

## Research Article

# Tier 2 Language Intervention for Diverse Preschoolers: An Early-Stage Randomized Control Group Study Following an Analysis of Response to Intervention

Trina D. Spencer,<sup>a</sup> Douglas B. Petersen,<sup>b</sup> and John L. Adams<sup>a</sup>

**Purpose:** The first purpose of this study was to explore the use of a whole class, test-teach-test, dynamic assessment of narratives for identifying participants. The second purpose was to examine the efficacy of a Tier 2 narrative language intervention for culturally and linguistically diverse preschoolers.

**Method:** A dynamic assessment was conducted with students from 3 Head Start classrooms. On the basis of the results of the dynamic assessment, 22 children were randomly assigned to treatment ( $n = 12$ ) and control ( $n = 10$ ) groups for intervention. Participants received a small-group (4:1), differentiated, narrative intervention for 15–20 min, twice a week, for 9 weeks. Interventionists used weekly

progress monitoring data to explicitly focus on individualized narrative and linguistic targets.

**Results:** The treatment group showed significant improvement over the control group on proximal and distal measures of narrative retells, with large effect sizes. Group differences on a measure of children's language in the context of personal stories were not statistically significant.

**Conclusions:** This early-stage study provides evidence that narrative language intervention is an effective approach to improving the language skills of preschoolers with diverse language needs. Furthermore, the evidence supports the use of dynamic assessment for reducing overidentification and identifying candidates for small-group language intervention.

The central role that language plays in academic performance cannot be understated, and the ability to understand and produce complex oral language, such as that found in narratives (stories), appears to be particularly important for academic success (Bishop & Edmundson, 1987; Catts, Adolf, & Ellis Weismer, 2006; Clarke, Snowling, Truelove, & Hulme, 2010; Collozzo, Gillam, Wood, Schnell, & Johnston, 2011; Dickinson & McCabe, 2001; Feagans & Short, 1984; Kaderavek & Sulzby, 2000; Larney, 2002; Nation, Clarke, Marshall, & Durand, 2004; Nation & Snowling, 1997, 1998a, 1998b; Scott & Windsor, 2000; Stothard & Hulme, 1992). Often, oral narratives use detailed, structurally complex language that is very similar to the written language students encounter

in school (Dickinson & McCabe, 2001; R. B. Gillam & Johnston, 1992; Greenhalgh & Strong, 2001; Griffin, Hemphill, Camp, & Wolf, 2004; Nippold & Taylor, 1995; Nippold, Ward-Lonergan, & Fanning, 2005; Strong & Shaver, 1991; Westby, 1985). In addition to the use and comprehension of complex, literate language, a clear understanding and proficient use of narrative story grammar are essential for the comprehension and production of narratives, which are frequently emphasized in education (Hudson & Shapiro, 1991). Children who have difficulty with narration tend to have academic difficulty, yet research has clearly indicated that the production and understanding of complex language and narrative structure can be improved with intervention (Clarke et al., 2010; De La Paz & Graham, 1997; S. L. Gillam, Gillam, & Reece, 2012; Graham & Harris, 1993; Hayward & Schneider, 2000; Petersen, 2011; Petersen, Brown, et al., 2014; Petersen, Gillam, Spencer, & Gillam, 2010; Spencer, Petersen, Slocum, & Allen, 2014; Spencer & Slocum, 2010). Young children who are accurately identified as having difficulty with narrative language can benefit from early, intensive language instruction, which may result in improved academic performance.

<sup>a</sup>Northern Arizona University, Flagstaff

<sup>b</sup>University of Wyoming, Laramie

Correspondence to Trina D. Spencer: trina.spencer@nau.edu

Editor: Krista Wilkinson

Associate Editor: RaMonda Horton

Received July 15, 2014

Revision received October 13, 2014

Accepted May 28, 2015

DOI: 10.1044/2015\_AJSLP-14-0101

**Disclosure:** Trina D. Spencer and Douglas B. Petersen developed the intervention examined in this study and may be entitled to royalties related to its sale.

## Response to Intervention

Students have been traditionally identified as needing supplemental language support in U.S. public schools by using a remedial model that requires documentation of a language impairment and poor academic performance (Individuals with Disabilities Education Act, 2004). This model was designed to identify and serve the limited number of students who have an academically relevant language impairment. Because language impairment is likely polygenic (Tomblin & Buckwalter, 1998), its prevalence in a population is expected to fall within the lower end of the normal curve (Spaulding, Plante, & Farinella, 2006). Thus, the percentage of children in a population who have language impairment—even when such a classification is culturally relativistic and arbitrary—would not be high (e.g., 7%; Tomblin et al., 1997).

Under the current model, although misclassifications occur more often than desired (Peña, Iglesias, & Lidz, 2001), many children with language impairment receive special services in the public school system. There are, however, a far greater number of students in the United States who have difficulty producing and understanding the oral and written language expected of them, even though they do not have a language impairment (Petersen & Spencer, 2014). There are many reasons for this, including limited English-language proficiency, dialectal differences, and cultural and environmental differences (including lower socioeconomic status [SES]).

Apart from English language learners, who can receive supplemental English language instruction through English-as-a-second-language services, there are few systems in place to identify and support the academic language of all students—and it shows. The National Assessment of Educational Progress (2013), which primarily measures language-dependent reading comprehension (Catts, 2009; Kamhi, 2007), indicates that nearly 80% of students who are from culturally and linguistically diverse populations in the United States read below grade level. *Cultural and linguistic diversity*, in this context, refers to ethnic minorities and home languages that are mismatched with the language or dialect typically used in schools and can be related to low SES. This extremely high percentage of the student population that is experiencing language-related academic difficulty cannot possibly have language impairment, yet they still need additional language support.

