ABSTRACT

In an effort to understand cognitive foundations of oral language comprehension (i.e., listening comprehension), we examined how inhibitory control, theory of mind, and comprehension monitoring are uniquely related to listening comprehension over and above vocabulary and age. A total of 156 children in kindergarten and first grade from high-poverty schools participated in the study. Using structural equation modeling, results showed that all three cognitive skills (inhibitory control, theory of mind, and comprehension monitoring) were positively related to listening comprehension after accounting for vocabulary and age. In addition, inhibitory control had a direct relation to listening comprehension, not indirectly via theory of mind. Results are discussed in light of cognitive component skills for listening comprehension.

Oral language comprehension is the foundational skill not only in social interactions but also in literacy acquisition (e.g., Catts, Fey, Zhang, & Tomblin, 1999; Snow, Burns, & Griffin, 1998). Along with word reading, oral language comprehension is one of the two necessary skills for reading comprehension according to the simple view of reading (Hoover & Gough, 1990). Evidence for the role of oral language comprehension in reading comprehension has been demonstrated across languages with varying depths of orthography, including English, Spanish, Chinese, and Korean (Hoover & Gough, 1990; Joshi, Tao, Aaron, & Quiroz, 2012; Kendeou, van den Broek, White, & Lynch, 2009; Kim, Park, & Wagner, 2014; Vellutino, Tunmer, Jaccard, & Chen, 2007).

Successful text comprehension requires establishing a coherent mental representation called the situation model and involves integrating text content with prior knowledge (Gernsbacher, 1994; Kintsch, 1988, 1998). According to the construction–integration model of text comprehension, the situation model is built on construction and integration of propositions (Kintsch, 1988) and draws on linguistic skills, such as vocabulary and syntactic knowledge as well as background knowledge (Dickinson, Golinkoff, & Hirsh-Pasek, 2010; Florit, Roch, Altoè, & Levorato, 2009; Lepola, Lynch, Laakkonen, Silvén, & Niemi, 2012). In addition, higher level cognitive skills such as inference making and comprehension monitoring have been suggested to be important to text comprehension (e.g., Cain & Oakhill, 2007; Cain, Oakhill, & Bryant, 2004; Tompkins, Guo, & Justice, 2013).

However, many previous studies examined the relation of cognitive skills to reading comprehension but not listening comprehension (Cain & Oakhill, 2007; Cain et al., 2004; Seigueric & Ehrlich, 2005; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000; see National Institute of Child Health and Human Development, 2000, for a review). Consequently, we have a limited understanding about cognitive foundations of text comprehension in
oral texts (hereafter, listening comprehension). Therefore, the primary goal of the present study was to examine how multiple cognitive skills (inhibitory control, theory of mind [ToM], and comprehension monitoring) are related to listening comprehension, particularly narrative comprehension, for young children after accounting for oral language skill such as vocabulary.

Inhibitory Control

Inhibitory control is the ability to suppress a dominant response and initiate a subdominant response (e.g., an opposite response; Nigg, 2000; Rothbart, Ahdai, & Hershey, 1994). As part of self-regulatory executive functioning, the ability to control one's attention to behave in a particular manner is important for social behaviors and regulating emotions (Eisenberg et al., 2005; Posner & Rothbart, 2000), as well as academic achievement, such as reading (e.g., Cartwright, 2012; Kiefer, Vukovic, & Barry, 2013; St. Clair-Thompson & Gathercole, 2006). Because inhibitory control facilitates the ability to resist distraction and sustain focused concentration, it is reasonable to expect its relation to oral language development (Joseph, McGrath, & Tager-Flusberg, 2005). That is, during listening activities, inhibitory control would help one suppress irrelevant and extraneous information from entering working memory and allow for focusing attention on critical, relevant information and suppressing attention from returning to information that is no longer pertinent (Stoltzfus, Hasher, Zacks, Ulivi, & Goldstein, 1993).

