LSHSS

Research Note

# Do Live Versus Audio-Recorded Narrative Stimuli Influence Young Children's Narrative Comprehension and Retell Quality?

Young-Suk Grace Kim[a]

**Purpose:** The primary aim of the present study was to examine whether different ways of presenting narrative stimuli (i.e., live narrative stimuli versus audio-recorded narrative stimuli) influence children's performances on narrative comprehension and oral-retell quality.
**Method:** Children in kindergarten ($n$ = 54), second grade ($n$ = 74), and fourth grade ($n$ = 65) were matched on their performance on a standardized oral-language comprehension task and then were randomly assigned to 1 of the 2 conditions that differed in how narrative stimuli were presented to children: live narrative stimuli and audio-recorded narrative stimuli.
**Results:** Kindergartners and 2nd graders in the live condition had higher mean performance on narrative comprehension, with effect sizes of .43 and .39, respectively, after accounting for age, gender, and school. No differences were found in narrative comprehension for children in 4th grade. Children's oral-retell quality did not differ as a function of condition in any grade.
**Conclusion:** These results suggest that how narrative stimuli are presented to children (i.e., live versus audio-recorded narrative stimuli) may affect children's narrative comprehension, particularly for young children in kindergarten and Grade 2. Implications for assessment and instruction are discussed.

Narrative skills, including comprehension and production (and retell), are critical to daily interactional demands in the classroom (Snow, 1983, 1991; Westerveld & Gillon, 2010). Narrative comprehension, and retell and production,[1] are also strongly related to literacy skills such as reading comprehension (Dickinson & Porche, 2011; Hoover & Gough, 1990; Kendeou, van den Broek, White, & Lynch, 2009; Kim, 2015b; Tabors, Snow, & Dickinson, 2001) and writing (e.g., Berninger & Abbott, 2010; Juel, Griffith, & Gough, 1986; Kim, 2015a; Kim, Al Otaiba, Wanzek, & Gatlin, 2015). It is unsurprising that the Common Core State Standards (2010), implemented by many states in the United States, explicitly state expectations on children's narrative skills even in kindergarten. Moreover, recent efforts in narrative assessment have provided useful evaluation tools for children of various ages. Justice, Bowles, Pence, and Gosse (2010) developed the Narrative Assessment Protocol, which assesses the microstructure of children's narrative samples. Petersen, Gillam, and Gillam (2008) also developed a scoring system called the Index of Narrative Complexity, which evaluates the macro- and microstructure of narrative samples for progress-monitoring purposes.

Given the importance to academic success and increased attention to narrative skills in the school context, it is critical to have a clear understanding about procedural factors that influence children's performance in narrative comprehension and retell. Studies have shown that variation in procedures and methods does influence children's performance. For instance, variation in follow-up prompts influenced middle-school students' oral-retell performance such that students who were provided with follow-up prompts such as "Do you remember anything else?" produced a

[a]College of Education and Florida Center for Reading Research, Florida State University, Tallahassee
Correspondence to Young-Suk Grace Kim: ykim@fcrr.org

Editor: Marilyn Nippold
Associate Editor: Julie Masterson
Received May 15, 2015
Revision received August 20, 2015
Accepted November 6, 2015
DOI: 10.1044/2015_LSHSS-15-0027

[1]Narrative retell and production have been shown to capture a single dimension or construct (Gillam & Pearson, 2004; Kim, Park, & Park, 2015).

***Disclosure:*** *The author has declared that no competing interests existed at the time of publication.*

greater number of idea units than those who were not (Reed & Petscher, 2012). In addition, presentation of stories using oral and pictorial stimuli made a difference in children's story recall such that kindergartners and second graders recalled more content when stories were presented orally and pictorially than either orally or pictorially (Schneider & Dubé, 2005). A recent study investigated whether preschool children's oral narrative retell would differ when stories were presented by an animated movie compared to a traditional storybook. Results revealed that animated stories elicited a greater number of words and more diverse syntactic structure than still pictures (Diehm, Wood, Messier, & Callender, 2015). In contrast, color of illustrations in stories (color or black and white) does not appear to influence the elicited number of words or diversity of vocabulary (Schneider, Rivard, & Debreuil, 2011). It is interesting, though, that the author was not able to identify studies that examined the effect of varying assessment procedures and conditions on narrative comprehension.

Among several procedural factors relevant to narrative performance, the focal aspect in the present study was to vary how narrative stories or stimuli were presented to the child, whether the stimuli were read aloud by assessors or children listened to audio-recorded narrative stimuli. Assessments vary in this aspect. For instance, the assessment protocols in the Test of Narrative Language (TNL; Gillam & Pearson, 2004) and Renfrew Bus Story (Renfrew, 1969) require assessors to read aloud stories to children. On the other hand, in the Strong Narrative Assessment Procedure (Strong, 1998), children listen to audio-recorded stories corresponding to the Frog series of wordless picture books (Mayer, 1967; Mayer & Mayer, 1975). Research studies also vary in how narrative stories are presented to children. In some studies, stories were read aloud live by assessors (e.g., Curenton, Craig, & Flanigan, 2008; Dickinson & Tabors, 1991; Reese & Cox, 1999; Reese, Leyva, Sparks, & Grolnick, 2010), whereas in other studies, children heard audio- or video-recorded narrative stories (Scott & Windsor, 2000; Westerveld & Gillon, 2010; Westerveld, Gillon, & Miller, 2004).

Despite variability in research studies and standardized and normed instruments, it is not clear whether young children's comprehension and oral retell are influenced by how narrative stimuli are presented (live versus audio-recorded narrative stimuli). One motivation for using audio-recorded stories in assessment is related to measurement error. Live narrative stimuli are susceptible to measurement error attributable to assessors because assessors are likely to vary the way they read a story to different children. Despite training, some assessors may vary in terms of expressiveness (pitch and intonation). Furthermore, even the same assessor may read a story differently at different times (e.g., when the assessor is working with the first child of the day versus the 10th child). Using audio-recorded stories would greatly reduce these variabilities attributable to assessors, and thus reduce inconsistency in assessment, provided that the ambient environmental factors are controlled (e.g., quiet and undisturbed assessment location).