Recently, systems that integrate frequent assessments and multitiered interventions for word-level reading (e.g., decoding) and for math have been implemented to individualize the education process in order to support all students' learning, including those without disabilities. These systems, often referred to as *response to intervention* (RTI) or *multitiered systems of support* (MTSS), offer an alternative instructional approach that eliminates the traditionally dichotomous "general" and "special" education process (Ehren & Nelson, 2005; Justice, McGinty, Guo, & Moore, 2009). RTI/MTSS models focus on prevention, affording earlier identification of children who require additional

support and providing it quickly and intensely so that students have the opportunity to catch up with their peers. This differentiated approach stands in stark contrast to waiting until students qualify for special education to provide intensified instruction when remediation is substantially more challenging (Lyon et al., 2001).

Through RTI/MTSS, students who need additional help, for whatever reason, receive that help. Within this framework, the general student population receives evidence-based instruction (Tier 1), and each student's performance and growth are assessed relative to curriculum expectations. Students who demonstrate lower-than-expected performance are provided with more intense, smaller group interventions (Tier 2). These smaller group interventions should provide learners with the exposure required to improve and reenter the general curriculum. Students who continue to struggle are placed in the third (Tier 3) and most intensive tier of instruction (Fuchs, Compton, Fuchs, Bryant, & Davis, 2008). The preventative nature of RTI allows for evidence-based Tier 1 instruction and moderately intensive Tier 2 interventions to be deployed before referrals to special education are made, thereby increasing the validity of special education eligibility classifications (Fuchs, Fuchs, & Speece, 2002). Early detection of difficulties and provision of services are based directly on individual performance.

There are two challenges with RTI/MTSS that this research addresses. First, MTSS has not been widely applied to language. This oversight denies essential language instruction to a large percentage of the student population who, although operating with intact language learning abilities, struggle with understanding and producing the oral and written language expected of them in U.S. schools. Second, research is just barely beginning to emerge from early childhood settings (Greenwood et al., 2013), and there is good reason to avoid blindly applying findings from RTI/MTSS research on school-age children to preschoolers. The most obvious and most discussed challenge is that the population of preschools, which are mainly income based or needs based, is markedly different from the larger population in elementary schools (Ball & Trammell, 2011; Gettinger & Stoiber, 2008; Greenwood et al., 2013). Not every child in the United States attends preschool, and the children whose parents can afford private day care do not typically attend income-based preschools such as Head Start. Given that, compared with the general education elementary school curriculum, the early childhood curriculum explicitly targets language outcomes, early childhood settings may be fertile ground for the development and evolution of multitiered systems of language support (MTSLS).

## Identification in RTI Systems

Proper implementation of RTI/MTSS requires two essential components: (a) the accurate identification of students in need for more intensive instruction through regular progress monitoring and (b) the use of a multitiered system of interventions (Fuchs et al., 2008). Ideally, students are

accurately placed in the appropriate tier of service as soon as possible so that early intervention and preventative measures can be taken. A key feature of any RTI/MTSS framework is that decisions about tier placement are based on specific, predetermined assessment criteria referred to as *benchmarks*. The validity of benchmark measures for identifying present or potential academic difficulties in students who are culturally and linguistically diverse, including students from low-SES backgrounds or English language learners, is highly suspect, especially at the beginning of the school year, when limited information is available (Ball & Trammell, 2011; Barrera & Liu, 2010; Catts, Petersen, Schatschneider, Bridges, & Mendoza, 2009; Petersen, Allen, & Spencer, 2014). Most relevant to an RTI/MTSS context is the extent to which over- or underidentification takes place when there is limited assessment validity. Underidentification leaves children without the services they need to succeed (Fuchs et al., 2008), and overidentification leads to the exhaustion of resources, providing intensive services to children who could have been successful with less intensive instruction.

In some schools with diverse student populations, a majority of young students (e.g., preschoolers and kindergarteners) at the beginning of the school year perform far below the established benchmark criteria (Petersen, Brown, et al., 2014), some of whom may not actually require intensive services. Administering multiple assessments over time can minimize problems related to bias and overidentification. By waiting to make any decisions about increased intensity of services until a pattern emerges indicating limited growth (or response) over time, validity is strengthened. Although cumulative evidence is far superior to the results of static single-time assessment that has considerable confounds for culturally and linguistically diverse students, precious time can be lost. The drawback of using progress-monitoring data that require the accumulation of multiple data points over several weeks or months is the time lost in waiting to see which children fail. One key purpose of the RTI/MTSS process is to greatly improve the antiquated “wait to fail” discrepancy model (Gersten & Dimino, 2006; Lyon et al., 2001), yet the need to wait—even a few months—that exists in the current RTI/MTSS systems translates to delayed language services. Although with this model the wait period in some cases has decreased from a few years to several months, for a child with limited language skills this delay could be academically detrimental.