Despite this potential importance of inhibitory control in language acquisition (Bishop & Norbury, 2005; Wolfe & Bell, 2004), however, only a few studies have examined the relation of inhibitory control to language skills. Inhibitory control was related to preschoolers' vocabulary (Moriguchi, Okanda, & Itakura, 2008; Wolfe & Bell, 2004) and adults' nonliteral language comprehension, which requires suppression of literal meaning (e.g., she has a heavy heart; Champagne, Desautels, & Joanette, 2004).

Although these studies suggest a relation between inhibitory control and language skills, it is unclear whether inhibitory control is related to listening comprehension, a discourse-level language skill. If inhibitory control is related to listening comprehension, an important corollary is mechanism of the influence. In particular, inhibitory control has been suggested to be a predictor of ToM, even after accounting for working memory and planning (Carlson & Moses, 2001; Carlson, Moses, & Claxton, 2004; Chasiotis, Kiessling, Hofer, & Campos, 2006; Flynn, O'Malley, & Wood, 2004). Understanding others' mental states (ToM) might require some minimum level of inhibitory control to direct and sustain attention and suppress unimportant information. However, little is known about whether inhibitory control is directly related to listening comprehension over and above other cognitive skills, such as ToM, or whether the potential relation between inhibitory control and listening comprehension is mediated by ToM.

In the present study, we tested three alternative models to understand the relation of inhibitory control to listening comprehension: a model in which inhibitory control has a direct relation to listening comprehension, a second model in which inhibitory control has an indirect relation via ToM, and a third model in which inhibitory control has a direct and indirect relation to listening comprehension. An indirect relation is plausible because ToM relies on comprehension of statements requiring the child to suppress inappropriate literal meaning or irrelevant information and make inferences about characters' intentions or understanding. A direct relation between inhibitory control and listening comprehension over and above ToM is plausible because listening comprehension might draw on one's ability to suppress contextually inappropriate information and activate relevant and appropriate information (i.e., inhibitory control) to a larger extent than does ToM.

ToM

ToM is the ability to infer others' mental states and perspectives and predict behaviors, and is typically measured by false-belief tasks (Howlin, Baron-Cohen, & Hadwin, 1999). False-belief tasks measure children's understanding or recognition of mental states of others using a specific framework of telling a child a short story and then asking the child to predict, for instance, where the protagonist will search for a hidden object or enact a particular action, irrespective of where the object is really located. Evidence indicates that whereas 3-year-old children score no better than chance on first-order false-belief measures (a commonly used indicator of ToM in early childhood), 4- and especially 5-year-old children typically demonstrate mastery on these tasks (Wellman, Cross, & Watson, 2001; Wellman, Fang, & Peterson, 2011).

In first-order false-belief tasks, children are asked to identify what someone else thinks (e.g., a character in a presented vignette), given that the character has access to different information than the child. Progression to mastering second-order false-belief tasks begins thereafter and appears to solidify in typical children in middle childhood (Coull, Leekam, & Bennett, 2006; Miller, 2009). These more challenging tasks require a child to identify the mental state of a third party from the perspective of a second party (e.g., if Jane is read a
story about a hidden toy, can she correctly ascertain what the character Devin thinks the other character Maya thinks about where the toy is?). Constructing a coherent mental representation of the meaning in oral texts requires making inferences of unstated ideas to build propositions and connecting propositions for local and global coherence (Kintsch, 1988; van den Broek, Kendoue, Lousberg, & Visser, 2011). Recent studies suggested that inference making is uniquely related to listening comprehension (Kendoue, Bohn-Gettler, White, & van den Broek, 2008; Lepola et al., 2012; Tompkins et al., 2013). Because ToM captures the ability to make inferences about others’ intentions and thoughts, it is reasonable to hypothesize that ToM is related to listening comprehension. That is, ToM might be important for capturing perspectives of interlocutors, story characters, and storytellers, which will lead to constructing the situation model, particularly in narrative stories. However, previous studies have focused on the relation of ToM either to specific aspects of language skills that require more social cognition or to discrete and relatively simple semantic and syntactic skills (e.g., Caillies, Hody, & Calmus, 2012; Johnston, Miller, & Tallal, 2001; Nilsen, Glenwright, & Huyder, 2011; Norbury, 2005; Slade & Ruffman, 2005). The present study aimed to fill this gap and examine whether ToM is related to listening comprehension, specifically narrative comprehension, over and above other potential language and cognitive skills.

