Language and False-Belief Task Performance in Children With Autism Spectrum Disorder

M. Jeffrey Farrar, a Hye Kyeung Seung, b and Hyeonjin Lee c

Purpose: Language is related to false-belief (FB) understanding in both typically developing children and children with autism spectrum disorder (ASD). The current study examined the role of complementation and general language in FB understanding. Of interest was whether language plays similar or different roles in the groups’ FB performance.

Method: Participants were 16 typically developing children (mean age = 5.0 years; mental age = 6.7) and 18 with ASD (mean age = 7.3 years; mental age = 8.3). Children were administered FB and language tasks (say- and think-complements), receptive and expressive vocabulary tests, and relative clauses.

Results: When mental age and receptive and expressive vocabulary were used as separate covariates, the typical control group outperformed the children with ASD in FB task performance. Chi-square analyses indicated that passing both complementation tasks was linked to the FB understanding of children with ASD. Children with ASD who passed FB tasks all passed say- and think-complement tasks. However, some children in the control group were able to pass the FB tasks, even if they failed the say- and think-complement tasks.

Conclusion: The results indicate that children with ASD relied more on complement understanding to pass FB than typically developing children. Results are discussed regarding the developmental pathways for FB understanding.

Autism is one of the most prevalent developmental disorders. Children with autism spectrum disorder (ASD) present deficits in two domains: deficits in social communication, and repetitive patterns of behaviors/ restricted interests and activities (American Psychiatric Association, 2013). Children with ASD who are high functioning (i.e., have near average cognitive level) experience social deficits, language delay, language pragmatic deficits, and executive function difficulties (Gernsbacher, Morson, & Grace, 2015; Rapin & Dunn, 2003; Tager-Flusberg, 2007). These social and pragmatic difficulties often have been attributed to theory of mind deficits in children with ASD (Baron-Cohen, 1995; Happé, 1993). It has long been recognized that theory of mind consists of many components, and children with ASD do not always show difficulty in all aspects of it (Broekhof et al., 2015). Furthermore, limitations in theory of mind development do not account for all characteristics of ASD (Tager-Flusberg, 2007). For instance, pragmatic language difficulties may be more related to basic language skills, such as syntax, rather than to theory of mind (e.g., Gernsbacher & Pripas-Kapit, 2012; Whyte, Nelson, & Sherer, 2014). Despite the expansion of the explanations of ASD and some of its characteristics, children with autism often struggle in different theory of mind tasks (e.g., Hutchins et al., 2016).

The current study explored how language contributes to theory of mind development. The role of language in facilitating theory of mind development, particularly false-belief (FB) understanding, has generated considerable debate about both children with ASD and typically developing children (J. G. de Villiers & de Villiers, 2014; Durrleman et al., 2016; San Juan & Astington, 2012). In both groups, language is related to FB performance. However, there is disagreement regarding whether general language or sentential complements is related to FB understanding and whether this relationship varies across populations.

Theory of Mind

Theory of mind has been extensively examined in typically developing children (S. A. Miller, 2015; Wellman, 2014). An important achievement in its development is FB...
understanding, the recognition that different people can have conflicting beliefs about the same situation, such as the content of a box. For instance, in the unexpected content task, children are shown a familiar crayon box and asked about its content. After predicting that there are crayons inside, the child is shown that there are actually pennies inside. After putting the pennies back, children are asked what a puppet will think is in the box even though he has never looked inside. If children understand FBs, they will predict that she will think there are crayons in the box, even though they know there are pennies.

Typically developing children pass FB tasks by approximately 4.5 to 5.0 years of age (Wellman, Cross, & Watson, 2001). In contrast, certain clinical populations, such as children with specific language impairment, ASD, and hearing impairment, are delayed in FB reasoning (P. A. de Villiers, 2005; Farrar et al., 2009; Schick, de Villiers, de Villiers, & Hoffmeister, 2007; Siegal & Peterson, 2008; Tager-Flusberg, 2007). A variety of explanations have been offered for the development of FB understanding, including a specialized theory of mind module (Baron-Cohen, 1995), executive function (Devine & Hughes, 2014; Tager-Flusberg, 2007), and language (Milligan, Astington, & Dack, 2007).

Language and False-Belief

Several different explanations regarding the influence of language on FB reasoning have been proposed. P. A. de Villiers (2005) views the development of sentential complement structures as necessary for representing a FB. Mental state verbs (e.g., think, believe) contain a proposition that builds into a mental representation an embedded clause structure (e.g., “John believes that it is a dog, but it is a wolf.”). This complementation structure allows a split between what John believes and the reality of the situation, which is the essence of FB comprehension.