There is a potential solution to this problem, whereby both validity and earlier identification are possible, and that is through the use of *dynamic assessment* (Fuerstein & Fuerstein, 1991; Jenkins & O'Connor, 2002). Dynamic assessment is conceptually related to the RTI/MTSS process, yet it is in abbreviated form (Grigorenko, 2009). Using a test–teach–test format, dynamic assessment provides information on a student’s present performance as well as the student’s ability to learn something new. Dynamic assessment can take place over a relatively brief period, and the extent to which a child learns during the brief dynamic process maps onto his or her ability to progress

academically throughout the school year (Petersen, Allen, & Spencer, 2014). The student’s ability to learn, often referred to as *modifiability*, can be an indicator of the level of intensity of intervention likely needed for academic success. This information on modifiability, in conjunction with static, pretest, or posttest information, can provide important information on current performance and future instructional needs. This information is both sensitive to any current needs a student might have (e.g., improving English language proficiency) and the likelihood of that student progressing adequately when given Tier 1 general education instruction. Through dynamic assessment, students with language intervention needs can potentially be identified earlier and accurately placed in the appropriate tier of service as soon as possible.

### *Intervention in RTI Systems*

In addition to frequent and accurate progress monitoring, an RTI/MTSS system requires multiple tiers of instruction that can cater to individual student needs. Standard estimates for school-age tiered systems indicate that 80% of students will respond satisfactorily to general instruction (Tier 1). Approximately 15% of school-age students can be adequately served through small-group intervention and increased exposure to material (Tier 2), and the remaining 5% will require the most concentrated level of intervention, in which individuals are served with the greatest degree of intensity (Tier 3; Mellard, McKnight, & Jordan, 2010). This distribution is patterned off of expectations derived from the normal curve, with Tier 3 serving children with the greatest needs. Difficulty with the production and comprehension of complex, academically related language, however, could have a strongly positive distribution, especially for young culturally and linguistically diverse children. These children are less likely to understand and produce language that is commensurate with the language expectations of the formal academic setting. A pattern that is much different from the 80%–15%–5% estimates for school-age populations may exist in many preschool settings (Ball & Trammell, 2011) because only children from low-income households, or those who are at risk of academic failure, qualify for enrollment in needs-based or income-based preschools such as Head Start. Although most of the preschool children enrolled in needs-based or income-based preschools do not have a language impairment, there may be a substantial portion of preschoolers who could benefit from explicit narrative-based language instruction that exposes them to the oral and written academic language encountered in the primary and secondary grades. Children attending preschools are often at risk for future academic difficulty and as a group do not represent a normal population distribution. Therefore, we might expect a greater percentage of these children to need supplemental language intervention. The exact percentage of preschoolers who need more intensive language instruction will depend on the community the preschool serves, the specific eligibility requirements, and the measures used to allocate interventions.

Given the ambiguity of the Tier 2 designation in preschool programs, we conceptually refer to Tier 2 as services that represent the middle tier of a multitiered language intervention program.

Only a few programs that address narrative language have been intentionally developed as multitiered interventions. Justice et al. (2009) described an early literacy curriculum called *Read It Again Pre-K* that can be used for Tier 1 or Tier 2 literacy instruction. Alongside vocabulary, print knowledge, and phonological awareness, narrative language is explicitly taught through storybook reading, questioning, and drawing attention to major story events. In a feasibility study of *Read It Again Pre-K*, preschool teachers implemented the program for 30 weeks in a whole-class format. The findings revealed meaningful language improvements for the preschoolers who received the program, resulting in medium to large effect sizes (Justice et al., 2010).

Another early childhood literacy curriculum developed for use within preschool RTI systems is Exemplary Model of Early Reading Growth and Excellence (EMERGE), which has three distinct tiers of intervention. As group size decreases (whole group, small group, individual), there is an increase in intensity on sound awareness (rhyming, alliteration, segmenting, blending), oral language (vocabulary, expressive language, listening comprehension), alphabet knowledge (letter recognition), and print awareness—the so-called SOAP skills. With respect to language outcomes, a preliminary evaluation of EMERGE revealed that children in the treatment group scored markedly better than children in the control group on a story retelling task (Gettinger & Stoiber, 2008). Of note is that the EMERGE developers recommend small-group (Tier 2) intervention be allocated to 50% of the students. They have found that because students who qualify for needs-based preschools have additional risk factors, small-group intervention can be an effective approach to addressing widespread needs.

In needs- and income-based early childhood settings, it may not be feasible for speech-language pathologists (SLPs) to provide direct services to every child who could benefit from Tier 2 language intervention. Nonetheless, SLPs engaged in the promotion of child language have much to offer preschool RTI systems in terms of language assessment and instruction. The same principles of effective language intervention applied to children with disabilities can benefit students who need additional experience with more complex, academically related language, and SLPs' expertise in these areas can be put to good use through a collaborative, interprofessional model of program development, service, and consultation.

### **The Current Study**

The intervention featured in the current study is an example of a multitiered narrative language program that has been developed and investigated through interprofessional collaboration. Stemming from a number of interdisciplinary projects between SLPs and educators, *Story*

*Champs* (Spencer & Petersen, 2012b) was designed to help prevent reading comprehension problems by teaching critical, academically related language skills early and in a multitiered fashion. In addition to children with language impairment, children who are culturally and linguistically diverse or from families with a low SES can also benefit from the type of narrative language intervention that is traditionally delivered by SLPs; therefore, there is a need to extend SLP interventions into settings in which more children can benefit (Petersen & Spencer, 2014). *Story Champs* has an intensive focus on many aspects of oral language production and comprehension through storytelling, whereas *Read It Again Pre-K* and EMERGE have only narrative components.