Comprehension Monitoring

Comprehension monitoring is to the ability to reflect on and evaluate one’s own comprehension of text (Baker, 1984; Kinnunen, Vauras, & Niemi, 1998; Ruffman, 1999). An inconsistency detection paradigm has been widely used to examine comprehension monitoring. In an inconsistency detection task, children are examined about the extent to which they detect whether ideas and thoughts in written or oral texts are consistent or contradictory (Baker, 1985; Ruffman, 1999; regarding other tasks, see Skarakis-Doyle, 2002; Skarakis-Doyle, Dempsey, & Lee, 2008). Children will notice and identify problems only if they are evaluating and monitoring their comprehension (Baker, 1985).

Constructing a coherent situation model might require comprehension monitoring because after children construct initial propositions based on the text, these potentially incorrect propositions have to be evaluated for errors and incoherence (Kintsch, 1988). Despite its recognized importance, the majority of previous studies have mainly examined comprehension monitoring in written texts (Cain et al., 2004; Oakhill & Cain, 2007; Oakhill, Cain, & Bryant, 2003), and comprehension monitoring is widely promoted as a reading comprehension strategy (e.g., Block & Pressley, 2002; Griffin, Malone, & Kame’enui, 1995).

In contrast, our understanding of comprehension monitoring in an oral language context is limited. Extant studies have shown that typically developing young children pay attention to consistency cues at the sentence or referential level (Baker, 1984; Revelle, Wellman, & Karabenick, 1985; Skarakis-Doyle, 2002). However, it is not clear how individual differences in comprehension monitoring are related to listening comprehension over and above other potential cognitive skills, such as ToM and inhibitory control, and a language skill such as vocabulary.

Present Study

In the present study, we aimed to expand our understanding of cognitive foundations of language comprehension. Although studies examined the relation of several cognitive skills, such as inhibitory control, ToM, and comprehension monitoring, to children’s language skills, these studies examined a respective cognitive skill separately and/or discrete aspects of language skills (Gillott, Furniss, & Walter, 2004; Moriguchi et al., 2008; Wolfe & Bell, 2004). Given interrelations among cognitive skills (e.g., Scullin & Bonner, 2006), it is important to examine how these cognitive component skills are independently related to discourse-level language skill that requires construction of the situation model (i.e., listening comprehension).

The following were specific research questions in the present study:

1. Are comprehension monitoring and ToM related to listening comprehension of narrative texts?
2. How is inhibitory control related to listening comprehension of narrative texts? Is it directly related to listening comprehension or indirectly related to listening comprehension via ToM?

Based on the literature reviewed previously, three alternative models were tested. In all the models, ToM and comprehension monitoring were hypothesized to have direct relations to listening comprehension, whereas the relation of inhibitory control to listening comprehension varied: In model 1, inhibitory control has a direct path to listening comprehension (see Figure 1a); in model 2, inhibitory control has only an indirect path to language comprehension via ToM (see Figure 1b); and in model 3, inhibitory control has a direct path as well as an indirect path via ToM to listening comprehension (see Figure 1c). Given the role of vocabulary in listening comprehension (Florit et al., 2009;
Method

Participants

A total of 156 children (95 boys) in kindergarten \((n = 74)\) and first grade \((n = 82)\) in high-poverty public schools in northern Florida participated in the study. Children from high-poverty schools were recruited because studies have shown that these children have low oral language skills (Fish & Pinkerman, 2003; Hart & Risley, 1995; Snow et al., 1998), and thus it is important to understand cognitive predictors that contribute to listening comprehension. This would allow us to have a preliminary understanding about instructional targets to improve their listening comprehension. The mean children's age was 74.19 months (standard deviation = 7.94 months), ranging from 60 to 97 months. Approximately 90% of these children were eligible for free or reduced-price lunch.