A memory for complements task is typically used to assess mastery of complements using both mental state verbs (e.g., think) and communication verbs (told, say; e.g., J. G. de Villiers & Pyers, 2002). For example, children are shown a picture and provided a verbal description, such as “She thought/said it was a bug in her hair” and are then shown a second picture with the description “but it was really a leaf.” The child is then asked: “What did she think/say was in her hair?” The use of communication verbs (e.g., say, tell), instead of mental state verbs (e.g., think) in the memory for complements task is important for testing the role of complement structures in FB. If mental state verbs are used, the task reflects both semantic (meaning of think) and syntactic structure components, making it difficult to separate the role of complement semantics from complement syntax per se. The use of communication verbs tests the role of complement syntax separate from meaning. Furthermore, say-complements develop earlier than think-complements and may be a precursor to FB reasoning (P. A. de Villiers, 2005; Perner, Zauner, & Sprung, 2005). They also provide a more stringent test of the role of complementation because memory for false complements involving mental state verbs may actually reflect an understanding of FB (Tardif, So, & Kaciroti, 2007).

Numerous studies have supported the importance of complementation in FB reasoning for typically developing children. In a longitudinal study of 28 typically developing children (age 3–5 years), the children’s performance on a memory for complements task predicted the children’s later FB more so than other language measures (J. G. de Villiers & Pyers, 2002; Low, 2010). Other researchers, however, have suggested that complementation is not required but rather that general language (semantic, syntax) is sufficient (Slade & Ruffman, 2005). These studies have demonstrated that general language using comprehensive standardized assessments (e.g., Test of Early Language Development [Hresko, Reid, & Hammill, 1999]; Clinical Evaluation of Language Fundamentals [Semel, Wiig, & Secord, 1992]; Astington & Jenkins, 1999), mean length of utterance as a morphosyntactic measure (J. Dunn, Brown, Slomkowski, Tesla, & Youngblad, 1991), Index of Productive Syntax (IPSYN; Scarborough, 1990), and receptive/expressive vocabulary (Ng, Cheung, & Xiao, 2010; Slade & Ruffman, 2005) were related to FB performance, but complementation played no unique role or was not examined.

In a similar way, social constructivist approaches to FB argue that children’s participation in linguistically mediated social interactions, involving mental state talk, facilitates FB understanding by highlighting different perspectives on situations (Nelson, 2005; Ruffman, 2014). Thus, there is clear support for the role of language in typically developing children, although there is continuing debate regarding what aspects of language are necessary (Astington & Baird, 2005; Slade & Ruffman, 2005). A meta-analysis of 107 studies of typically developing children reported that all language measures, vocabulary, syntax, and complementation were significantly associated with FB, with the largest effect for complementation. However, only four of the studies had included complementation measures (Milligan et al., 2007).

Language is also related to FB in clinical populations. Studies of verbal children with ASD (Fisher, Happé, & Dunn, 2005; Hale & Tager-Flusberg, 2005; Peterson, Wellman, & Liu, 2005), specific language impairment (Farrant, Fletcher, & Mayberry, 2006; C. Miller, 2006), and who are deaf have and hearing parents (Schick et al., 2007) all show FB delays that are related to their language ability. As with typically developing children, the issue of whether complementation or general language accounts for FB understanding in children with ASD has been a focus of interest, although the number of studies is limited.

Of particular interest are studies that examined the impact of both general language and complementation in the same study. Tager-Flusberg (2000) compared the roles of receptive vocabulary, as assessed by the Peabody Picture Vocabulary Test–Fourth Edition (PPVT-4; L. M. Dunn & Dunn, 2007), complementation structures involving communication verbs (e.g., say), and those involving mental state verbs (e.g., think) in predicting FB for children with ASD.
or intellectual disability. For children with ASD, complementation involving communication verbs was the only predictor of FB. For the children with an intellectual disability, complementation involving mental verbs (e.g., think) was the only predictor of FB. Lind and Bowler (2009) compared children with ASD, children with a learning disability, and typically developing children. For children with ASD, complement syntax of say and told predicted performance on the unexpected locations FB task after controlling for vocabulary, but not the unexpected contents task. Complement syntax was not related to FB performance in the comparison control groups. For these children, general vocabulary was sufficient.