Listening to stories via audio media is not unfamiliar to young children. With the increasing availability and popularity of digital media and devices such as CDs, DVDs, computers, and television shows targeting young children (Rideout & Hamel, 2006), children are exposed to spoken language via digital devices at an early age (Krcmar & Cingel, 2014; Linebarger & Vaala, 2010). This is true in the classroom context also, because audio-recorded stories are frequently used for instructional purposes. For instance, it is not uncommon that primary-grade children listen to audio-recorded stories during center time for story comprehension and reading fluency. Furthermore, extant literature about children's vocabulary learning suggests that young children can learn from digital devices such as television (Anderson et al., 2001; Linebarger & Piotrowski, 2010; Linebarger & Walker, 2005; Rice, Huston, Truglio, & Wright, 1990; Schmitt & Anderson, 2002). Young children watching TV shows with rich vocabulary such as *Arthur* and *Dragon Tales* had a greater expressive vocabulary than those watching TV shows with reduced linguistic stimuli such as *Teletubbies*, after accounting for parental education, children's home environment, and general cognitive ability (Linebarger & Walker, 2005).

These findings are important because studies have consistently shown that vocabulary knowledge, a foundational oral language skill (Lepola, Lynch, Laakkonen, Silvén, & Niemi, 2012), is fundamental to discourse-level oral language skills, including narrative comprehension (Florit, Roch, & Levorato, 2013; Kendeou, Bohn-Gettler, White, & van den Broek, 2008; Kim, 2015b, 2016; Kim & Phillips, 2014; Lepola et al., 2012; Strasser & del Río, 2014). In other words, narrative comprehension requires an understanding of meanings of words in the stories in addition to other cognitive skills (e.g., working memory and inferencing). For instance, vocabulary was independently related to narrative comprehension for 4- to 5-year-old children (Florit et al., 2013). In a similar vein, vocabulary was independently related to narrative comprehension after accounting for theory of mind, comprehension monitoring, and inhibitory control for children in kindergarten and first grade (Kim & Phillips, 2014).

However, despite consistency in stimuli presentation, and evidence that children can learn vocabulary from digital devices, using audio-recorded narrative stimuli might not be the best way to present information in order to assess children's narrative skills accurately. Studies on language acquisition indicate that the most effective vocabulary learning (receptive vocabulary) is through live, child-directed or joint-reference contexts, not via media (Krcmar, Grela, & Lin, 2007). Moreover, the potential effect of live versus audio-recorded narrative stimuli might depend on the developmental level of children. Children younger than 22 months were not able to identify the target item when a novel word was presented via the television, whereas older children were (Krcmar et al., 2007). Therefore, the effect of live versus audio-recorded narrative stimuli, if any, may be larger for younger children than older children.

To the author's knowledge, little evidence is available about whether variation in the way a story is presented

(read aloud live by an assessor or audio-recorded) influences young children's performance on narrative comprehension and retell tasks. Therefore, the primary goal of the present study was to identify conditions under which children perform better, with a particular focus on how narrative stimuli are presented to children. The research question in the present study was as follows: Are there differences in mean performances in narrative comprehension and narrative oral-retell quality for children who listen to stories presented in a live narrative-stimuli condition versus an audio-recorded narrative-stimuli condition?

This research question was addressed by assigning children in kindergarten and Grades 2 and 4 to one of the two conditions. On the basis of the previous finding that live interaction is beneficial to children's language acquisition, it was hypothesized that live narrative stimuli might be advantageous for children's narrative comprehension and consequently story-retell quality. It was also expected that the effect, if any, would be more likely to be observed for younger children, but no hypothesis was made regarding a specific age or grade. It should be noted that in oral retell, retell quality and retell amount (i.e., total number of utterances) were examined. However, the author hypothesized that effect would be found in retell quality, not quantity, because any observed effects on retell would likely be due to the child's comprehension of the story, which would primarily influence retell quality.

The findings of the proposed study may have important implications for researchers and practitioners (e.g., teachers, school psychologists, and speech-language pathologists). If findings do show an impact of narrative-stimuli presentation method, researchers and clinicians should carefully consider how to present narrative stimuli to children during narrative comprehension and retell assessments. Furthermore, narrative-stimuli presentation method may have to be considered as part of the instructional procedures when designing instruction on narrative ability.

## Method

### *Overview of the Study Design*

Children in kindergarten, Grade 2, and Grade 4 were matched on their performance on a general language-comprehension task and then randomly assigned to either the live or the audio-recorded narrative-stimuli condition. Children in the live condition heard narrative stories read aloud by assessors. Children in the audio-recorded condition heard recorded narrative stories via a CD player. Children were then asked to retell the stories and were asked comprehension questions. Children's retells were digitally recorded, transcribed, and coded for retell quantity.

### *Participants*

A total of 193 children—54 in kindergarten (31 girls, 23 boys), 74 in Grade 2 (39 girls, 35 boys), and 65 in Grade 4 (37 girls, 28 boys)—from two elementary schools (one public and one private) in a midsize city in Florida participated in the study (see Table 1). Three additional children participated in the study but were not included in the analysis reported here because they were not available for the key assessments (narrative comprehension and retell) after assignment to conditions. Children with severe intellectual and emotional disabilities identified from school records were excluded from the study. Children with limited English proficiency, as determined by school records, were initially included. In Florida, where the study was conducted, English language learners' proficiency is assessed by the Comprehensive Language Learning Assessment, which includes listening, speaking, reading, and writing. A couple of students who had severely limited English proficiency (i.e., it was clear that these students did not understand directions in the task) were excluded from the study. Children with language impairment, speech impairment, learning disabilities, and limited English proficiency (those who had sufficient proficiency to understand and conduct the tasks) were included in the study for ecological validity. According to school records, one child in kindergarten was identified to have language impairment, one had speech impairment, one was identified with developmental delay, and one had limited English proficiency. In Grade 2, three children were identified to have language impairment, two had speech impairment, and one had limited English proficiency. In Grade 4, one child had language impairment. All these children were included in the study. Assignment to the conditions was based on the child's performance on a normed oral comprehension task in each grade, not the child's impairment status (see Table 1 for number of children in each condition).