The children's racial and ethnic backgrounds were as follows: 41% white \((n = 64)\), 38% African American \((n = 59)\), 5% Hispanic \((n = 7)\), 14% multiracial \((n = 22)\), and 2% others \((n = 4)\). According to school records, 26 children had a speech impairment, three had a language impairment, one had a specific learning disability, and one was designated as gifted. Consistent with the school district's demographics, most of the children were native speakers of English, and less than 5% of the sample were English learners, students for whom a language other than English is used in the home and have limited English proficiency. Note that patterns of relations were essentially the same when children with a language or speech impairment were excluded in the data analysis. Therefore, all the children were included in the analysis.

Measures

Listening Comprehension

Children's listening comprehension was measured by four tasks. The first was the Listening Comprehension Scale of the Oral and Written Language Scales (OWLS; Carrow-Woolfolk, 1995). In this task, the child hears sentences and is asked to point to the picture that best describes each heard sentence. Cronbach's αs were reported to be .78 and .82 for 5- and 6-year old children, respectively. The second task was the narrative comprehension subtests of the Test of Narrative Language (TNL; R.B. Gillam & Pearson, 2004). The TNL includes three narrative comprehension tasks (tasks 1, 3, and 5) in which the child listens to stories and is asked comprehension questions. These questions were primarily related to the narrative text's structural elements, such as characters, problems, resolutions, and key details. Cronbach's αs were reported to be .84 and .82 for 5- and 6-year old children, respectively.

In addition, two experimental narrative listening comprehension tasks were developed. In these tasks, children heard a story and were asked comprehension questions related to characters, events, problem, resolution, and details. The stories were 247 and 365 words long and had eight and seven comprehension questions, respectively. Many questions were scored dichotomously, but several were scored on a 0–2 scale based on the following rubric: 0 for an incorrect answer, 1 for a partially correct answer, and 2 for a precise answer. Inter-rater agreement was .99. Note that language used...
in these tasks was primarily straightforward without much figurative language. Although many questions were factual recall questions (e.g., character’s name), some required making inferences (e.g., “Why did the character feel that way?”).

Inhibitory Control
The day/night task (Livesey, Keen, Rouse, & White, 2006) was adapted. In this task, the child was required to suppress his or her dominant response and respond the opposite, saying the opposite to the stimuli shown. Three pairs of stimuli were day/night, boy/girl, and snow/green. For example, the child was shown pictures of a boy and a girl. After confirming the pictures (a boy and a girl), the child was asked to point to an opposite picture (e.g., when the examiner points to a boy, the correct answer is a girl). The child’s response was scored on a scale of 0–3: 0 for failing to point, 1 for pointing at an incorrect answer, 2 for a corrected answer (self-correction), and 3 for a correct response. There were six practice items and 18 experimental items. The total possible maximum score was 54 (3 × 18). Cronbach’s α was .94.

ToM
Three second-order false-belief scenarios (e.g., “Sam believes that Maria thinks…”) were employed. Given that children were in kindergarten and first grade, second-grade false-belief tasks were used to examine children’s ability to infer a story character’s false belief about another character’s knowledge. The scenario context included a bake sale, a birthday celebration, and feeding ducks. The assessor presented a series of four illustrations to the child and explained the context. Then, the assessor asked the child a series of questions involving characters’ beliefs, such as “Maria has not gotten to the bake sale yet. Does Maria know they are selling pumpkin pie at the bake sale?” “What does Sam think they are selling at the bake sale?” and “Why does he think that?” There were six questions in each task, and thus the total possible maximum score was 6 for each task. Cronbach’s α estimates were .73, .73, and .71, respectively, for each task.

Comprehension Monitoring
An inconsistency detection task adapted from Baker’s (1984) work was used to assess children’s comprehension monitoring in oral language contexts. Baker’s study included 5-, 7-, and 9-year-old children, and the inconsistency sentences consisted of 7–9 sentences and 69–82 words. Our adaptation for younger children included shortening the stories to three or four sentences. In this task, the child was asked to identify whether the story made sense and, if not, to provide a brief explanation. An example of an inconsistent item is as follows: “Shelly’s favorite color is blue. She wears blue every day. Shelly has blue pants and even blue shoes. Shelly likes to have everything purple!” An example of a consistent and coherent story is as follows: “Deborah’s favorite thing to do is dance. Deborah dances everywhere she goes. She dances at home. She dances at a park. Deborah even dances when she watches TV.” The meaning of “not making sense” was explained as sentences not going together. Feedback and explanations were provided during practice items.