The *Story Champs* curriculum has lessons that can be delivered in various arrangements, such as in large groups, in small groups, and individually. It contains 12 carefully constructed stories that revolve around childhood themes such as losing an item or getting hurt. In addition to attractive visual materials (e.g., icons and illustrations), core components of *Story Champs* include flexible but manualized explicit teaching procedures, immediate corrective feedback, and story games to increase active participation. Children receive repeated practice retelling modeled stories and producing their own stories with systematic scaffolding of visual material and supportive prompting from an instructor.

The first examination of *Story Champs* featured the small-group arrangement (Spencer & Slocum, 2010). Participants included five diverse preschoolers enrolled in Head Start who had limited language skills. The five participants were distributed among three small groups of four students according to a multiple-baseline-across-groups experimental design. The narrative intervention focused on producing story grammar in both narrative retell and personal narrative generation formats. The first and second authors served as the interventionists, and they did not differentiate targets for children's various language needs. Teaching procedures included explicit instruction strategies such as frequent opportunities to retell and tell stories, modeling and shaping, immediate corrective feedback, and scaffolded visual support (e.g., icons and pictures; Ukrainetz, 2006a). Storytelling opportunities were embedded in games to enhance active listening and motivation. Groups received intervention 4 days a week for 7–18 min. All five preschoolers retold longer and more complete stories following intervention, and improvements in personal stories were also documented. Spencer and Slocum (2010) noted the potential for differentiated intervention within *Story Champs* sessions because each participant presented with various oral language needs (e.g., vocabulary, story grammar, lengthening utterance, use of subordination).

In an examination of *Story Champs* with whole classes of preschoolers, Spencer, Petersen, et al. (2014) completed a quasi-experimental control group study with four Head Start classes. The personal themed stories, pictures, icons, and teaching procedures (e.g., modeling, corrective feedback) were the same as in the small-group arrangement; differences included an exclusive focus on



retelling, a peer-tutoring component in which every child retold a modeled story to a classmate, and choral responding to produce parts of the story. An educator delivered daily 15- to 20-min sessions for 3 weeks with consultation from an SLP. Each participant's narrative language was assessed preintervention, postintervention, and at follow-up using retell, personal generation, and story comprehension measures. The results indicated that the treatment group's retell and story comprehension scores were statistically significantly higher than the control groups at postintervention and follow-up, but the intervention had a minimal impact on children's personal generation skills.

A unique aspect of this study was the completion of a responsiveness analysis. On the basis of the notion that diverse children may need differentiated intervention, children were sorted into post hoc groups according to how they responded to the intervention. Despite statistically significant results with large effect sizes, Spencer, Petersen, et al. (2014) found that only 64% of the preschoolers responded to the intervention as expected and that the remaining students would have likely needed a more intense teaching arrangement (i.e., Tier 2 or Tier 3). This pattern of response was consistent for the subset of treatment participants who were English language learners (59%), suggesting that they were not the only children who might benefit from a more intense intervention. Spencer et al. suggested that future research should attempt a similar responsiveness analysis for identifying Tier 2 candidates, drawing parallels to a dynamic assessment process.

Because culturally and linguistically diverse preschoolers, including those from low-income households, may need explicit language instruction in order to meet the language demands of the academic setting, classroom-based language interventions that SLPs and educators deliver cooperatively may lead to improved language outcomes and enhanced feasibility of tiered service models. Narrative interventions have been used traditionally by SLPs for treating children with language impairment (Petersen, 2011), but research indicates that narratives can be integrated into teacher-delivered, tiered literacy interventions (Gettinger & Stoiber, 2008; Justice et al., 2010), or they can be the foundation of teacher-delivered, tiered language intervention (Spencer, Petersen, et al., 2014; Spencer & Slocum, 2010). With respect to RTI/MTSS for language in early childhood settings, the majority of which serve children from low-income communities, little is known about how best to identify children for Tier 2 intervention, how Tier 1 and Tier 2 should be organized, or whether sufficient gains can be made when interventionists are not certified language professionals. Furthermore, and perhaps most urgent, there are not enough manualized language programs that have been validated through high-quality research designs to promote evidence-based tiered language interventions (Ukrainetz, 2006b).

Early childhood RTI/MTSS researchers need to identify a valid process for identifying children who need Tier 2 interventions. In the current study, we explored the use of a dynamic assessment framework to help identify participants for intervention. Thus, the first research question was "What

percentage of preschoolers attending Head Start are identified as needing Tier 2 language intervention using a dynamic assessment approach?" Considering the importance of the issues raised for SLPs who practice in early childhood settings, there is a need to document the efficacy of Tier 2 language interventions using stronger research designs. Therefore, in the current study we used an early-stage randomized control group design to investigate the efficacy of *Story Champs* small-group language intervention. The second research question was "Compared with a control group, do children who receive Tier 2 language intervention show improved narrative language as measured through narrative retells and personal generations?"

## Method

To answer both research questions, we conducted this study in two phases. First, we used a large-group dynamic assessment procedure to identify candidates for small-group intervention. Spencer, Petersen, et al. (2014) completed a responsiveness analysis at the completion of their study to determine which children had not responded to their low-intensity large-group intervention. On the basis of children's differentiated responses to the interventions, they made predictions about which children would likely benefit from more intense intervention (Tier 2). In the current study, we used the same responsiveness analysis strategy to identify appropriate participants for Tier 2 intervention. We considered this a test-teach-test dynamic assessment, with individualized pretest and posttest assessments and whole-class instruction. In the second phase of this study, the children identified for Tier 2 intervention through dynamic assessment were randomly assigned to treatment or control groups. Children in the treatment group received approximately 18 oral narrative language-intervention sessions, delivered in groups of four children, whereas children in the control group did not receive language intervention other than what had been provided through their Head Start classroom.