Complementation is related to FB in children with ASD who speak other languages as well. For instance, mental verb (e.g., think) complementation played a significant role in FB task performance in Korean-speaking children with ASD (Lee et al., 2009). Complementation of both communication and cognition verbs was uniquely related to FB understanding in a sample of 17 French-speaking children and adolescents with ASD on verbal FB tasks but not on nonverbal ones (Durrleman & Franck, 2015). Performance on FB tasks in the ASD group was also significantly lower than that in the control group, but not on nonverbal FB and control tasks. However, a subsequent study did report an effect of complementation on nonverbal FB reasoning (Durrleman et al., 2016). In the only longitudinal study of children with ASD, Tager-Flusberg and Joseph (2005) included both complementation and general language measures (IPSYN) as predictors of FB. Age and IPSYN at Time 1 were the strongest predictors of FB at Time 2. Complement understanding involving communication verbs accounted for a unique additional variance.

Thus, across studies there is some inconsistency in which aspect of language is associated with FB in terms of the type of FB task, verbs used, and populations studied (e.g., Cheung, 2006; Low, 2010). The general pattern for children with ASD is that sentential complement syntax may play a necessary role in FB, with some studies finding that only say-complements are linked to FB performance (Tager-Flusberg, 2007) whereas others find that both say- and think-complementation are related (Durrellmann & Franck, 2015). The purpose of the current study was to further examine the role of language in FB understanding in children with ASD. Of particular interest was whether language affects FB development differently in children with ASD than in typically developing children. There is some evidence that complementation is critical to FB understanding in children with ASD (Durrellmann & Franck, 2015; Lind & Bowler, 2009; Tager-Flusberg & Joseph, 2005).

The present study used multiple language measures to assess current language levels. FB performance was compared between ASD and control groups on two types of FB tasks (change of location and change of content tasks) because some studies have demonstrated different effects depending on the FB task (Lind & Bowler, 2009). The study aimed to evaluate the contribution of a set of language variables (receptive vocabulary, expressive vocabulary, complementation, and parents’ perception of their child’s overall syntax) on FB task performance. Furthermore, because children with ASD differ from typically developing children on several dimensions, such as mental age, it is important to either match or control for these factors. Two specific research questions were addressed: (a) What is the relationship between FB performance and language in children with ASD and in typically developing younger children? (b) Does general language and complementation differentially affect FB understanding in the two groups of children?

Method

Participants

Flyers about the study were sent to agencies (regional center, psychology private practice clinics, speech-language clinics, workshops, special educators) in a large urban community. Informed consent forms were mailed to parents who were interested in having their children participate in the study. A total of 42 children were recruited: 24 children with high-functioning ASD (i.e., IQ standard scores above 70 and verbal communicators) between ages 4 and 11 years in the ASD group, and 18 typically developing children between ages 3 and 7 years in the control group. Participants in the ASD group were diagnosed by community professionals on the basis of DSM-IV criteria (American Psychiatric Association, 1994). It should be noted that the data collection was completed before the release of the DSM-5 (American Psychiatric Association, 2013). The diagnosis of those in the ASD group was confirmed by having parents complete the Lifetime form of the Social Communication Questionnaire (SCQ; Rutter, Bailey, Berument, Lord, & Pickles, 2003) and a clinical observation by the second author, who has experience with ASD diagnosis as a certified and licensed speech-language pathologist. Bishop and Norbury (2002) reported a good agreement between total cutoff scores on the Autism Diagnostic Interview–Revised (ADI-R; Le Couteur, Lord, & Rutter, 2003) and the SCQ. Several children in the control group were siblings of the children in another ASD study conducted by the second author.

Consistent with the ASD population, the number of male participants in the ASD group outnumbered the female participants (15 boys and three girls). In the control group, the male participants also outnumbered the female participants (14 boys and two girls). Of the 24 recruits for the ASD group, six were excluded for various reasons. Three recruits did not meet an autism diagnosis on the basis of SCQ score and clinical observation, two had IQ scores lower than 70 and low vocabulary scores (≤ 2 SD on PPVT-4 and/or Expressive Vocabulary Test–Second Edition (EVT-2; Williams, 2007), and one discontinued after vocabulary testing. Thus, the final sample size of the ASD group consisted of 18 children with the following heritages: three Asian, six European, six Hispanic, and...
three Middle Eastern (mean age = 7.3 years; range: 4–11 years). Of the 18 recruits for the control group, two were excluded: One child withdrew from participation after completing vocabulary tests, and one child’s IQ score was lower than 70. Thus, the final sample size of the control group consisted of 16 children with the following heritages: seven Asian, five European, three Hispanic, and one Middle Eastern (mean age = 5 years; range: 3–7 years).

**Materials and Procedures**

Trained graduate students in the Communicative Disorders program tested the participants at a university speech-hearing clinic room equipped with a one-way mirror. Each participant came to the clinic three times. The participants were tested individually while their parents observed their children’s testing through the one-way mirror.

