, Frith U. Does the autistic child have a "theory of mind"? Cognition. 1985; 21:37–46. [PubMed: 2934210]

Bartsch, K.; Wellman, HM. Children Talk about the Mind. NY: Oxford University Press; 1995.

Bloom P, German T. Two reasons to abandon the false belief task as a test of theory of mind. Cognition. 2000; 77:25–31.

Bloom L, Rispoli M, Gartner B, Hafitz J. Acquisition of complementation. Journal of Child Language. 1989; 16:101–120. [PubMed: 2925807]

Bracken, BA.; McCallum, RS. Universal Nonverbal Intelligence Test (UNIT). Chicago: Riverside Publishing; 1998.

Carlson SM, Moses LJ. Individual Differences in Inhibitory Control and Children's Theory of Mind. Child Development. 2001; 72(4):1032–1053. [PubMed: 11480933]

Carlson SM, Moses LJ, Breton C. How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. Infant and Child Development. 2002; 11:73–92.

Carrington S, Bailey A. Are there Theory of Mind regions in the brain? A Review of the Neuroimaging Literature. Human Brain Mapping. 2009; 30:2313–2335. [PubMed: 19034900]

Carrol-Woolfolk, E. Comprehensive Assessment of Spoken Language (CASL). Circle Piness, Minn: American Guidance Service; 1999.

Case, R. A neo-Piagetian analysis of the child's understanding of other people, and the internal conditions which motivate their behaviour. Paper presented at the Biennial Meeting of the Society for Research in Child Development; Kansas City, MO. 1989.

Case R, Kurland M, Goldberg J. Operational efficiency and the growth of short term memory span. Journal of Experimental Child Psychology. 1982; 33:386–404.

Cattell, RB.; Cattell, AKS. A culture-fair intelligence test. Champaign, IL: Insitute for Personality and Ability Testing; 1963.

Chandler M, Fritz AS, Hala SM. Small-scale deceit: Deception as a marker of two, three-, and four-year-olds' early theories of mind. Child Development. 1989; 60:1263–1277. [PubMed: 2612240]

Davis HL, Pratt C. The development of children's theory of mind: The working memory explanation. Australian Journal of Psychology. 1995; 47:25–31.

Dennett D. Beliefs about Beliefs. Behavioral & Brain Sciences. 1978; 1:568–570.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

de Villiers, JG. On acquiring the structural representations for false complements. In: Hollebrandse, B., editor. New perspectives on language acquisition. Amherst: University of Massachusetts, Graduate Linguistic Student Association; 1998. p. 125-136.

de Villiers, JG. Language and Theory of mind: what are the developmental relationships?. In: Baron-Cohen, S.; Tager-Flusberg, H.; Cohen, D., editors. Understanding other minds: perspectives from autism and developmental cognitive neuroscience. NY: Oxford University Press; 2000. p. 83-123.

de Villiers, JG. Can language acquisition give children a point of view?. In: Astington, J.; Baird, J., editors. Why Language Matters for Theory of Mind. NY: Oxford University Press; 2005. p. 186-219.

de Villiers JG. The interface of language and theory of mind. Lingua. 2007; 117:1858–1878. [PubMed: 17973025]

de Villiers, JG.; de Villiers, PA. Language for Thought: coming to understand False Beliefs. In: Gentner, D.; Goldin-Meadow, S., editors. Language in Mind: Advances in the Study of Language and Cognition. Cambridge, MA: MIT press; 2003. p. 335-384.

de Villiers, J.; Pyers, J. Complementing Cognition: the relationship between language and theory of mind. In: Hughes, E.; Hughes, M.; Greenhill, editors. Proceedings of the twenty-first annual Boston University conference on language development. Vol. 1. Sommerville, MA: Cascadilla Press; 1997. p. 136-147.

de Villiers JG, Pyers JE. Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. Cognitive Development. 2002; 17:1037–1060.

de Villiers, PA. The role of language in Theory of Mind development: what deaf children tell us. In: Astington, J.; Baird, J., editors. Why Language Matters for Theory of Mind. NY: Oxford University Press; 2005. p. 266-297.