As shown in Table 1, the majority of children were African American or Caucasian, which corresponds to the population of the area where the study was conducted. A total of 52%, 65%, and 51% of children, respectively, in kindergarten, Grade 2, and Grade 4 were eligible for free or reduced-price lunch. Note that the vast majority of children from the public school were African American, and all were eligible for free or reduced-price lunch, whereas the majority of children from the private school were Caucasian and none were eligible for free or reduced-price lunch. Differences in schools and associated demographic factors and socioeconomic status differences were taken into account in the statistical analysis by including school as a control variable. Given participating children's demographic characteristics in each school, including school as a control variable essentially accounted for differences in free and reduced-price lunch status as well as racial background of the children.

### *Measures*

#### Oral Language Comprehension

The Listening Comprehension subscale of the Oral and Written Language Scales–Second Edition (OWLS-II; Carrow-Woolfolk, 2011) was used for kindergarteners. In this task, the child heard sentences of increasing length and then was asked to point to a picture that represented

**Table 1.** Number of children and demographic information by condition.

| Grade | *n* | Mean (*SD*) age (years) | Girls/boys | Live condition | Audio-recorded condition | Racial backgrounds |
|---|---|---|---|---|---|---|
| Kindergarten | 54 | 5.64 (0.55) | 31/23 | 27 (1 language; 1 speech) | 27 (1 developmental delay; 1 LEP) | 50% African American; 43% Caucasian; 7% other |
| Grade 2 | 74 | 8.14 (0.64) | 39/35 | 36 (1 language; 2 speech) | 38 (2 language; 1 LEP) | 61% African American; 31% Caucasian; 8% other |
| Grade 4 | 65 | 9.97 (0.49) | 37/28 | 32 (1 language) | 33 | 49% African American; 42% Caucasian; 9% other |

*Note.* Language = language impairment; speech = speech impairment; LEP = limited English proficiency.

the heard sentence. Reliability was reported to be .98 for 3- to 7-year-old children (Carrow-Woolfolk, 2011). For children in Grades 2 and 4, the Woodcock-Johnson III Tests of Achievement (WJ-III) Oral Comprehension subtest (Woodcock, McGrew, & Mather, 2001) was used. The WJ-III Oral Comprehension subtest is a cloze task in which the child completes orally presented sentences (e.g., People sit in ____) and short stories. It has been shown to be related to other language skills such as verbal comprehension ($r$ = .59) and story recall ($r$ = .47) for children 6 to 8 years old (Woodcock et al., 2001). Reliability was reported to be .82 for 4- to 7-year-olds and .74 for 8- to 10-year-olds.

Using different listening-comprehension measures for kindergartners and children in Grades 2 and 4 is a limitation of the present study (see later). This was primarily due to the limited time that was available to the research team to work with children in Grades 2 and 4.

### Narrative Comprehension

Children's narrative comprehension was assessed by their responses to comprehension questions in Tasks 1, 3, and 5 of the TNL (Gillam & Pearson, 2004). In Task 1, the child heard a story without visual stimuli. In Task 3, a series of five pictures was presented to the child along with a story. In Task 5, a single picture was presented to the child along with a story. After hearing each story in either a live or an audio-recorded condition, the child was asked to retell the story and was then asked comprehension questions about the story. Comprehension questions were related to the narrative key elements and details (i.e., *Who were the main characters? Where did this story take place? What was the problem in this story?*). There were a total of 30 comprehension questions across the three tasks. Children's responses were scored according to the manual, with a total possible score of 39; although the vast majority of questions were worth 1 point, some were worth 2 or 3 points. Interrater agreement (exact agreement) was 1.00.

### Narrative-Retell Quality

Children's narrative retell was assessed by asking the child to retell the stories in TNL Tasks 1, 3, and 5. According to the TNL protocol, Tasks 3 and 5 do not have a retell portion, but these were adapted in the present study such that children were asked to retell the story immediately after they heard the stories in the three tasks. This was done to obtain retell data on multiple stories. Therefore, unlike with narrative comprehension, standard scores were not available for the retell portion of the TNL. Children were given the following directions for the retell: "Now I want you to tell the story back to me. Tell me everything that you remember. Try to say the story the same way as it was in the story." If the child paused for a few seconds, the assessor asked, "Is there anything else that you would like to add?"

Children's retells were recorded using a digital recorder (Olympus VN 8100 pc) and were transcribed verbatim following Systematic Analysis of Language Transcripts (SALT; Miller & Iglesias, 2008) guidelines. Transcriptions were segmented into C-units following those guidelines. Using transcribed data, children's narrative qualities were measured by the extent to which they included key narrative elements (e.g., main characters, setting, events, problem, and resolution) as well as key details of the stories in the narrative-retell tasks (e.g., Barnes, Kim, & Phillips, 2014). Children's performance on each narrative element was rated on a scale of 0 to 3, except the resolution element for Task 1, which was on a scale of 0 to 2. The child received a 0 for no inclusion of the story elements; 1 for a partially correct or implicitly stated retell; 2 for a correct but imprecise statement; and 3 for a precise statement. Children also received 1 point for each key detail that was a priori determined. The Appendix shows examples of a rubric for character delineation and problem identification for TNL Task 1 (i.e., McDonald's Story).

Graduate-student research assistants who were unaware of the study purpose and conditions were trained for coding on the basis of the rubric described. Interrater agreements (exact percent agreement) using 40 transcripts were .96, .94, and .95 for TNL Tasks 1, 3, and 5, respectively.