**Baseline Testing**

Each participant took a vocabulary test during the first visit and a cognitive test during the second visit. Participants’ receptive vocabulary level was tested using Form A of the PPVT-4 and expressive vocabulary level was tested using the EVT-2. Age-equivalent scores of the PPVT-4 and the EVT-2 were used for analysis. Participants’ cognitive level was assessed using a brief IQ score on the Kaufman Assessment Battery for Children—Second Edition (KABC-II; Kaufman & Kaufman, 2004). The KABC-II was selected because it reduces ethnically and culturally based differences. The brief IQ scores were derived from the Mental Processing Index, which focuses on mental processing instead of acquired knowledge. The Mental Processing Index consists of sequential processing, simultaneous processing, learning ability (ages 3–6 years), and planning ability (administered to children ages 7–18 years). The subtests administered (i.e., Atlantis, Conceptual Thinking, Story Completion, Face Recognition, Number Recall, Rover, Rebus, Triangle, Word Order, and Pattern Reasoning) varied depending on the age of the participants. Composite IQ standard scores and subscale scores were obtained from the subtests administered. The mental age for each child was computed and was used in the analyses in order to control for any group differences.

Parents in both groups completed the Lifetime form of the SCQ. At the end of their first visit, the research assistant gave the SCQ to the parents and had them complete it at home and bring it to their second visit. A total score higher than 15 on the SCQ is indicative of an ASD diagnosis.

Parents also completed the Children’s Communication Checklist—Second Edition (CCC-2; Norbury, Nash, Baird, & Bishop, 2004), which screens overall communication in speech, syntax, semantics, coherence, initiation, scripted language, context, nonverbal communication, social relations, and interest. The CCC-2 is appropriate to use for children older than 4 years of age who can produce sentences. Therefore, three parents of 3-year-olds in the control group did not complete the CCC-2. General Communication Composite standard scores and scaled scores were obtained. General Communication Composite and syntax subscale scores were used for analysis in order to assess general language abilities. The syntax subscale scores were used in order to examine whether syntax, rather than complementation per se, was associated with FB understanding.

**Experimental Tasks**

Each participant received the experimental tasks during the third visit. Four sets of stimuli (Batteries A–D) were presented to each participant, and the order of the stimuli presentation was counterbalanced. Each set consisted of two FB tasks (change of location and change of content), a language task (say- and think-complementation), and a control language task (relative clause task). The inclusion of the relative clause control task was to assess whether any relation between complementation and FB was a result of complementation per se and not general grammatical ability.

A blank screen was presented at the beginning of each task. Each task was presented with a brief narrative while showing cartoonlike pictures on a computer screen. After presenting an event picture on the computer screen, a research assistant read a brief script of the presented picture (see the Appendix for sample stimuli).

**Change of location task.** A sample narrative was:

*John was playing with a toy plane. John wanted to drink water. John put the toy plane in his backpack. Then he went to the kitchen. John’s father took the toy plane and put it in a box. Then John’s father left the room. John went to the room to play with the toy plane.*

Following the presentation of a narrative, four questions were asked.

Q1. When John goes to the room, where would he look for the toy plane? (FB)
Q2. Why did John look there?
Q3. Where did John put the toy plane?
Q4. Where is the toy plane now?

Q1 above was the FB question and questions 2, 3, and 4 were control questions to assess the participant’s memory.  

**Change of content task.** An examiner presented a closed chocolate box and asked the participant “What do you think is inside the box?” The box was then opened to reveal there are really keys in the box. The box was then closed while the child was looking. Following this presentation, the examiner asked the following five questions.

Q1. What is inside the box?
Q2. When you first saw the box, all closed up like this, what did you think was inside the box? (FB)
Q3. Can you remember what was in the box?

The experimenter takes a puppet out of a closed bag and lets her look at the closed box.
Q4. The puppet has never looked inside the box. What does she think is inside the box? (FB)

Q5. Why does she think that?

Responses on two of the above questions (Q2: self as an agent; Q4: a puppet as an agent) assessed FB understanding. Questions 1 and 3 were control questions.

Complementation language tasks. A sample narrative was:

Paul’s parents were talking. Paul’s father asked, “What is Paul doing?” Paul’s mother said that Paul was sleeping. However, Paul’s mother thought that Paul was playing computer games.

Following the presentation of the narrative, two questions were asked.

Q1: What did Paul’s mother say Paul was doing?
Q2. What did Paul’s mother think Paul was doing?

Responses to Q1 and Q2 were scored for data analysis.

Relative clause control language tasks. A sample narrative was:

This is Helen. These are Helen’s friends. She is brushing her hair. She is brushing her teeth. Helen hugs the girl who is brushing her hair.

Following the presentation of the narrative, language comprehension question was asked.