Farrant BM, Fletcher J, Maybery MT. Specific Language Impairment, Theory of Mind, and Visual Perspective Taking: Evidence for Simulation Theory and the Developmental Role of Language. Child Development. 2006; 77(6):1842–1853. [PubMed: 17107464]

Farrar JM, Johnson B, Tompkins V, Easters M, Zilisi-Medus A, Beningo J. Language and theory of mind in preschool children with specific language impairment. Journal of Communication Disorders. 2009; 42:428–441. [PubMed: 19647837]

Flavell JH, Green FL, Flavell ER. Development of knowledge about the appearance-reality distinction. Monographs of the Society for Research in Child Development. 1986; 51:1–68.

Fodor JA. A theory of the child’s theory of mind. Cognition. 1992; 44:283–296. [PubMed: 1424495]

Frye D, Zelazo P, Palfai T. Theory of mind and rule-based reasoning. Cognitive Development. 1995; 10:483–527.

Gathercole, SE.; Baddeley, AD. The Children’s Test of Nonword Repetition. London: Psychological Corporation; 1996.

Gillot A, Furniss F, Walter A. Theory of mind ability in children with specific language impairment. Child Language Teaching and Therapy. 2004; 20(1):1–11.

Gopnik A, Astington JW. Children’s understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. Child Development. 1988; 59:26–37. [PubMed: 3342716]

Gopnik, A.; Wellman, H. The theory theory. In: Hirschfeld, LA.; Gelman, SA., editors. Mapping the mind: domain specificity in cognition and culture. Cambridge, MA: Cambridge University Press; 1994. p. 257-293.

Hala S, Hug S, Henderson H. Executive functioning and false-belief understanding in preschool children: Two tasks are harder than one. Journal of Cognition and Development. 2003; 4:275–298.

Hammill, DD.; Newcomer, PL. Test of Language Development: Intermediate. Austin, TX: PRO-ED; 1982.

Happé F. The role of age and verbal ability in the theory of mind task performance of subjects with autism. Child Development. 1995; 66:843–855. [PubMed: 7789204]

Harris, PL. Desires, beliefs and language. In: Carruthers, P.; Smith, PK., editors. Theories of theories of mind. Cambridge, MA: Cambridge University Press; 1996. p. 200-220.

Harris PL. Simulation (mostly) rules: A commentary. British Journal of Developmental Psychology. 2009; 27:555–559. [PubMed: 19994568]

Hintikka, J. Models for Modalities. Dordrecht, Netherlands: D. Reidel; 1969.

Hogrefe GJ, Wimmer H, Perner J. Ignorance versus false belief: a developmental lag in attribution of epistemic states. Child Development. 1986; 57:567–582.

Hughes C, Dunn J. Understanding mind and emotion: longitudinal associations with mental-state talk among young friends. Developmental Psychology. 1998; 34:1026–1037. [PubMed: 9779748]

Jenkins JM, Astington JW. Cognitive factors and family structure associated with theory of mind development in young children. Developmental Psychology. 1996; 32:70–78.

Karmiloff-Smith, A. Beyond Modularity: A Developmental Perspective on Cognitive Science. Cambridge, MA: MIT Press; 1992.

Keenan T. Memory span as a predictor of false belief understanding. New Zealand Journal of Psychology. 1998; 27:36–43.

Keenan T, Olson DR, Marini Z. Working memory and children's developing understanding of the mind. Australian Journal of Psychology. 1998; 50:76–82.

Leslie AM. Pretense and representation: The origins of "theory of mind". Psychological Review. 1987; 94:412–426.

Lohmann H, Tomasello M. The role of language in the development of false-belief understanding: A training study. Child Development. 2003; 74:1130–1144. [PubMed: 12938709]

Mayer, M. Frog, Where are You?. NY: Dial; 1969.

Miller CA. False belief understanding in children with specific language impairment. Journal of Communication Development. 2001; 34:73–86.

Miller CA. False belief and sentence complement performance in children with specific language impairment. International Journal of Language and Communication Disorders. 2004; 39(2):191–213. [PubMed: 15204451]

Milligan K, Astington JW, Dack LA. Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. Child Development. 2007; 78:622–646. [PubMed: 17381794]

Mitchell, P. Acquiring a Conception of Mind: A Review of Psychological research and Theory. Hove: Psychology Press; 1996.