## *Procedure for Child Assessment and Assignment to Conditions*

Children were assessed in quiet places in the participating schools. The oral language comprehension tasks were administered first. On the basis of the children's performance on the oral comprehension task, a child was matched with another child from the same grade in the same school whose performance was highly similar in terms

of raw score and standard score (age based). Then, one of the two matched children was randomly assigned to the live narrative-stimuli condition and the other to the audio-recorded narrative-stimuli condition. Raw scores, in addition to standard scores, were used in matching children in order to ensure that children's oral language comprehension ability was comparable in an absolute sense, because standard scores incorporate children's age, and therefore children with similar standard scores have different oral language comprehension ability. The number of children in each condition by grade was as follows: 27, 36, and 32 children were in the live narrative-stimuli condition in kindergarten, second grade, and fourth grade, respectively; and 27, 38, and 33 were in the audio-recorded narrative-stimuli condition in kindergarten, second, and fourth grade, respectively (see Table 1). Note that the numbers of participating children in each condition in Grades 2 and 4 are not equal. This is because, as noted earlier, three children (two in Grade 2 and one in Grade 4) who were originally assigned to the live condition were excluded from analysis because they were not available for the narrative comprehension and retell assessments after assignment to the condition. Assessment procedures and locations were identical except for the live versus audio-recoded narrative-stimuli presentation. Volume in the audio-recorded stimuli condition was adjusted for comfort level for each child prior to the child's listening to the stories in each session. Headphones were not used in the audio-recorded condition for ecological validity, as is typically the case in many school settings.

For the audio-recorded stimuli condition, the stories of TNL Tasks 1, 3, and 5 were digitally recorded by a woman from the region where the study was conducted. The woman was asked to read the stories at "a slow, pleasant, oral reading pace" without "extra emphasis to the words that will be scored," according to the TNL manual (Gillam & Pearson, 2004, p. 13). Two graduate students who were unaware of the study's goal judged the recorded reading of the stories to be natural reading expected in a story reading. Children heard the recorded stories via a CD player (Emerson PD6548SL).

Graduate-student research assistants were rigorously trained on the administration procedures of assessments. Assessors in the live narrative-stimuli condition were asked to read the stories in the same manner just described, following the TNL manual. The same assessors administered all the tasks to children in the live and audio-recorded stimuli conditions in each grade. Prior to working with children, research assistants had to meet a minimum of 95% fidelity[2] in following assessment procedures (e.g., following directions accurately).

[2]Assessment fidelity was conducted by a master assessor who had extensive experience working with children and with the assessments included in the study. The master assessor had a checklist to observe for each assessment and observed accuracy of assessment administration when research assistants conducted mock assessments with another assessor.

## Results

### *Descriptive Statistics*

Descriptive statistics for the entire sample and for each condition (i.e., live and audio-recorded narrative stimuli) are shown in Table 2. In addition to standard scores, raw scores are reported because children's performance on the language comprehension tasks were matched using both standard scores and raw scores, and the multiple regression analysis was conducted using raw scores. Recall that the TNL comprehension raw scores are a composite of the three subtasks. TNL retell score was retell quality (see Appendix A for an example rubric).

Overall, the sample children's mean performances on the oral comprehension tasks (measured by the OWLS-II and WJ-III Oral Comprehension) were in the average range as indicated by mean standard scores ranging from 96.57 to 100.39. Their narrative-comprehension skill measured by the TNL was in the low average or average range (see Table 2). As expected, results of multivariate analysis of variance indicated no differences in the raw scores and standard scores in any of the grades in the oral comprehension tasks that were used in the condition assignment: Wilks's $\lambda = .995$, $F(51, 2) = .095$, $p = .91$ in kindergarten; Wilks's $\lambda = .998$, $F(70, 2) = .007$, $p = .93$ in Grade 2; Wilks's $\lambda = .973$, $F(62, 2) = .861$, $p = .43$ in Grade 4. Children's total number of utterances is also reported in Table 2. Although the number of utterances increased from kindergarten to Grade 4, no differences were found between live and audio-recorded narrative-stimuli conditions: $t(52) = 0.89$, $p = .38$ for kindergartners; $t(72) = 1.09$, $p = .28$ for second graders; $t(63) = 0.48$, $p = .63$ for fourth graders.

### *Differences in Narrative Comprehension and Retell as a Function of Condition*

In order to examine whether there are mean differences in children's performances on the narrative comprehension and retell quality as a function of the live versus audio-recorded narrative-stimuli conditions, multiple regression was conducted in each grade. Multiple regression is appropriate because it incorporates a data analytic strategy such as analysis of covariance (ANCOVA; Cohen, Cohen, West, & Aiken, 2003) but has fewer assumptions (e.g., homogeneity of variances) and more flexibility[3] than ANCOVA, and yields identical results (Tabachnick & Fidell, 2001). Demographic variables (age, gender, and school) were included as control variables in order to increase precision in estimation. Age was included as a covariate because raw scores were used in the analysis. Gender was included as a covariate to remove any potential differences in the outcome attributable to gender. As a final and

[3]"Any data analyzable by ANOVA/ANCOVA may be analyzed by multiple regression/correlation, whereas the reverse is not the case" (Cohen et al., 2003, p. 4). In fact, ANCOVA estimates are typically produced as part of the multiple regression outputs.