Q1. Who did Helen hug? If the participants did not respond to Q1, Q2 was presented.
Q2. Did Helen hug the girl who was brushing her hair? Or did Helen hug the girl who was brushing her teeth? 

Response to Q1 or Q2 was scored for analysis.

Scoring

Raw scores, age-equivalent scores, and standard scores were obtained for the PPVT-4 and the EVT-2. Age-equivalent scores were used for analysis because we were interested in the absolute level of performance. For each subtest of the KABC-II, raw scores, scaled scores, percentile ranks, and age-equivalent scores were also computed. Subscale standard scores of sequential, simultaneous, learning, planning processes and composite IQ score (i.e., Mental Processing Index) were also obtained. The KABC-II provides only the composite IQ score and does not provide the age-equivalent score of the composite IQ. Therefore, mental age was derived as a mean score of each subtest age-equivalent score administered on the KABC-II as suggested in the manual (p. 46). The General Communication Composite standard score and syntax subscale score were obtained for the CCC-2.

The range of scores for the FB location task was 0–4, the FB content task was 0–8 (4 for self as an agent, and 4 for puppet as an agent). The range of scores for the FB location control task was 0–8, and the FB content control task was 0–4. To be credited with passing the FB task, children had to pass the control questions. If they failed the control questions on a specific FB task, they were not credited with passing that task. If they missed all the control questions, they were excluded from analyses. We also removed those children who did not pass the control questions but passed the FB tasks. This resulted in a loss of one child from each group. For language tasks, the range for the say-complement was 0–4, the think-complement was 0–4, and the range for the relative clause (language control task) was 0–4. All raw scores were converted to percentages of correct responses for analyses. Trained research assistants completed scoring, and the second author checked their accuracy. Any disagreements were resolved by discussion.

Results

Baseline Data

All data were analyzed using the Statistical Package for Social Sciences (SPSS; IBM Corp, 2014). For all analyses, we conducted Bonferroni corrections to guard against Type 1 errors. For the baseline data on participant characteristic variables, the error rate was set at .01 per variable and the inflation of the error rates was controlled by dividing .01 by the number of comparisons per variable. The four participant characteristic variables were age (chronological age and mental age), IQ (KABC-II composite and subscale standard scores), SCQ, and CCC-2 (General Communication Composite and syntax subscale scores). A descriptive summary of participants’ characteristics within the ASD and control groups is presented in Table 1. These characteristics were compared between groups using analyses of variance.

The SCQ scores were significantly higher in the ASD group than in the control group, $F(1, 31) = 53.13, p = .000, \eta_p^2 = .64$. This result is consistent with the diagnosis of ASD as shown by the mean total score higher than 15. Children in the ASD group were older than those in the control group, $F(1, 32) = 13.92, p = .001, \eta_p^2 = .30$. However, their mental ages were similar, $F(1, 31) = 3.70, p = .064$. We matched the participants on their overall cognitive level by using their mental age. The IQ composite score between groups was not significant, $F(1, 32) = 7.02, p = .012$ (error rate is .0025 by dividing .01 by four comparisons). Overall, their cognitive level was similar. It should be noted that KABC-II subscale scores were based on $n = 18$ in the ASD group and $n = 13$ in the control group because three participants in the control group were 3 years old and subscale scores were not available for children less than 4 years of age. The planning subscale score is applicable only to children older than 7 years of age ($n = 9$ in the ASD group, and $n = 2$ in the control group).

The overall communication composite score on the CCC-2 of the children in the ASD group was significantly lower than the scores in the control group, $F(1, 27) = 30.31, p = .000, \eta_p^2 = .54$. The mean of the CCC-2 syntax subscale score in the ASD group is below 1 SD and lower than the scores in the control group. However, the syntax subscale score was not statistically different between the ASD
and control groups, $F(1, 26) = 5.59, p = .026$. It should be noted that six of the participants did not have CCC-2 scores because their chronological age did not meet the age range (4–16 years).

Table 2 provides the means and standard deviations on the FB and language measures. The overall percent-age correct response rates were examined using analyses of variance and are shown in Figure 1. The percent correct on the total FB task performance was not significantly different between the ASD and control groups, $F(1, 30) = 1.06, p = .31$, or between the FB location and content tasks. The differences between the FB location and content task performance were further examined within each group using paired $t$ tests, and there were no significant differences between the ASD group, $t(16) = −1.81, p = .088$, and the control group, $t(14) = −1.24, p = .24$. The receptive and expressive vocabulary age-equivalent scores also did not differ between groups, nor were comparisons between $say$- and $think$-complement task performances significantly different between the ASD and control groups. Paired $t$ tests between $say$- and $think$-complement were not significant in either the ASD, $t(16) = .35, p = .73$, or control groups, $t(14) = −.53, p = .61$. Finally, there were no group differences on the relative clause control task, $t(30) = .59, p = .46$.