Olson DR. Making up your mind. Canadian Psychology. 1989; 30:617–627.

Olson DR. The development of representations: The origins of mental life. Canadian Psychology. 1993; 34:1–14.

Onishi K, Baillargeon R. Do 15-month-old infants understand false beliefs? Science. 2005; 308:255–258. [PubMed: 15821091]

Pascual-Leone J. A mathematical model for the transition rule in Piaget's developmental stages. Acta Psychologia. 1970; 32:301–345.

Paynter J, Peterson C. Language and ToM development in autism versus Asperger syndrome: Contrasting influences of syntactic versus lexical/semantic maturity. Research in Autism Spectrum Disorders. 2010; 4(3):377–385.

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

Peterson CC, Siegal M. Insights into theory of mind from deafness and autism. Mind & Language. 2000; 15:123–145.

Pinker S, Bloom P. Natural language and natural selection. Behavioral and Brain Sciences. 1990; 13:707–784.

Plante E. Criteria tor SLI: The Stark and Tallal legacy and beyond. Journal of Speech, Language, and Hearing Research. 1998; 41:451–957.

Premack D, Woodruff G. Does the chimpanzee have a theory of mind? Behavioral and Brain Science. 1978; 1:515–526.

Pyers JE, Senghas A. Language promotes false-belief understanding: evidence from learners of a new sign language. Psychological Science. 2009; 20:805–812. [PubMed: 19515119]

Rakhlin, N.; Felsenfeld, S.; Babyonyshev, M.; Kornilov, S.; Reich, J.; Grigorenko, E. Spoken language impairments in a geographically isolated population: illustrative pedigrees. Poster presented at the 2009 Annual Meeting of the Society for the Scientific Study of Reading; Boston, MA. 2009.

Rakhlin, N.; Kornilov, SA.; Palejev, D.; Koposov, R.; Babyonyshev, M.; Chang, J.; Grigorenko, EL. The Language Phenotype of a Small Geographically Isolated Russian-speaking Population: Implications for Genetic and Clinical Studies of Developmental Language Disorder. 2010. Manuscript submitted for publication

Reilly J, Losh M, Bellugi U, Wulfeck B. "Frog, where are you?" Narratives in children with specific language impairment, early focal injury, and Williams syndrome. Brain and Language. 2004; 88:229–247. [PubMed: 14965544]

Russell, J. Agency: Its Role in Mental Development. Hove: Erlbaum; 1996.

Ruffman T, Slade L, Rowlandson K, Rumsey C, Garnham A. How language relates to belief, desire and emotion understanding. Cognitive Development. 2003; 18:139–158.

Schick B, de Villiers J, de Villiers P, Hoffmeister R. Language and theory of mind: A study of Deaf children. Child Development. 2007; 78:376–396. [PubMed: 17381779]

Segal M, Varley R. Neural Systems Involved in 'Theory of Mind'. Nature Neuroscience. 2002; 3:463–471.

Semel, E.; Wiig, EH.; Secord, WA. Clinical Evaluation of Language Fundamentals. 3. San Antonio, TX: The Psychological Corporation, Harcourt Brace and Co; 1995.

Slade L, Ruffman T. How language does (and does not) relate to theory of mind: A longitudinal study of syntax, semantics, working memory and false belief. British Journal of Developmental Psychology. 2005; 23:117–141.

Smith M, Apperly I, White V. False Belief Reasoning and the Acquisition of Relative Clause Sentences. Child Development. 2003; 74(6):1709–1719. [PubMed: 14669891]

Spelke, ES. What makes humans smart? Core knowledge and natural language. In: Gentner, D.; Goldin-Meadow, S., editors. Language in Mind. Cambridge, MA: MIT Press; 2003. p. 277-311.