### Phase I

#### Participants and Settings

Three Head Start classrooms of preschool students were selected as a convenience sample to participate in this study. Two of the classrooms were traditional Head Start classrooms with a maximum of 20 students, one teacher, and one teaching assistant. Approximately 10% of the students in these classrooms received special education services through the local school district. The third classroom was a hybrid Head Start and special education class. Children enrolled in this classroom had to meet one of two criteria: (a) qualify for Head Start due to low income and have parent(s) who attended adult English language classes provided by the school or (b) have a current individualized education program (IEP). There were 16 students in this class, with one special education teacher, one Head Start teacher, and three teaching assistants. Of these students, 25% ( $n = 4$ ) had identified developmental disabilities with IEPs. Children in

all three classrooms attended school for 4 hr a day, 4 days a week. Two students from these three classrooms did not return parent permission forms and 12 students did not attend school during the entire dynamic assessment phase. As a result, Phase I participants across the three classrooms were limited to those who had signed permission forms and completed the dynamic assessment process ( $N = 41$ ).

This study took place in a southwestern state where there are many Native American and Mexican American families. All classroom teachers followed the Creative Curriculum for Early Childhood (Dodge, Colker, & Heroman, 2002) as their Tier 1 instruction. Although storytelling was not targeted explicitly, one of the early childhood learning objectives states, "retell a story with the beginning, middle, and end." To help describe participants, parents completed a family survey that asked them to report on child characteristics such as ethnicity and dominant language. Surveys were provided in English or Spanish depending on the parent's preferred language. A bilingual SLP was available to read the questions in English or Spanish to any parents who needed additional help. On the basis of the family survey results, the Phase I participant group was 48.9% Latino, 12.2% Native American, 12.2% White, 2.4% African American, and 24.3% who reported more than one ethnicity. English was the dominant language reported for 49% of the group, Spanish was reported as the dominant language for 17%, and 34% of the children's parents reported that they were equally dominant in Spanish and English. All assessment and intervention procedures took place in the Head Start classrooms or areas near the classrooms.

### Dynamic Assessment Procedures

A test-teach-test format was used to sort children into tiers based on their language learning abilities. The Narrative Language Measure (NLM; Petersen & Spencer, 2012) was used to assess children's language skills before (pretest) and after (posttest) 3 days of whole-class narrative instruction (teach). The NLM is a general outcome measure (Deno, 2003) with 25 parallel forms for each grade level (preschool to third grade). It is designed to assess children's narrative language growth. It involves standardized administration and scoring procedures. The NLM preschool forms (Spencer & Petersen, 2012a) have adequate alternate-form reliability ( $r = .85, p < .0001$ ) and strong evidence of concurrent validity ( $r = .88-.93$ ; Petersen & Spencer, 2012). To administer the test, research assistants (one undergraduate student, one graduate student, and one bilingual SLP) followed the script, read a model story in English, asked the child to retell it in English, and listened to the child's story while providing only neutral prompts. Pictures were not used in the elicitation of the narrative retells. Children's stories were recorded using digital audio recorders and scored immediately following administration. Three story retell opportunities were given to each child in the pretest and posttest sessions. Using the NLM, retells were scored for the clarity and completeness of story grammar elements (character, setting, problem, feeling, action, consequence, and ending) on a 0–2 scale with weighted points

for episodic elements (e.g., problem, action, consequence). Language complexity features such as the use of temporal coordinating conjunctions (*then*), causal subordinating conjunctions (*because*), and temporal subordinating conjunctions (*after*, *when*) were scored for their frequency. Total NLM retell scores were calculated by summing the story grammar, language complexity, and episodic points. The time required for individual administration of three stories was approximately 3–5 min, and scoring took another 4–5 min for each participant.

All children were administered the three pretest NLM retell stories on a Thursday. On the subsequent Monday, Tuesday, and Wednesday, large-group narrative intervention following the same procedures in Spencer, Petersen, et al. (2014) was delivered to each of the classes. This could be considered Tier 1 instruction on narratives, although it was not a typical part of the Creative Curriculum. Three sessions were selected for the teaching phase on the basis of preliminary findings that preschoolers show notable responses to story grammar interventions within three sessions (Spencer, Petersen, et al., 2014; Spencer & Slocum, 2010). Each large-group intervention session lasted approximately 15–20 min. The first author (English-speaking teacher) served as the interventionist. Each session followed the same steps as described in Spencer, Petersen, et al. (2014) and can be viewed online (<http://www.youtube.com/watch?v=0M-IKtJVg7s>). With pictures displayed so the whole class could see, the interventionist modeled a story while pointing to corresponding pictures and attaching brightly colored story grammar icons to the pictures. She had the children name each of the parts of the story (e.g., character, problem, feeling, action, ending) and then retold the story while children produced gestures representing each part of the story. Next, the interventionist called for individual turns, in which children answered questions about parts of the story (e.g., "Who was this story about?" and "What did he do to fix his problem?"). Once a child answered the question, the whole class repeated the answer using group responding. Finally, children were paired up to tell the story in its entirety to a peer (i.e., peer tutoring). Partners helped monitor, and when one partner finished telling the story, the roles were reversed. On the next day (Thursday), all children were administered three NLM retell posttest stories.