**Table 1. Participant characteristics.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ASD group ($n = 18$)</th>
<th>Control group ($n = 16$)</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Age</td>
<td>7.33</td>
<td>2.17</td>
<td>5.00</td>
</tr>
<tr>
<td>Mental age</td>
<td>8.26</td>
<td>2.89</td>
<td>6.68</td>
</tr>
<tr>
<td>KABC-IIa</td>
<td>98.78</td>
<td>18.38</td>
<td>114.13</td>
</tr>
<tr>
<td>Sequential</td>
<td>98.28</td>
<td>14.92</td>
<td>103.08</td>
</tr>
<tr>
<td>Simultaneous</td>
<td>96.67</td>
<td>20.28</td>
<td>113.23</td>
</tr>
<tr>
<td>Learning</td>
<td>101.71</td>
<td>17.39</td>
<td>122.15</td>
</tr>
<tr>
<td>Planningb</td>
<td>97.00</td>
<td>15.46</td>
<td>114.00</td>
</tr>
<tr>
<td>SCQc</td>
<td>19.31</td>
<td>7.04</td>
<td>5.13</td>
</tr>
<tr>
<td>CCC-2d</td>
<td>75.93</td>
<td>11.30</td>
<td>119.23</td>
</tr>
<tr>
<td>Syntax</td>
<td>6.69</td>
<td>3.28</td>
<td>9.55</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Measure</th>
<th>ASD group ($n = 17$)</th>
<th>Control group ($n = 15$)</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>False-belief total</td>
<td>55.88</td>
<td>30.86</td>
<td>66.63</td>
</tr>
<tr>
<td>False-belief location</td>
<td>42.65</td>
<td>45.72</td>
<td>58.33</td>
</tr>
<tr>
<td>False-belief content</td>
<td>62.50</td>
<td>34.23</td>
<td>70.93</td>
</tr>
<tr>
<td>Receptive vocabularya</td>
<td>8.04</td>
<td>2.94</td>
<td>5.86</td>
</tr>
<tr>
<td>Expressive vocabularya</td>
<td>7.78</td>
<td>2.57</td>
<td>6.22</td>
</tr>
<tr>
<td>Complementation-say</td>
<td>72.06</td>
<td>35.22</td>
<td>65.00</td>
</tr>
<tr>
<td>Complementation-think</td>
<td>70.59</td>
<td>34.50</td>
<td>70.00</td>
</tr>
<tr>
<td>Relative clauses</td>
<td>80.58</td>
<td>20.21</td>
<td>75.16</td>
</tr>
</tbody>
</table>

Note. ASD = autism spectrum disorder.

*aAge-equivalent score, $p < .004$ (adjusting the number of comparisons).
compared with the ASD group. As shown in Table 3, these effects were found for the total and location FB scores, but not for the content task scores.

Similar significant effects were found when receptive and expressive vocabulary scores were used as covariates in separate ANCOVAs. Children with ASD scored lower than the control group in estimated FB total and location performance scores. No group differences were found for the FB content task. Thus, controlling for mental age and vocabulary created a suppressor effect, indicating that FB, language, and mental age are related. Of particular interest was whether the ASD and control groups differed on FB performance, controlling for complementation performance. Controlling for either the say- or think-complement had no impact on FB scores for the groups, as shown in Table 3. Finally, although not reported in Table 3, no FB differences were found when CCC-2 syntax or relative clauses were used as covariates (all ps > .16).

**Pearson Chi-Square Analyses**

We further examined the association between the number of children who passed or failed the say-complement and the think-complement tasks (i.e., percentage of correct performance ≥ 75%) and passed or failed the FB tasks. Separate chi-square analyses were performed for the children with ASD and the control group for the two complement tasks. As seen in Table 4, passing the say-complement task was associated with passing the FB tasks in the ASD group. Every child with ASD who passed the FB task also passed the say-complement task, and no one who failed the say-complement task passed the FB tasks. There were children with ASD who passed the say-complement task but failed the FB tasks. There are likely other cognitive achievements, such as inhibitory control, that also contribute to understanding FB, particularly with children with ASD (Devine & Hughes, 2014; Tager-Flusberg, 2007). However, in the control group, 29% of the children who failed the say-complement task were still able to pass the FB task, and 75% of the children who passed the FB task passed the say-complement task.