**Table 2.** Descriptive statistics.

| Grade | Measure | Entire sample | | Live condition | | Audio-recorded condition | |
|---|---|---|---|---|---|---|---|
| | | *M* (*SD*) | Range | *M* (*SD*) | Range | *M* (*SD*) | Range |
| Kindergarten | OWLS-II raw | 58.96 (15.04) | 15–90 | 59.67 (13.74) | 33–89 | 58.26 (16.48) | 15–90 |
| | OWLS-II SS | 100.39 (16.86) | 49–127 | 101.37 (15.94) | 62–126 | 99.41 (17.98) | 49–127 |
| | TNL comprehension raw | 19.43 (7.72) | 0–31 | 21.07 (6.96) | 4–31 | 17.78 (8.21) | 0–29 |
| | TNL comprehension SS | 8.61 (3.20) | 2–14 | 9.30 (3.15) | 3–14 | 7.93 (3.16) | 2–13 |
| | TNL retell total | 18.93 (10.75) | 0–45 | 20.67 (11.19) | 3–45 | 17.18 (10.20) | 0–41 |
| | Total number of utterances | 25.94 (11.20) | 7–56 | 27.30 (12.98) | 7–56 | 24.59 (9.11) | 10–42 |
| Grade 2 | WJ-III Oral Comprehension raw | 14.34 (4.59) | 1–24 | 14.17 (4.37) | 6–24 | 14.50 (4.84) | 1–24 |
| | WJ-III Oral Comprehension SS | 96.57 (17.47) | 37–132 | 95.11 (16.03) | 66–128 | 97.00 (18.93) | 37–132 |
| | TNL comprehension raw | 25.43 (6.58) | 1–35 | 26.72 (5.31) | 17–35 | 24.21 (7.46) | 1–34 |
| | TNL comprehension SS | 8.03 (3.57) | 1–15 | 8.53 (3.65) | 4–15 | 7.55 (3.48) | 1–14 |
| | TNL retell total | 31.11 (11.97) | 0–50 | 32.08 (12.48) | 0–50 | 30.18 (11.55) | 5–46 |
| | Total number of utterances | 34.89 (15.53) | 5–80 | 36.92 (17.88) | 5–80 | 32.97 (12.88) | 6–57 |
| Grade 4 | WJ-III Oral Comprehension raw | 18.32 (4.16) | 9–26 | 18.69 (4.13) | 10–26 | 17.97 (4.22) | 9–26 |
| | WJ-III Oral Comprehension SS | 96.71 (13.81) | 68–127 | 97.53 (13.68) | 70–124 | 95.91 (14.10) | 68–127 |
| | TNL comprehension raw | 30.37 (3.85) | 16–38 | 30.56 (3.95) | 16–38 | 30.18 (3.80) | 22–36 |
| | TNL comprehension SS | 9.65 (2.89) | 2–17 | 9.78 (2.82) | 2–17 | 9.52 (2.99) | 5–15 |
| | TNL retell total | 39.49 (6.76) | 24–51 | 39.38 (6.84) | 24–51 | 39.61 (6.79) | 28–51 |
| | Total number of utterances | 42.38 (11.20) | 17–67 | 43.06 (12.68) | 17–67 | 41.73 (9.69) | 21–59 |

*Note.* OWLS-II = Listening comprehension of Oral and Written Language Scales–Second Edition; SS = standard score; TNL = Test of Narrative Language; WJ-III = Woodcock-Johnson III Tests of Achievement.

important tactic, school was included as a covariate (dichotomously coded: private school = 1; public school = 0) to control for differences between the two schools, which differed in terms of children's socioeconomic and racial backgrounds. As noted earlier, controlling for schools essentially controlled for racial and socioeconomic backgrounds of the participating children.

The outcome variables in the multiple regression were TNL comprehension and retell quality; results are presented in Table 3. The primary predictor of interest was the live condition (see the boldface variable in Table 3). Because the live-condition predictor was a dichotomous variable (live condition = 1; audio-recorded condition = 0), the regression coefficients represent mean differences in the outcome between the live and audio-recorded narrative-stimuli conditions after accounting for the other variables in the model. For kindergarten and Grade 2, children in the live condition had statistically higher performance on narrative comprehension than those in the audio-recorded condition. On average, kindergarteners and second graders in the

**Table 3.** Multiple regression results: Narrative comprehension and retell outcomes predicted by live versus audio-recorded conditions after accounting for age, female gender, and school.

| Grade | Variable | TNL comprehension | | TNL retell quality | |
|---|---|---|---|---|---|
| | | β (*SE*) | *p* | β (*SE*) | *p* |
| Kindergarten | Intercept | 11.02 (14.93) | .46 | 26.73 (24.01) | .27 |
| | Age (months) | 0.02 (0.20) | .93 | −0.18 (0.32) | .58 |
| | Female gender | 3.27 (1.91) | .09 | 1.77 (3.06) | .57 |
| | School | 7.10 (1.85) | < .001 | 4.77 (2.98) | .12 |
| | **Live condition** | **3.68 (1.82)** | **.048** | **3.49 (2.92)** | **.24** |
| | Effect size | .43 | | .33 | |
| Grade 2 | Intercept | 36.26 (9.46) | < .001 | 51.29 (16.82) | .003 |
| | Age (months) | −0.14 (0.09) | .14 | −0.27 (0.17) | .12 |
| | Female gender | −1.14 (1.37) | .41 | 0.60 (2.43) | .80 |
| | School | 5.93 (1.49) | < .001 | 12.00 (2.64) | < .001 |
| | **Live condition** | **3.10 (1.36)** | **.03** | **2.43 (2.41)** | **.32** |
| | Effect size | .39 | | .16 | |
| Grade 4 | Intercept | 45.23 (10.62) | < .001 | 73.84 (19.66) | < .001 |
| | Age (months) | −0.14 (0.09) | .11 | −2.9 (0.16) | .07 |
| | Female gender | 0.09 (0.96) | .93 | −2.70 (1.78) | .14 |
| | School | 2.94 (0.94) | .003 | 4.20 (1.74) | .02 |
| | **Live condition** | **0.61 (0.88)** | **.49** | **0.61 (1.63)** | **.71** |
| | Effect size | .10 | | −.03 | |

*Note.* TNL = Test of Narrative Language; *SE* = standard error.

live narrative-stimuli condition scored 3.68 ($p = .048$) and 3.10 ($p = .025$) points higher on the TNL narrative comprehension outcome, respectively, than those in the audio-recorded narrative-stimuli condition after accounting for age, gender, and school. In contrast, no statistically significant difference was found in TNL narrative comprehension for children in Grade 4 ($p = .49$). Effect sizes for adjusted means[4] (see Lipsey & Wilson, 2000) are also reported in Table 3. When the outcome was oral-retell quality, no differences were observed as a function of live versus audio-recorded stimuli condition in any grade ($ps \geq .24$) after accounting for children's age, gender, and school.