### Phase I Analysis

Children's highest pretest and posttest retell scores on the preschool NLM were analyzed for responsiveness following the procedures outlined by Spencer, Petersen, et al. (2014). Children who scored at or above a total score of 8 at pretest were classified as *levelers*. Children who scored below 8 at pretest and above 8 at posttest were classified as *responders*. Children who made gains from pretest to posttest but never earned a score of 8 or higher were classified as *gainers*. Children who made no gains from pretest to posttest and scored below 8 were classified as *minimal responders*. A retell score of 8 or above was selected as the cutoff because it translates into a basic episode consisting of a problem, action, and consequence. Five children were unable to

be tested using the NLM because of extremely limited verbal skills and attention associated with developmental disabilities. These children had already been identified as needing special education services and had current IEPs. Because they were unable to participate in the testing or attend in the large-group context (teaching phase), they were automatically grouped with the minimal responders and designated for Tier 3 language instruction. It should be noted that a few children were unable to produce scorable retells in English because of limited English language proficiency; however, these children were included in the responsiveness analysis because they were able to participate in the testing and attend during whole-class instruction. Note that we were interested in identifying children for a preventative small-group language intervention. Children with significant developmental disabilities who were unable to participate in the testing and who were already receiving individual speech-language services would not have been appropriate for a preventative intervention. In contrast, children who are typically developing but speak English as a second language are the exact population for which a Tier 2 small-group language intervention should be designed.

### Phase I Results

On the basis of this dynamic, test–teach–test classification process, we calculated the percentage of preschool students who were sorted into responsiveness groups. Seventeen percent of the children were levelers, meaning they scored above the cut score at pretest and did not need intervention (posttest  $M = 14.7$ ). Another 17% of the children were responders, suggesting that they did not need supplemental language intervention because they made sufficient gains from pretest to posttest and scored above the cut score at posttest ( $M = 11.0$ ). Twenty-four percent of the students were classified as gainers because they improved from pretest to posttest, but they did not perform above the cut score at posttest ( $M = 4.5$ ). The remaining children (42%) were classified as minimal responders ( $M = 0.5$ ), which included the five children with minimal verbal skills. It is important to note that not all children who had identified disabilities were automatically categorized as minimal responders. Two children with developmental disabilities were among the levelers and responders. All seven children who spoke Spanish as their dominant language were among the gainers or minimal responders.

The responsiveness analysis that was based on the dynamic assessment results made it possible to estimate how each child would respond to long-term language intervention, leading to tier assignment and selection of Tier 2 participants for Phase II of the current study. Because the five children with minimal verbal skills were unable to participate in the testing, they were automatically sorted into Tier 3 language supports (12%), which they received through their IEPs. They were not considered viable candidates for preventative Tier 2 language intervention. Levelers and responders (34%) did not demonstrate a need for narrative language intervention and were allocated to Tier 1, which they received through their classroom

instruction. Gainers and the minimal responders (54%) who were able to participate in the testing were considered appropriate candidates for Tier 2 language intervention and advanced to Phase II of the study (see Figure 1).

### Phase II

#### Participants

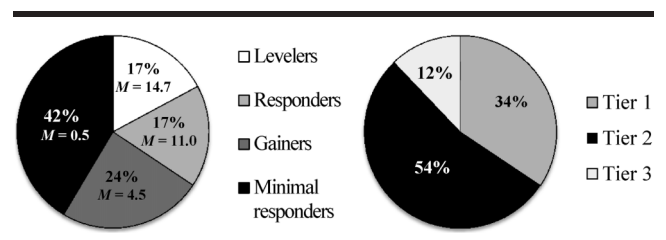
Twenty-two children were identified as candidates for Tier 2 language support. Because intervention was delivered within each classroom in a small group of four students, we randomly assigned 12 children, four from each classroom, to receive the intervention and 10 children to serve as the control group. After 2 weeks of intervention, one of the treatment children chose not to participate in the intervention and engaged in problem behavior when he was invited to participate. Based on our internal review board policy, this constituted declination of assent. As a result, he was removed from the study, leaving 11 children in the treatment group. All of the children in the treatment and control groups were administered the Clinical Evaluations of Language Fundamentals–Preschool (CELF-P; Wiig, Secord, & Semel, 2004) in English to document their overall language and English language skills prior to intervention. Participant characteristics, by group, are shown in Table 1.

#### Research Design and Measures

A pretest/posttest randomized control group design was used to answer the second research question. Children in both the treatment and control groups were assessed before and after the 9-week intervention condition as well as after a 4-week maintenance period. The NLM retell was considered the primary outcome measure because it most closely aligned with the focus of the intervention. Rather than readministering the preschool NLM to each of the Phase II participants, we used the children's dynamic assessment posttest scores from Phase I as their preintervention scores. All testing and intervention was completed in English. If children attempted to respond in Spanish, they were asked to do their best in English.

The Renfrew Bus Story (Cowley & Glasgow, 1994) was used as a secondary outcome measure. The Renfrew Bus Story is a preschool narrative retell measure involving a series of 12 pictures and an accompanying story about a bus that drives away when the driver tries to fix it. It is a

**Figure 1.** Percentage of participants who were sorted into each responsiveness category with group mean scores from the dynamic assessment posttest (left) and each intervention tier on the basis of the dynamic assessment results (right).