### Table 3. ANCOVA results for false-belief (FB) performance (percentage correct) in ASD and control groups.

<table>
<thead>
<tr>
<th>Covariate and DV</th>
<th>F-ratio</th>
<th>$p$</th>
<th>Estimated means</th>
<th>Partial $\eta^2$</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>groups</td>
<td></td>
<td>Control</td>
<td>ASD</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FB total</td>
<td>6.32</td>
<td>.02</td>
<td>72.52</td>
<td>50.71</td>
<td>.18</td>
</tr>
<tr>
<td>FB location</td>
<td>9.83</td>
<td>.005</td>
<td>69.00</td>
<td>33.24</td>
<td>.26</td>
</tr>
<tr>
<td>FB content</td>
<td>1.73</td>
<td>.20</td>
<td>74.28</td>
<td>59.46</td>
<td>.25</td>
</tr>
<tr>
<td>Receptive vocabulary</td>
<td>11.20</td>
<td>.002*</td>
<td>72.65</td>
<td>47.40</td>
<td>.28</td>
</tr>
<tr>
<td>FB total</td>
<td>13.59</td>
<td>.001*</td>
<td>74.16</td>
<td>28.66</td>
<td>.32</td>
</tr>
<tr>
<td>FB location</td>
<td>3.25</td>
<td>.085*</td>
<td>77.32</td>
<td>56.77</td>
<td>.10</td>
</tr>
<tr>
<td>FB content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expressive vocabulary</td>
<td>5.55</td>
<td>.02</td>
<td>72.64</td>
<td>50.60</td>
<td>.16</td>
</tr>
<tr>
<td>FB total</td>
<td>7.85</td>
<td>.01</td>
<td>69.26</td>
<td>33.01</td>
<td>.21</td>
</tr>
<tr>
<td>FB location</td>
<td>1.68</td>
<td>.21</td>
<td>74.35</td>
<td>59.40</td>
<td>.06</td>
</tr>
<tr>
<td>FB content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Say-complement</td>
<td>3.13</td>
<td>.087</td>
<td>51.78</td>
<td>67.29</td>
<td>.09</td>
</tr>
<tr>
<td>FB total</td>
<td>2.79</td>
<td>.105</td>
<td>37.35</td>
<td>59.54</td>
<td>.08</td>
</tr>
<tr>
<td>FB location</td>
<td>1.44</td>
<td>.244</td>
<td>58.96</td>
<td>71.17</td>
<td>.04</td>
</tr>
<tr>
<td>FB content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Think-complement</td>
<td>1.40</td>
<td>.246</td>
<td>53.59</td>
<td>65.24</td>
<td>.04</td>
</tr>
<tr>
<td>FB total</td>
<td>1.38</td>
<td>.249</td>
<td>40.00</td>
<td>56.55</td>
<td>.02</td>
</tr>
<tr>
<td>FB location</td>
<td>0.53</td>
<td>.471</td>
<td>60.51</td>
<td>69.60</td>
<td>.04</td>
</tr>
</tbody>
</table>

Note. DV = dependent variable (FB total, FB location, FB content); ASD = autism spectrum disorder.

*aCovariate is significant. Receptive and expressive vocabulary scores are age-equivalent scores. bCovariate is not significant.

*Significance level set at .003.
Finally, Table 5 displays the chi-square results for the think-complement. The association of passing think-complement with passing FB tasks was statistically significant in the ASD group but not in the control group. As with the say-complement task, every child with ASD who passed the think-complementation task passed the FB task. However, in the control group, 50% of the children who passed the think-complement task failed FB tasks. As with say-complementation, passing the think-complementation task was not necessary to pass FB tasks in the control group.

Table 4. Chi-square analyses of likelihood of passing false-belief (FB) tasks as a function of say-complement scores.

<table>
<thead>
<tr>
<th>Group</th>
<th>Say fail</th>
<th>Say pass</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FB fail</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>FB pass</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>12</td>
<td>17</td>
</tr>
</tbody>
</table>
| χ²(1, n = 17) = 3.86.  
  p < .05 |
| Control |          |          |       |
| FB fail | 5       | 2        | 7     |
| FB pass | 2       | 6        | 8     |
| Total | 7       | 8        | 15    |
| χ²(1, n = 15) = 3.23.  
  p = ns |

Note.  ASD = autism spectrum disorder.

The chi-square analyses indicated, however, that complementation was critical for children with ASD to pass the FB tasks. Not a single child with ASD passed the FB tasks without passing both the say- and think-complement tasks. This was not the case for control children, who were able to pass the FB tasks without necessarily passing the complementation task.

Table 5. Chi-square analyses of likelihood of passing false-belief (FB) tasks as a function of think-complement scores.