There were two practice items and 20 experimental items. Consistent and inconsistent stories (10 items each) were randomly spread across items. For inconsistent stories, the accuracy of children’s explanation was scored, and thus a total possible score was 30 (20 for 20 total items + 10 for explanation of inconsistent stories). In the inconsistent example above, to receive credit for the accuracy of explanation, the child has to indicate that the last sentence does not go together with the fact that Shelly’s favorite color is blue. Children’s responses were scored dichotomously (1 = correct; 0 = incorrect). Acceptable correct responses for explanations were clearly delineated based on children’s responses from a pilot study. Inter-rater reliability was .99, and Cronbach’s α was .89.

Vocabulary
Children’s vocabulary was assessed using the Picture Vocabulary subtest of the Woodcock–Johnson III (Woodcock, McGrew, & Mather, 2001), which requires children to name pictured objects. Reliability (Cronbach’s α) was reported to be .70 for 6-year-olds.

Procedures
Graduate research assistants were rigorously trained about assessments and working with children. Children were assessed individually in a quiet room in the schools at the beginning of the school year. Research assistants were also trained on scoring using examples from a previous study until they reached 99% inter-rater agreement with a master scorer.

Results

Descriptive Statistics and Preliminary Analysis
Table 1 shows means, standard deviations, and minimum and maximum values of the measures. Standard scores and percentile ranks are also reported where available. Children’s listening comprehension skills, on average, were in the low average range with mean percentile ranks of 28 and 34 on the OWLS and TNL listening comprehension tasks, respectively. In contrast,
children’s mean vocabulary score was higher, at the 48th percentile. Children’s performance on the inhibitory control was negatively skewed (−2.52) with a sign of a ceiling effect, a limitation of the present study. Because transformations did not make a difference in terms of the distribution pattern, raw scores were used in the present study. In addition, the ToM tasks were skewed right with skewness values of 1.05, 1.05, and 1.28 for tasks 1, 2, and 3, respectively.

Difficult items in the ToM tasks were those that required explanations that second characters’ knowledge is not updated. To illustrate, a dad and his child want to go to the zoo. While the dad is gone, the child learns that the zoo is closed, and decides to go to the swimming pool instead. However, the dad does not know about this new information. When the student in the present study was asked where the dad thinks his son thinks that they will be going (correct answer: zoo), about a third of the sample children in the present study were able to provide correct responses. However, only few children (e.g., 6%) were able to explain why the dad thinks that (because the dad did not know about his child’s change of plan). Although not ideal, the skewness values are borderline, and the use of a latent variable approach reduces its impact because the latent variable represents common variance among observed variables. Partial bivariate correlations between observed variables after controlling for age are presented in Table 2. Listening comprehension measures were moderately to fairly strongly related with one another (.35 ≤ rs ≤ .64). The cognitive measures were weakly to moderately related to listening comprehension measures (.26 ≤ rs ≤ .55), and similar relations were observed among cognitive measures (.14 ≤ rs ≤ .57). Raw scores were used in the subsequent analyses.

Relations of Cognitive Skills to Listening Comprehension

Latent variables were constructed for listening comprehension and ToM, which had multiple measures. The confirmatory factor analysis model in which ToM and listening comprehension were considered as two