**Table 1.** Phase II participant characteristics

Characteristics	Treatment group ( <i>n</i> = 11)	Comparison group ( <i>n</i> = 10)
Age in months, <i>M</i> ( <i>SD</i> )	50.36 (8.54)	46.90 (6.69)
Gender, <i>n</i> (%)		
Male	6 (55)	4 (40)
Female	5 (45)	6 (60)
Dominant language ( <i>n</i> , %)		
English	6 (55)	6 (60)
Spanish	4 (36)	3 (30)
Bilingual	1 (9)	1 (10)
Ethnicity ( <i>n</i> , %)		
White	2 (18.2)	2 (20)
Latino	3 (27.3)	3 (30)
Native American	1 (9)	1 (10)
Multiethnic	5 (45.5)	4 (40)
CELF-P SS, <i>M</i> ( <i>SD</i> )	74.36 (15.39)	81.00 (26.20)
NLM preintervention, <i>M</i> ( <i>SD</i> )	2.27 (2.49)	2.80 (2.25)

*Note.* CELF-P = Clinical Evaluation of Language Fundamentals–Preschool; SS = standard score; NLM = Narrative Language Measure.

norm-referenced, standardized test with adequate test–retest reliability (.72–.79) and interrater reliability (.70–.92). Alternate forms of the Renfrew Bus Story are unavailable so it was administered only at preintervention and postintervention assessment points. Research assistants administered the tests and recorded children’s retells using digital voice recorders. Children’s stories were later transcribed and scored according to the standardized scoring procedures in the manual. Analyses were conducted for information and sentence length. The information analysis involves scoring for the amount of content the child includes in his or her retell that was also in the model story about a bus. For sentence length, each sentence is examined for word count, then a mean word count for the story is calculated.

In addition to eliciting a narrative retell using the NLM, personal stories were elicited. Research assistants played individually with the children in a quiet room and talked informally with them while they played. During the play session, examiners told stories (similar to the preschool NLM model stories) in first person and asked, “Has something like that ever happened to you?” If the child told a personal story, the examiner listened and provided only neutral prompts. When it appeared as if the child was done telling the story, the examiner said, “Are you finished?” Examiners shared three personal stories in each session, giving each child three opportunities to tell his or her own personal story. All personal stories were recorded using digital voice recorders and transcribed before they were scored.

The personal story transcripts were scored using a flowchart scoring rubric that is available by contacting the authors (Petersen & Spencer, 2013). The flowchart featured yes–no questions about the inclusion and complexity of story grammar elements and linguistic features (e.g., temporal and causal subordination), descending in quality. For each component, scores between 0 and 3 were possible, depending

on where a yes was recorded. This range of scores was larger than the scale for the preschool NLM retell because without a model story to constrain the content in personal stories, there were more possible features available.

*Administration fidelity and scoring reliability.* Prior to the study, research assistants were trained in the administration and scoring of the tests. After they read the manuals, the first author conducted an hour-long training on the preschool NLM, the personal story elicitation procedures, and the Renfrew Bus Story. One research assistant was an undergraduate student in psychology, one was a master’s student in psychology, and the third was a bilingual SLP who worked at one of the schools. These research assistants administered all of the tests but scored only the NLM story retells. Before qualifying to work on the study, the research assistants demonstrated accurate administration of all the tests and 90% or higher scoring agreement with the first author on five of the NLM parallel forms. These three research assistants served as primary scorers for the NLM retells and the secondary scorers for stories for which they were not the primary scorer. A fourth research assistant, who was a doctoral student in school psychology, transcribed and scored the retells produced through the Renfrew Bus Story administrations. To learn to score the Renfrew Bus Story responses, he read the manual and co-scored with the first author for the first few responses. For the personal story scoring, the second author served as the primary scorer and the first author served as the reliability scorer. The first and second authors developed the NLM and the personal story elicitation approach and have extensive experience administering, scoring, and interpreting language tests.

Thirty percent of participants’ retell and personal narratives from all assessment times (preintervention, postintervention, and follow-up) were randomly selected to be



scored by an independent scorer. The research assistants listened to participants' audio recordings of the narratives that were initially scored by a different research assistant. For each test, there were 11 opportunities for agreement, and to agree, both scorers had to rate the item exactly the same. The following formula was used to calculate percentage agreement: number of agreements divided by agreements plus disagreements, multiplied by 100. The mean agreement was 96.4% (range = 64%–100%) for retells and 84.9% (range = 55%–100%) for personal stories.

From preintervention, postintervention, and follow-up assessment times, 30% of all of the retell narratives and personal narratives were randomly selected for a fidelity examination. An independent research assistant listened to each of the audio recordings and completed an eight-step administration checklist for each test. For each one, the percentage of steps completed correctly was calculated. The overall mean fidelity of administration was 96.5% (range = 88%–100%) for retells and 94.8% (range = 76%–100%) for the personal story elicitation approach.

## Tier 2 Narrative Intervention

Children in the treatment group received approximately 18 sessions (depending on absences) of small-group narrative intervention, twice a week for 9 weeks. Small groups comprised one interventionist and four children. For two of the groups, all four children were research participants. For the third group, because one child dissented, other children in the class who were not control participants filled in as the fourth child on a rotating basis. If a child was absent, other children in the group who were not control participants were invited to join the group. This was done to ensure that the group size remained the same for each session.