Surian L, Caldi S, Sperber D. Attribution of beliefs by 13-month- old infants. Psychological Science. 2007; 18:580–586. [PubMed: 17614865]

Tager-Flusberg, H. Language acquisition and theory of mind: contributions from the study of autism. In: Adamson, LB.; Romski, MA., editors. Research on communication and language disorders: Contributions to theories of language development. Baltimore, MD: Paul Brookes Publishing Company; 1997. p. 135-160.

Tager-Flusberg, H.; Joseph, R. Understanding in children with autism. In: Astington, J.; Baird, J., editors. Why language matters for theory of mind. NY: Oxford University Press; 2005. p. 289-313.

Tager-Flusberg H, Sullivan K. Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. Applied Psycholinguistics. 1995; 16:241–256.

Wellman, HM. Understanding the psychological world: Developing a theory of mind. In: Goswami, U., editor. Handbook of Childhood Cognitive Development. Oxford: Blackwell; 2002. p. 167-187.

Wellman HM, Cross D, Watson J. Meta-analysis of theory-of-mind development: The truth about false belief. Child Development. 2001; 72:655–684. [PubMed: 11405571]

Wimmer H, Perner J. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition. 1983; 13:103–128. [PubMed: 6681741]

Woolfe T, Want SC, Siegal M. Signposts to development: Theory of mind in deaf children. Child Development. 2002; 73:768–778. [PubMed: 12038550]

Yamaguchi M, Kuhlmeier VA, Wynn K, vanMarle K. Continuity in social cognition from infancy to childhood. Developmental Science. 2009; 12:746–752. [PubMed: 19702767]

Yirmiya N, Erel O, Shaked M, Solomonica-Levi D. Meta-analyses comparing theories of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. Psychological Bulletin. 1998; 124:283–307. [PubMed: 9849110]

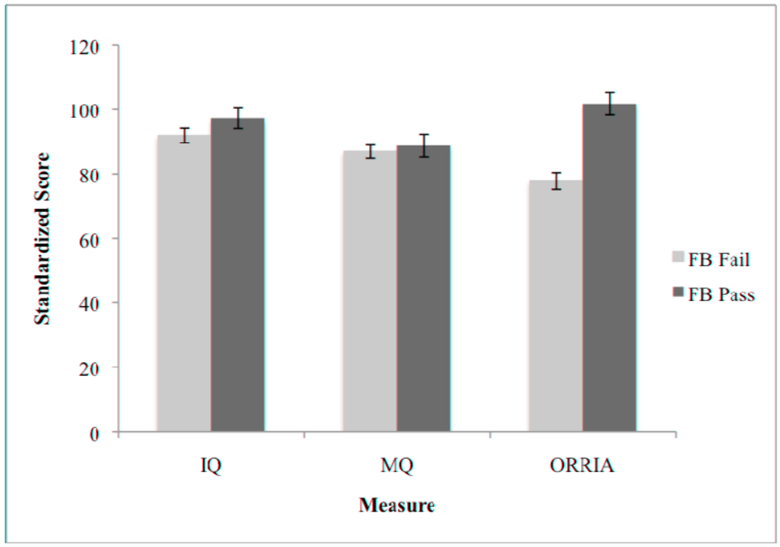

Figure 1.
Mean values for general cognitive development (nonverbal IQ as assessed by UNIT or CFIT; MQ – short-term memory as measured by UNIT) and language functioning (as assessed by ORRIA) measures among FB fail and pass groups (with standard error bars).