## Discussion

Because narrative skill is critical to daily interactions as well as literacy acquisition, a precise understanding is needed about procedural factors that influence children's narrative performance. Previous studies have shown that assessment procedures such as the nature of follow-up prompts, story presentation method (i.e., oral and pictorial presentation), and use of media can influence children's performance on narrative tasks (Reed & Petscher, 2012; Schneider & Dubé, 2005). In particular, given the widespread and increasing use of media and digital devices in the home and classroom, it is important for researchers and clinicians to understand whether the presentation method of narrative stories (i.e., live versus audio-recorded stimuli) influences children's narrative comprehension and oral retell. Although audio-recorded narrative stimuli would promote consistency of presentation by reducing measurement error associated with assessors, little is known about whether it would influence children's narrative comprehension and retell quality.

Overall, the results of the present study show that narrative-stimuli presentation conditions do influence children's narrative comprehension, but not retell quality. To be specific, children in kindergarten and Grade 2 performed better in the live stimuli condition than in the audio-recorded stimuli condition. Effect sizes were .43 and .39, respectively, which are considered medium (Cohen, 1988). In contrast, no differences were found in narrative comprehension for children in Grade 4. These findings suggest that live reading of narrative stories does facilitate children's comprehension of the stories compared with audio-recorded stimuli presentation, at least for children in kindergarten and Grade 2. The results do not appear to be a by-product of the poor condition of audio recording, because if that were the case, a similar result would be expected in Grade 4. According to previous studies on children's language acquisition, although young children can learn language (e.g., vocabulary) via media, live child-directed context is more beneficial (Krcmar et al., 2007). Inferring from these studies, the present finding may be interpreted to suggest that children in kindergarten and second grade are likely to comprehend narrative stories better when stories are presented live by a person. These findings add to the growing literature about various procedural aspects of narrative assessment that influence children's performance on narrative tasks.

Unclear from the present study, however, is the mechanism because of which different performances were observed for the two different conditions. One potential factor might be attentional resources (Krcmar et al., 2007; Krcmar & Cingel, 2014). Attention has been shown to be directly and indirectly related to children's story comprehension (Kim, 2016; Strasser & del Río, 2014). Listening to live presentation of narrative stimuli may facilitate children's attentiveness to a greater extent than audio-recorded narrative stimuli because children are more accustomed to direct interactions. Despite the increasing popularity of digital technologies, the primary mode of interaction in the home and school is live and direct rather than via digital devices. Therefore, less familiarity with language experience via digital devices may demand greater mental processes and capacity, resulting in reduced cognitive resources (e.g., attention and working memory) left for comprehension (Bickham, Schmidt, & Huston, 2012). In support of this hypothesis, preschoolers with greater experience with television learned more from television than those with less experience (Krcmar, 2011). However, this speculation requires a future study.

Another remaining question is why differences were observed in narrative comprehension but not in retell quality. Although there was a favorable trend toward the live narrative-stimuli condition in retell quality, particularly for younger children (e.g., the effect size for kindergartners was .33), mean differences did not reach statistical significance (see Table 3). As a fundamental matter, children's oral retell should be based upon on memory and comprehension of the story, and therefore, differences in comprehension might translate to oral-retell quality. However, evidence suggests that comprehension and retell are not the same and instead may be two dissociable skills (e.g., Gillam & Pearson, 2004; Wagner, Sahlén, & Nettelbladt, 1999; see also Mar, 2004). Furthermore, children who have average oral language comprehension skills have impairment in expressive language including retell (Cornish & Munir, 1998; Rescorla & Ratner, 1996). Taken together, oral retell of narrative stories (e.g., the ability to retrieve and organize many elements of story) may draw on somewhat different skills than does narrative comprehension. An understanding of the different mechanisms involved in narrative comprehension and oral retell could explain the present findings of no effect of the conditions on oral retell quality.

### *Limitations, Future Directions, and Implications*

The results of the present study should be interpreted in the context of the study design and associated limitations.

[4]
$$d = \frac{\overline{X}_1 - \overline{X}_2}{s_{pooled}}$$
$$s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}.$$

One of the limitations includes different listening-comprehension tasks used between kindergarten and the two older grades. As noted earlier, this was primarily due to the practical constraints of conducting the study with limited time to work with children in Grades 2 and 4. There were some differences between the two listening-comprehension tasks. The OWLS-II Listening Comprehension task was a receptive task, whereas the WJ-III Oral Comprehension subtest was a receptive and expressive task. In addition, reliability estimates of these two tasks differed. Although different results were found for children in Grade 2 versus Grade 4 despite using the same general listening-comprehension task, a future study is needed using the same general listening-comprehension task across grades and age.

It would be informative to replicate the present study with different populations, including younger children (i.e., prekindergarteners) and somewhat older children (e.g., Grade 5), in order to extend our understanding about developmental differences in the effect of narrative-stimuli presentation on narrative comprehension. Furthermore, it could be explored whether the present findings are replicated for children with language impairment. It is important to stress that mechanisms of performance difference should be investigated to explain why live narrative stimuli versus audio-recorded narrative stimuli make a difference in comprehension.

The findings of the present study show that children's narrative comprehension was influenced by modes of narrative-stimuli presentation (i.e., live versus audio-recorded presentation). These results have important implications for assessments and instruction of children's narrative comprehension. For instance, speech-language pathologists who assess young children such as those in kindergarten and Grade 2 might consider live presentation of narrative stimuli, although audio-recorded stories are recommended (e.g., Strong, 1998). Furthermore, in developing or delivering instruction on children's narrative ability, how narrative stories are presented to children might have to be considered. According to the present findings, the teacher might not want to rely solely on children's listening to audio-recorded narrative stories in center activities. Instead, the teacher can supplement and combine it with live reading of narrative stories.

## Acknowledgments

This work was supported by Institute of Education Sciences, U.S. Department of Education Grant R305A130131. The views expressed herein are those of the author and have not been reviewed or approved by the granting agencies. The author wishes to thank the participating schools and children.