<table>
<thead>
<tr>
<th>TABLE 1 Descriptive Statistics</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Listening comprehension</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OWLS listening comprehension—raw</td>
<td>40.65</td>
<td>8.95</td>
<td>16</td>
<td>72</td>
</tr>
<tr>
<td>OWLS listening comprehension—SS</td>
<td>89.27</td>
<td>11.78</td>
<td>59</td>
<td>126</td>
</tr>
<tr>
<td>OWLS listening comprehension—PR</td>
<td>28.27</td>
<td>22.04</td>
<td>3</td>
<td>96</td>
</tr>
<tr>
<td>TNL narrative comprehension—raw</td>
<td>20.05</td>
<td>6.79</td>
<td>3</td>
<td>33</td>
</tr>
<tr>
<td>TNL narrative comprehension—SS</td>
<td>8.10</td>
<td>3.34</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>TNL narrative comprehension—PR</td>
<td>33.53</td>
<td>27.76</td>
<td>1</td>
<td>98</td>
</tr>
<tr>
<td>Experimental story comprehension 1</td>
<td>4.46</td>
<td>2.28</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Experimental story comprehension 2</td>
<td>3.84</td>
<td>2.56</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td><strong>Theory of mind</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theory of mind task 1</td>
<td>1.49</td>
<td>1.65</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Theory of mind task 2</td>
<td>1.36</td>
<td>1.51</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Theory of mind task 3</td>
<td>1.31</td>
<td>1.52</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Inhibitory control (day/night)</td>
<td>48.96</td>
<td>8.32</td>
<td>18</td>
<td>54</td>
</tr>
<tr>
<td>Comprehension monitoring (inconsistency detection)</td>
<td>13.73</td>
<td>6.82</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>WJ–III picture vocabulary—raw</td>
<td>17.39</td>
<td>2.82</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td>WJ–III picture vocabulary—SS</td>
<td>97.75</td>
<td>12.67</td>
<td>20</td>
<td>126</td>
</tr>
<tr>
<td>WJ–III picture vocabulary—PR</td>
<td>47.61</td>
<td>21.51</td>
<td>5</td>
<td>96</td>
</tr>
</tbody>
</table>


*The TNL SS is on a scale of 10 (standard deviation = 3).
separate latent variables yielded an excellent model fit: χ²(13) = 19.58, p = .11; comparative fit index (CFI) = .99; Tucker–Lewis index (TLI) = .98; root mean square error of approximation (RMSEA) = .057; and standardized root mean square residual (SRMR) = .03. In contrast, the model in which ToM and listening comprehension indicators were forced to be a single latent variable yielded a poorer fit: χ²(14) = 52.73, p < .001; CFI = .92; TLI = .89; RMSEA = .13; and SRMR = .053. The two-latent-variable model had a superior fit: Δχ² (Δdf = 1) = 33.15, p < .001.

Measurement models for ToM and listening comprehension latent variables were all appropriate, with significant loadings of all the indicators (ps ≤ .001), and factor loadings are similar to what is shown in the final model in Figure 2.

Structural equation modeling was employed to examine the unique relations of cognitive skills to listening comprehension over and above children’s vocabulary and age. Model fits were evaluated by the following indexes: chi-square, CFI, TLI, RMSEA, and SRMR. Generally, RMSEA values below .085, CFI and TLI values greater than .95, and SRMR values below .05 indicate an excellent model fit (Hu & Bentler, 1999; Kline, 2005). CFI and TLI values greater than .90 and SRMR values less than .10 indicate a good model fit (Kline, 2005).

The three alternative models shown in Figure 1 were tested. The fit of the direct model (Figure 1a) was good: χ²(33) = 52.89, p = .02; CFI = .97; TLI = .96; RMSEA = .06; and SRMR = .04. The indirect model (Figure 1b) yielded a poor fit to the data: χ²(38) = 228.66, p < .001; CFI = .71; TLI = .63; RMSEA = .18; and SRMR = .31. The third model (direct and indirect model; Figure 1c) also yielded a poor fit: χ²(36) = 114.06, p < .001; CFI = .88; TLI = .84; RMSEA = .12; and SRMR = .14. The second and third models were both statistically poorer than the first model (ps < .001), with chi-square differences of 175.77 and 61.17 and associated Δdf of 5 and 3, respectively. In other words, the model in which inhibitory control is directly related to listening comprehension (Figure 1a) was superior to the indirect model (Figure 1b) and the indirect and direct model (Figure 1c). Therefore, the direct model (Figure 1a) was chosen as the final model.