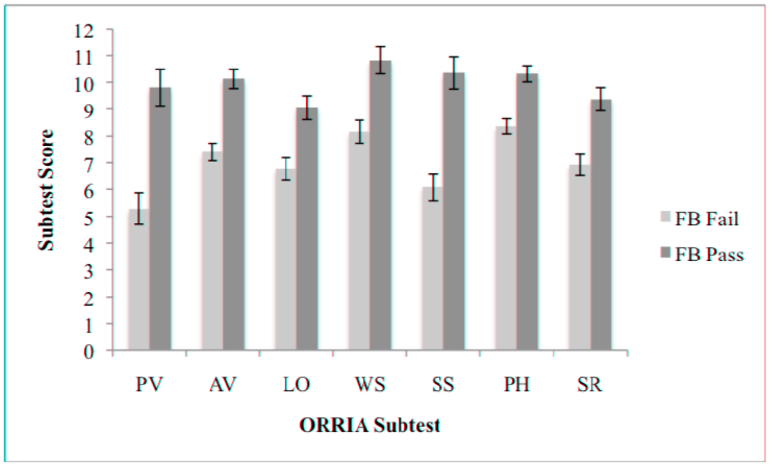

Figure 2.
Marginal mean values for ORRIA subtests among FB fail and pass groups (with standard error bars) when controlled for nonverbal IQ and short-term memory (UNIT MQ). PV – passive vocabulary, AV – active vocabulary, LO – linguistic operators, WS – word structure, SS – sentence structure, PH – phonology, SR – sentence repetition.

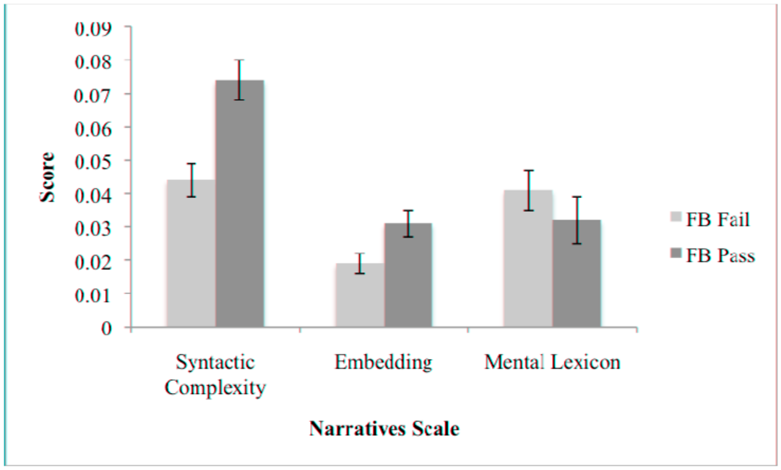

Figure 3.
Marginal mean values for narrative measures among FB fail and pass groups (with standard error bars) when controlled for nonverbal IQ, short-term memory (UNIT MQ), and language development (passive and active vocabulary as measured by ORRIA).

Table 1

Summary Statistics and Psychometric Properties of the Study Measures

		FB Performance				Impairment Status							
		Fail (*n*=32)		Pass (*n*=22)		TD (*n*=22)		DLD (*n*=21)		DLD+LowIQ (*n*=4)		LowIQ (*n*=5)	
Variable	***α***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
Age		86.03	17.95	124.68	21.39	120.23	23.47	81.76	13.61	85.00	15.43	128.20	16.72
Cognitive functioning													
IQ	.89	91.97	12.78	97.27	14.69	100.68	12.37	95.52	10.45	76.50	4.51	75.60	8.65
MQ	.87	86.90	11.31	88.77	16.68	91.27	12.93	88.60	14.50	73.00	3.61	78.60	12.78
SM Words	.92	−.04	.90	.06	1.12	.20	1.08	−.03	.99	.23	.93	−.81	.37
SM Numbers	.88	−.01	.88	.02	1.15	.24	1.10	.13	.72	−.47	.64	−.81	1.35
SM Spatial	.45	−.06	1.02	.08	.96	−.06	.96	−.06	1.03	.61	.76	−.02	1.42
Language development													
ORRIA	.89	77.77	14.56	101.82	16.53	105.20	11.57	72.29	6.34	65.68	19.71	95.65	10.84
PV	.84	5.28	3.12	9.74	2.88	9.56	2.62	4.26	2.70	3.58	2.68	9.81	1.93
AV	.97	7.93	1.77	10.16	1.67	10.14	1.69	6.97	1.36	6.51	1.84	9.90	1.24
LO	.85	6.65	2.04	9.13	2.15	9.80	1.30	5.74	1.44	5.97	1.54	7.70	2.40
WS	.84	8.09	2.54	10.80	1.41	11.14	1.01	7.08	1.16	6.28	4.54	11.10	1.02
SS	.83	6.11	2.75	10.33	2.59	10.55	1.89	5.11	1.70	3.74	4.62	9.19	1.85
PH	.83	8.27	1.21	10.36	1.81	10.71	1.32	7.61	.69	7.50	.55	9.82	1.10
SR	.93	6.80	2.09	9.46	2.00	9.80	1.62	6.20	1.63	6.51	2.56	7.92	2.51
Syntactic Complexity		.04	.02	.07	.02	.07	.03	.04	.02	.04	.03	.06	.02
Embedding		.02	.01	.03	.02	.03	.02	.02	.01	.02	.01	.02	.01
Mental Lexicon		.04	.03	.03	.01	.04	.03	.04	.02	.02	.03	.03	.01