## References

**Anderson, D. R., Huston, A.C., Schmitt, K. L., Linebarger, D. L., Wright, J. C., & Larson, R.** (2001). Early childhood television viewing and adolescent behavior: The recontact study. *Monographs of the Society for Research in Child Development, 66*(1, Serial No. 264), 1–154.

**Barnes, A. E., Kim, Y.-S., & Phillips, B. M.** (2014). The relations of proper character introduction to narrative quality and listening comprehension for young children from high poverty schools. *Reading and Writing, 27,* 1189–1205. doi:10.1007/s11145-013-9481-0

**Berninger, V. W., & Abbott, R. D.** (2010). Listening comprehension, oral expression, reading comprehension, and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. *Journal of Educational Psychology, 102,* 635–651. doi:10.1037/a0019319

**Bickham, D. S., Schmidt, M. E., & Huston, A. C.** (2012). Attention, comprehension, and the educational influences of television and other electronic media. In D. G. Singer & J. L. Singer (Eds.), *Handbook of children and the media* (2nd ed., pp. 113–137). Los Angeles, CA: Sage.

**Carrow-Woolfolk, E.** (2011). *Oral and Written Language Scales–Second Edition*. Torrance, CA: Western Psychological Services.

**Cohen, J.** (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.

**Cohen, J., Cohen, P., West, S. G., & Aiken, L. S.** (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

**Common Core State Standards.** (2010). *National Governors Association and Council of Chief School Officers.* (2010). Retrieved from http://www.corestandards.org/

**Cornish, K. M., & Munir, F.** (1998). Receptive and expressive language skills in children with cri-du-chat syndrome. *Journal of Communication Disorders, 31,* 73–81.

**Curenton, S. M., Craig, M. J., & Flanigan, N.** (2008). Use of decontextualized talk across story contexts: How oral storytelling and emergent reading can scaffold children's development. *Early Education and Development, 19,* 161–187. doi:10.1080/10409280701839296

**Dickinson, D. K., & Porche, M. V.** (2011). Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. *Child Development, 82,* 870–886. doi:10.1111/j.1467-8624.2011.01576.x

**Dickinson, D. K., & Tabors, P. O.** (1991). Early literacy: Linkages between home, school and literacy achievement at age five. *Journal of Research in Childhood Education, 6,* 30–46.

**Diehm, E., Wood, C., Messier, J., & Callender, M.** (2015). *Technology enhanced narrative assessment: Effect of presentation modality on preschooler's narrative retell*. Manuscript submitted for publication.

**Florit, E., Roch, M., & Levorato, M. C.** (2013). The relationship between listening comprehension of text and sentences in preschoolers: Specific or mediated by lower and higher level components? *Applied Psycholinguistics, 34,* 395–415. doi:10.1017/S0142716411000749

**Gillam, R. B., & Pearson, N.** (2004). *Test of Narrative Language*. Austin, TX: Pro-Ed.

**Hoover, W. A., & Gough, P. B.** (1990). The simple view of reading. *Reading and Writing, 2,* 127–160.

**Juel, C., Griffith, P. L., & Gough, P. B.** (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78,* 243–255.

**Justice, L. M., Bowles, R., Pence, K., & Gosse, C.** (2010). A scalable tool for assessing children's language abilities within a narrative context: The NAP (Narrative Assessment Protocol). *Early Childhood Research Quarterly, 25,* 218–234. doi:10.1016/j.ecresq.2009.11.002

**Kendeou, P., Bohn-Gettler, C., White, M. J., & van den Broek, P.** (2008). Children's inference generation across different media.

*Journal of Research in Reading, 31,* 259–272. doi:10.1111/j.1467-9817.2008.00370.x

**Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S.** (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101,* 765–778. doi:10.1037/a0015956

**Kim, Y.-S. G.** (2015a). *Expanding the developmental models of writing: A direct and indirect effects model of development writing (DIEW)*. Manuscript submitted for publication.

**Kim, Y.-S.** (2015b). Language and cognitive predictors of text comprehension: Evidence from multivariate analysis. *Child Development, 86,* 128–144.

**Kim, Y.-S. G.** (2016). Direct and mediated effects of language and cognitive skills on comprehension or oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology, 141,* 101–120. doi:10.1016/j.jecp.2015.08.003

**Kim, Y.-S., Al Otaiba, S., Wanzek, J., & Gatlin, B.** (2015). Toward an understanding of dimensions, predictors, and the gender gap in written composition. *Journal of Educational Psychology, 107,* 79–95. doi:10.1037/a0037210

**Kim, Y.-S. G., Park, C., & Park, Y.** (2015). Dimensions of discourse-level oral language skills and their relations to reading comprehension and written composition: An exploratory study. *Reading and Writing: An Interdisciplinary Journal, 28,* 633–654.

**Kim, Y.-S., & Phillips, B.** (2014). Cognitive correlates of listening comprehension. *Reading Research Quarterly, 49,* 269–281.

**Krcmar, M.** (2011). Can past experience with television help US infants learn from it? *Journal of Children and Media, 5,* 235–247. doi:10.1080/17482798.2011.584373

**Krcmar, M., & Cingel, D. P.** (2014). Parent–child joint reading in traditional and electronic formats. *Media Psychology, 17,* 262–281. doi:10.1080/15213269.2013.840243

**Krcmar, M., Grela, B., & Lin, K.** (2007). Can toddlers learn vocabulary from television? An experimental approach. *Media Psychology, 10,* 41–63. doi:10.1080/15213260701300931

**Lepola, J., Lynch, J., Laakkonen, E., Silvén, M., & Niemi, P.** (2012). The role of inference making and other language skills in the development of narrative listening comprehension in 4–6-year-old children. *Reading Research Quarterly, 47,* 259–282. doi:10.1002/rrq.020

**Linebarger, D. L., & Piotrowski, J. T.** (2010). Structure and strategies in children's educational television: The roles of program type and learning strategies in children's learning. *Child Development, 81,* 1582–1597. doi:10.1111/j.1467-8624.2010.01493.x