Results with standardized coefficients of the final model are presented in Figure 2. ToM (β = .45, p < .001), comprehension monitoring (β = .21, p < .001), and inhibitory control (β = .20, p < .001) were all related to listening comprehension over and above vocabulary and the other cognitive variables. A total 82% of variance in listening comprehension was explained.

**Discussion**

Listening comprehension is a complex cognitive activity. Although linguistic foundational skills (e.g., vocabulary) are required, they are not sufficient for successful
listening comprehension. Thus, in the present study, we examined cognitive component skills of listening comprehension of narrative texts, such as inhibitory control, ToM, and comprehension monitoring, after accounting for vocabulary and age. Inhibitory control, ToM, and comprehension monitoring were weakly to moderately related to one another (.14 ≤ \( r_s \) ≤ .57; see Table 2), confirming previous studies (e.g., Bjorklund & Pellegrini, 2002; Chasiotis et al., 2006; Scullin & Bonner, 2006). However, all three cognitive skills in the present study were independently related to listening comprehension of narrative texts after accounting for one another as well as vocabulary and age. Altogether, the predictors in the present study explained a large amount of total variance in listening comprehension (i.e., 82%).

Comprehending oral texts (i.e., listening comprehension) requires focused attention to process a large amount of linguistic information. Therefore, the ability to suppress irrelevant stimuli (i.e., inhibitory control) was hypothesized to be a necessary cognitive skill to build the situation model. The present findings support this hypothesis. A related, important question was the path of influence of inhibitory control on listening comprehension. Previous studies suggested that inhibitory control, particularly a conflict inhibition task such as the one used in the present study, might be implicated for ToM development (Chasiotis et al., 2006). When we tested whether the influence of inhibitory control on listening comprehension was direct, indirect via ToM (e.g., Carlson & Moses, 2001; Carlson et al., 2004; Chasiotis et al., 2006), or direct as well as indirect, data supported the direct model. Thus, although individual differences in inhibitory control are related to ToM (\( r = 34 \)), it appears that inhibitory control has a direct influence on listening comprehension over and above ToM, suggesting that the child’s ability to inhibit or discount distracting, nonrelevant information or thoughts is important to listening comprehension over and above ToM. Given that inhibitory control is critical for attentional control (i.e., supervisory attention; see Lijffijt, Kenemans, Verbaten, & van Engeland, 2005, for a review) and attention has been suggested to be important for language development (Mundy, Sigman, & Kasari,
and therefore the comprehension aspect captured in the inconsistency detection task might have led to the present findings. However, this does not appear to explain the relation found in the present study. Because ToM tasks also required comprehending the heard sentences, if the comprehension aspect required in comprehension monitoring and ToM was the driving force for the observed relations, then the effects of comprehension monitoring and ToM would have largely overlapped and thus would not be independently related to listening comprehension. In other words, the unique aspect of comprehension monitoring (i.e., detecting inconsistencies) is likely to be the explanation for the findings in the present study.

Limitations, Future Directions, and Implications

Several limitations and related potential future directions are worth noting. First, potential other important language and cognitive skills were not included in the present study. Working memory has been shown to be important for listening comprehension (Florit et al., 2009) but was not included in the present study. Working memory has been hypothesized to be important for the integration process to establish local and global coherence (Cain et al., 2004), and thus it will be important to include in a future study. Note, however, that a recent study revealed that the influence of working memory on listening comprehension was completely mediated by ToM and comprehension monitoring (Kim, in press). In addition, studies have shown that syntactic skills are important for ToM as well as listening comprehension (Kim, in press; Slade & Ruffman, 2005). Thus, future studies should include syntactic knowledge.

Second, some tasks suffered from some floor (ToM) and ceiling effects (inhibitory control). As for the ToM tasks, given the known association between language ability and ToM, the sample children’s relative weaknesses in language skills likely suppressed performance on the false-belief tasks. This may be particularly true given that the tasks were second-order false-belief measures, which are known to be more challenging than those that tap the earlier developing first-order skills (Miller, 2009). Additionally, all the ToM tasks for ToM were of a similar nature, and thus inclusion of other tasks (e.g., appearance/reality; Gopnik & Astington, 1988) is needed. As for inhibitory control, although statistically significant relations were found between inhibitory control and listening comprehension despite reduced variations, a future replication is needed.