Note. FB – False Belief task. TD – typical development, ORRIA > 85, IQ > 85; DLD – developmental language disorder, ORRIA < 85, IQ > 85; DLD+LowIQ – developmental language disorder (ORRIA < 85) and IQ below 85; LowIQ – ORRIA > 85 and IQ < 85. MQ – Memory Quotient (UNIT). PV – passive vocabulary, AV – active vocabulary, LO – linguistic operators, WS – word structure, SS – sentence structure, PH – phonology, SR – sentence repetition. Reliability coefficient for MQ was obtained by averaging reliabilities for Symbolic Memory, Spatial Memory and Object Memory subtests of UNIT. Reliability coefficient for IQ was obtained by averaging CFIT and UNIT average subtest reliabilities. The total n is 52 for FB Fail and Pass groups combined for all of the language development scores due to missing data described in Method.

Table 2

Intercorrelations

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.
1. Age	1																
2. IQ	.02	1															
3. MQ	.09	.80**	1														
4. SM Words	−.04	.45**	.38**	1													
5. SM Numbers	−.05	.38**	.37*	.53**	1												
6. SM Spatial	.00	.04	.15	.22	.04	1											
7. ORRIA	.80	.34*	.32*	.20	.16	.10	1										
8. PV	.78**	.11	.23	.04	.03	.18	.89**	1									
9. AV	.79**	.25	.29*	.17	.20	.13	.90**	.87**	1								
10. LO	.61**	.30*	.27	.17	.22	.12	.85**	.67**	.72**	1							
11. WS	.72**	.10	.11	−.05	−.04	.06	.88**	.77**	.82**	.78**	1						
12. SS	.79**	.24	.28	.02	−.11	.12	.93**	.86**	.86**	.74**	.85**	1					
13. PH	.78**	.33*	.37**	.14	.19	.01	.98**	.83**	.84**	.79**	.76**	.87**	1				
14. SR	.65**	.31*	.28*	.38**	.28	.16	.83**	.74**	.74**	.77**	.73**	.70**	.73**	1			
15. Syntactic Complexity	.56**	.03	.07	.03	.19	−.09	.50**	.38*	.48**	.33*	.37*	.47**	.49**	.35*	1		
16. Embedding	.44**	.16	.08	.13	.19	.06	.51**	.33*	.46**	.45**	.38**	.44**	.48**	.42*	.78**	1	
17. Mental Lexicon	.02	.11	.09	.14	.09	−.26	.10	.11	−.09	.02	.12	.22	.05	.07	.05	−.01	1

Note. MQ – Memory Quotient (UNIT), PV – passive vocabulary, AV – active vocabulary, LO – linguistic operators, WS – word structure, SS – sentence structure, PH – phonology, SR – sentence repetition.

$p < .05$.

** $p < .01$

Table 3

Relationship between False- Belief Performance and Impairment Status

FB Performance	Impairment Status			
	TD	DLD	DLD+LowIQ	LowIQ
Fail	7 (32%)	18 (86%)	3 (75%)	2 (40%)
Pass	15 (68%)	3 (14%)	1 (25%)	3 (60%)
Total	22 (100%)	21 (100%)	4 (100%)	5 (100%)

Note: percentages for those who passed and failed the False Belief (FB) task in each impairment status group are presented in parentheses. TD – typical development, ORRIA > 85, IQ > 85; DLD – developmental language disorder, ORRIA < 85, IQ > 85; DLD+LowIQ – developmental language disorder (ORRIA < 85) and IQ < 85; LowIQ – ORRIA > 85 and IQ < 85.