**Linebarger, D. L., & Vaala, S. E.** (2010). Screen media and language development in infants and toddlers: An ecological perspective. *Developmental Review, 30,* 176–202. doi:10.1016/j.dr.2010.03.006

**Linebarger, D. L., & Walker, D.** (2005). Infants' and toddlers' television viewing and language outcomes. *American Behavioral Scientist, 48,* 624–645. doi:10.1177/0002764204271505

**Lipsey, M. W., & Wilson, D. B.** (2000). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

**Mar, R. A.** (2004). The neuropsychology of narrative: Story comprehension, story production and their interrelation. *Neuropsychologia, 42,* 1414–1434. doi:10.1016/j.neuropsychologia.2003.12.016

**Mayer, M.** (1967). *A boy, a dog, and a frog*. New York, NY: Penguin.

**Mayer, M., & Mayer, M.** (1975). *One frog too many*. New York, NY: Penguin.

**Miller, J. F., & Iglesias, A.** (2008). *Systematic analysis of language transcripts* (SALT). Madison, WI: University of Wisconsin.

**Petersen, D. B., Gillam, S. L., & Gillam, R. B.** (2008). Emerging procedures in narrative assessment: The Index of Narrative Complexity. *Topics in Language Disorders, 28,* 115–130.

**Reed, D. K., & Petscher, Y.** (2012). The influence of testing prompt and condition on middle school students' retell performance. *Reading Psychology, 33,* 562–585. doi:10.1080/02702711.2011.557333

**Reese, E., & Cox, A.** (1999). Quality of adult book reading affects children's emergent literacy skills. *Developmental Psychology, 35,* 20–28.

**Reese, E., Leyva, D., Sparks, A., & Grolnick, W.** (2010). Maternal elaborative reminiscing increases low-income children's narrative skills relative to dialogic reading. *Early Education and Development, 21,* 318–342. doi:10.1080/10409289.2010.481552

**Renfrew, C. E.** (1969). *The bus story: A test of continuous speech*. Old Headington, United Kingdom: Oxford.

**Rescorla, L., & Ratner, N. B.** (1996). Phonetic profiles of toddlers with specific expressive language impairments (SLI-E). *Journal of Speech and Hearing Research, 39,* 153–165.

**Rice, M. L., Huston, A. C., Truglio, R., & Wright, J. C.** (1990). Words from "Sesame Street": Learning vocabulary while viewing. *Developmental Psychology, 26,* 421–428.

**Rideout, V. J., & Hamel, E.** (2006). *The media family: Electronic media in the lives of infants, toddlers, preschoolers and their parents*. Menlo Park, CA: Kaiser Family Foundation.

**Scott, C. M., & Windsor, J.** (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research, 43,* 324–339.

**Schmitt, K. L., & Anderson, D. R.** (2002). Television and reality: Toddlers' use of visual information from video to guide behavior. *Media Psychology, 4,* 51–76.

**Schneider, P., & Dubé, R. V.** (2005). Story presentation effects on children's retell content. *American Journal of Speech-Language Pathology, 14,* 52–60. doi:10.1044/1058-0360(2005/007)

**Schneider, P., Rivard, R., & Debreuil, B.** (2011). Does colour affect the quality or quantity of children's stories elicited by pictures? *Child Language Teaching & Therapy, 27,* 371–378. doi:10.1177/0265659011414278

**Snow, C.** (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review, 53,* 165–189.

**Snow, C. E.** (1991). The theoretical basis for relationships between language and literacy in development. *Journal of Research in Childhood Education, 6,* 5–10.

**Strasser, K., & del Río, F.** (2014). The role of comprehension monitoring, theory of mind, and vocabulary depth in predicting story comprehension and recall of kindergarten children. *Reading Research Quarterly, 49,* 169–187. doi:10.1002/rrq.68

**Strong, C. J.** (1998). *The Strong Narrative Assessment Procedure*. Eau Claire, WI: Thinking Publications.

**Tabachnick, B. G., & Fidell, L. S.** (2001). *Using multivariate statistics* (4th ed.). New York, NY: Allyn & Bacon.

**Tabors, P. O., Snow, C. E., & Dickinson, D. K.** (2001). Homes and schools together: Supporting language and literacy development. In D. K. Dickinson & P. O. Tabors (Eds.), *Beginning literacy with language: Young children learning at home and school* (pp. 313–334). Baltimore, MD: Brookes.

**Wagner, C. R., Sahlén, B., & Nettelbladt, U.** (1999). What's the story? Narration and comprehension in Swedish preschool

children with language impairment. *Child Language Teaching & Therapy, 15,* 113–137.

**Westerveld, M. F., & Gillon, G. T.** (2010). Profiling oral narrative ability in young school-aged children. *International Journal of Speech-Language Pathology, 12,* 178–189. doi:10.3109/17549500903194125

**Westerveld, M. F., Gillon, G. T., & Miller, J. F.** (2004). Spoken language samples of New Zealand children in conversation and narration. *Advances in Speech Language Pathology, 6,* 195–208. doi:10.1080/14417040400010140

**Woodcock, R. W., McGrew, K. S., & Mather, N.** (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside.

## Appendix

Oral-Retell Quality Coding Example for the Test of Narrative Language Task 1 (McDonald's Story)

| Score | Character delineation | Problem |
|---|---|---|
| 0 | No utterance related to the characters of the story; or mentioning unrelated characters or unspecified, generic pronoun, "they" or "he" or "she" | No utterance related to the problem of the story |
| 1 | Utterances about characters are partial and imprecise ("the girl" or "the boy" or "the kids," but their names are not mentioned) | Problem is implicitly alluded to but not specific enough (e.g., They did not have money) |
| 2 | Utterances about characters are partial; either Lisa or Raymond and the mother are named correctly | Problem is implicitly noted (e.g., Mom cannot find wallet/purse) |
| 3 | Both Lisa and Raymond must be correctly named as well as their mother | Explicitly mentions the key problem (e.g., Mom left her purse on the kitchen counter) |