Third, inhibitory control and comprehension monitoring in the present study were observed variables. Although their reliability estimates were high, future
replications using multiple tasks and a latent variable approach are needed.

Fourth, all the listening comprehension measures in the present study were narrative in nature, and thus replications are necessary in expository text comprehension. Narrative and expository texts are different in terms of structure, language used in the texts, and the extent to which inferences are required (S.L. Gillam, Fargo, & Robertson, 2009). For instance, differences were found in what children noted after reading narrative versus expository texts. Children’s comments in the think-aloud data for narrative texts consisted of inferences (65–80%), whereas the majority of comments (68–74%) were paraphrases of expository texts (S.L. Gillam et al., 2009). Therefore, the nature of relations between cognitive skills and listening comprehension might differ for expository versus narrative texts.

Finally, the findings of the present study are limited to concurrent relations for kindergarten and first-grade children from high-poverty schools. Therefore, a study with children from diverse socioeconomic status backgrounds would be informative to examine generalizability. In addition, a longitudinal study would clarify directionality of relations. We examined a directional flow from ToM to listening comprehension in the present study. However, longitudinal and other studies suggest that the relation between ToM and language skill are likely to be bidirectional at multiple developmental levels (e.g., Astington & Jenkins, 1999; de Villiers, 2000, 2005; Grazzani & Ornaghi, 2012; Slade & Ruffman, 2005).

Given that the present study is correlational in nature, implications are preliminary and should be taken with caution. With this caution in mind, we believe that the findings of the present study, in conjunction with previous studies, preliminarily suggest that to improve listening comprehension of narrative texts for children from high-poverty schools, attention is needed beyond linguistic skills, particularly to multiple cognitive skills such as ToM, inhibitory control, and comprehension monitoring. In other words, there are multiple sources of difficulty with listening comprehension, including not only linguistic skills, such as vocabulary, but also cognitive skills.

Studies have indicated that children from high-poverty backgrounds have depressed oral language and cognitive skills (Arnold & Doctoroff, 2003; Arriaga, Fenson, Cronan, & Pethick, 1998; Lonigan, Burgess, Anthony, & Barker, 1998). Therefore, diagnostic assessment for difficulties with listening comprehension and instruction on oral language intervention should include cognitive skills such as inhibitory control, ToM, and comprehension monitoring. Note that the majority of previous oral language intervention has focused on linguistic skill, particularly vocabulary (e.g., Coyne, McCoach, & Kapp, 2007; Lesaux, Kieffer, Faller, & Kelley, 2010; Silverman, 2007). Emerging evidence suggests that ToM, comprehension monitoring, and inhibitory control are malleable (Bianco et al., 2010; Guajardo & Watson, 2002; Rueda, Rothard, Mc Candliss, Saccomanno, & Posner, 2005; Slaughter & Gopnick, 1996; Tsai, 2009). Given that listening comprehension is critical to reading comprehension and that difficulties with listening comprehension can be identified at an early age, whereas reading comprehension is not, it is crucial to identify children who are struggling with listening comprehension at an early age and provide intervention with an attention to cognitive skills. These noted instructional implications require future investigations.

NOTES
This work was supported by a grant from the Institute of Education Sciences, U.S. Department of Education (R305F100027). The views expressed herein are those of the authors and have not been reviewed or approved by the granting agencies.

1 Post hoc statistical power analysis (see Soper, 2013, for the formula used) showed that for a statistical power of .80 with two latent variables and 11 observed variables, minimum sample sizes to detect effect was 152 and 138 for detecting small effect sizes of .10 and .20, respectively.

REFERENCES

278 Reading Research Quarterly, 49(3)


Submitted October 2, 2013
Final revision received February 24, 2014
Accepted March 4, 2014

YOUNG-SUK KIM and BETH PHILLIPS are associate professors in the College of Education and the Florida Center for Reading Research at Florida State University, Tallahassee, USA; e-mail ykim@fcrr.org and bphillips@fcrr.org.