<table>
<thead>
<tr>
<th>Group</th>
<th>Think fail</th>
<th>Think pass</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FB fail</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>FB pass</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>12</td>
<td>17</td>
</tr>
</tbody>
</table>
| χ²(1, n = 17) = 386.  
  p < .05 |
| Control |          |            |       |
| FB fail | 3       | 5          | 8     |
| FB pass | 3       | 5          | 8     |
| Total | 6       | 10         | 16    |
| χ²(1, n = 15) = .134,  
  p = ns |

Note.  ASD = autism spectrum disorder.

Discussion

The present study investigated whether general language and complementation abilities differentially affected FB understanding in children with ASD compared with a typically developing control group. In initial group comparisons, the groups did not differ on FB understanding. In a series of ANCOVAs, the children with ASD scored significantly lower on FB tasks compared with the control group. To be specific, when mental age and receptive or expressive vocabulary age scores were used as covariates, there were significant FB performance differences between the ASD and control groups. These results suggest that controlling for mental age and vocabulary affected FB development. This has been demonstrated in numerous studies of typically developing children and children with ASD (Cheung, 2006; see Milligan et al., 2007, for a review). One limitation of the use of vocabulary measures as a proxy for general language is that it does not assess general syntactic ability, which is also related to FB (Slade & Ruffman, 2005). In the current study, controlling for a parental measure of children’s expressive syntax level and relative clause performance was unrelated to FB understanding. However, a more standardized assessment of syntax may be informative in future studies.

Of particular theoretical interest was whether controlling for complementation affected FB understanding. Prior studies have suggested that memory for complements is particularly critical for children with ASD. Although studies have differed on finding whether say- or think-complements are required, Tager-Flusberg and Joseph (2005) reported that complementation measures involving communication verbs (say and told) were the only predictor of FB in children with ASD. Lind and Bowler (2009) found that complement syntax with communication verbs predicted performance on the unexpected locations FB task after controlling for vocabulary, but not on the unexpected contents task. Lind and Bowler did not test for mental state verbs. In contrast, Lee et al. (2009) reported that mental state verb (think) complements, but not say-complements, were a significant contributor to FB task performance in Korean-speaking children with ASD. A recent study of French-speaking children found effects of both say- and think-complementation (Durreleman & Franck, 2015).

In the present study, controlling for either complement measure did not affect group difference in FB performance in the ANCOVAs. The chi-square analyses, however, clearly demonstrated that both say- and think-complements were significantly associated with FB understanding in children with ASD. Children with ASD who passed FB tasks...
clinical implications/future directions

future studies should address some of the other limitations of the current study by including (a) more homogeneous age ranges in both groups (ages 6–8 years in the ASD group; ages 4–6 years in the control group), (b) ADI-R and/or the Autism Diagnostic Observation Schedule (ADOS) measures, (c) direct syntax measures rather than parent report measures (CCC-2 syntax), and (d) longitudinal studies with larger sample sizes to investigate the emergence of FB in both populations. Participants should understand the semantic distinction between say- and think-complements as well as syntactic differences between them. Our tasks are not completely independent. In future studies, the linguistic stimuli should be more carefully selected to avoid any overlap between complement stimuli.

These results also have intervention implications for children with ASD (Hutchins & Prelock, 2008). A few training studies have examined whether training complement understanding affects FB performance in typically developing children (Hale & Tager-Flusberg, 2003; Lohmann & Tomasello, 2003) or low-income preschoolers (Tompkins, 2015b). Hale and Tager-Flusberg (2003) reported that typically developing children who received training on sentential complements improved performance on sentential complements as well as on a FB task. Thus, FB as an identified deficit skill could be targeted in a training study for children with ASD by training on complementation, although it certainly would be expected that not all children could be trained because of various demographic or cognitive characteristics. The result of training studies in typically developing children provides a basis for exploring the effectiveness of training on FB performance in children with ASD.

References


Dunn, L. M., & Dunn, D. M.

Farrar, M. J., Johnson, B., Tompkins, V., Easters, M., Zilis-


Scarborough, H. S.
Appendix (p. 1 of 4)
Sample Stimuli

Figure A1. John plays with the toy plane.

Figure A2. John wanted to drink water.
Sample Stimuli

**Figure A3.** John put the toy plane in his backpack.

**Figure A4.** John went to the kitchen.
Appendix (p. 3 of 4)

Sample Stimuli

**Figure A5.** John’s father went into the room. He thought “the toy plane might break in the backpack.”

**Figure A6.** John’s father took the toy plane from the backpack and put it in a box.
Appendix (p. 4 of 4)

Sample Stimuli

**Figure A7.** John's father left the room.

**Figure A8.** John went to the room to play with the toy plane.