*Program and materials.* The Tier 2 intervention adhered to the small-group procedures of *Story Champs* (Spencer & Petersen, 2012b). The program includes 12 personal-themed stories with accompanying pictures (five for each story). Pictures were large enough to spread across a small table for all four children to see them. Additional visual materials included brightly colored story grammar icons, representing the major parts of the story. Story games were used to increase children's active engagement while they listened to their peer tell a story individually. Materials for story games included small wooden sticks with the icons on them, small cubes with the icons on them, and bingo cards with the icons on them. Story gestures were also used in a game format, but materials are not required to play.

*Intervention steps.* A six-step procedural sequence was followed in each session. Visual material was systematically withdrawn so children told the story initially with pictures and icons for support and by the end of the session told the story without pictures or icons. The steps also moved from the interventionist modeling the story (Step 1), the group retelling the story (Step 2), individuals retelling the story (Steps 3–4), and finally to individuals generating personal stories (Steps 5–6). These steps are outlined in the Appendix and described in detail in the *Story Champs* manual (Spencer & Petersen, 2012b) and in Spencer and Slocum (2010).

*Differentiation.* Spencer and Slocum (2010) did not differentiate the intervention for their participants; they targeted the same story grammar elements for each participant and did not prompt linguistic targets. In the current study, each child's preintervention NLM retell performances were examined for the absence of developmentally relevant language features. This was done so that children could receive an individualized language intervention embedded in a small-group intervention session. For example, in one of the three participating groups, all four children were Spanish-speaking English language learners, and each had different narrative and linguistic targets. The child with the most advanced English language worked on using complex linguistic features (e.g., temporal and causal subordination); another child worked on English verbs because when she retold stories in English, she always used Spanish verbs; a third child worked on increasing the number of story grammar elements (e.g., problem, attempt, consequence); and the fourth child was encouraged to increase the complexity of simple utterances using more prepositions and subject–verb–object clauses. Each child's targets were addressed in the same 15-min session led by a single interventionist. Interventionists added developmentally appropriate targets as children progressed through the intervention phase and mastered earlier targets.

*Prompting and corrections.* The *Story Champs* manual provides guidance about how to verbally prompt children to produce their specific targets. This guidance is primarily principle based, meaning that principles (not specific behaviors) were followed to ensure that children received verbal prompting and corrections specific to their needs. The principles included the following:

- (a) Never prompt more than necessary, but prompt enough to help the child be successful. For this study, the interventionist judged whether an indirect prompt, such as “What happened next?” or a direct prompt such as “John was sad because he got hurt. Now you say that” was necessary to get the child to use the target. Overprompting leads to dependency on prompts and discourages independent storytelling.
- (b) Prompt or correct immediately. If a child did not use the word *because* when it was his or her target and there was an opportunity for its use, the interventionist stopped the child and prompted use of the target.
- (c) Ensure that each child had several opportunities to practice his or her language targets in the session. This was done through repetition of the stories and parts of the stories. For example, the interventionist might have said, “Great. You used the word *because*. Now start at the problem and tell me the parts of the story again.”
- (d) Always encourage and praise and never tell a child he or she did something wrong. During corrections, interventionists only modeled and suggested what the child should say. They never said, “No. You said it wrong.”

Motivation is a key part of storytelling and we wanted to ensure children were engaged, felt socially comfortable,

and were supported. This is especially important considering that these children had significantly low language (or English language) skills.

**Take-home activities.** After each intervention session, the interventionists sent a storytelling activity home with the children that corresponded with the story featured in the intervention session. The take-home activities included (a) a very brief introduction to the parts of the stories in Spanish or English and (b) a series of five pictures. The purpose of the take-home activities was to involve the families in a rich language and literacy activity without requiring them to read books to their children. At the end of the study, the researchers sent a brief survey home asking parents about their use of the take-home activities and the interactions they had with their children. The items were scored on a 1–5 Likert scale, where 1 represents *never/strongly disagree* and 5 represents *often/strongly agree*. Ten of the 11 parents returned the completed survey; five completed them in Spanish. The items and mean responses are as follows:

1. My child told stories at home using the take-home activities ( $M = 3.9$ ).
2. My child enjoyed the storytelling activities ( $M = 4.4$ ).
3. I enjoyed the storytelling interaction with my child ( $M = 4.3$ ).
4. My child's language improved as a result of the storytelling activities ( $M = 4.4$ ).

**Interventionists and fidelity of implementation.** The three research assistants previously described served as interventionists, one for each of the three groups. Before serving as interventionists, the research assistants read the *Story Champs* manual, practiced with nonparticipant children, and received coaching and feedback from the first author. Throughout the intervention phase, the first author observed a third of the sessions, ensuring that each interventionist was observed at least five times. She completed a 33-step fidelity checklist (see the Appendix) each time and used the results to give feedback to the interventionists following the session. The average fidelity of intervention implementation was 97.8%, with a range of 91% to 100%.

## Phase II Results

This study involved random assignment to one of two groups: control or treatment. Analysis of covariance (ANCOVA) was selected to minimize error variance and control for variance in pretest performance, thus providing a clearer picture of the effectiveness of the intervention than would otherwise be possible. In addition, ANCOVA provides more statistical power than alternative analyses such as a factorial analysis of variance (ANOVA), which was desirable given the small sample size. ANOVAs were initially conducted to determine whether there were significant differences between groups at pretest for each outcome measure at the alpha .05 level. Data were then screened to assess the appropriateness of conducting an ANCOVA. This was determined by testing the homogeneity of slopes and homogeneity

of variances assumptions. Homogeneity of slopes was investigated by examining the interaction between preintervention measures and the experimental group, and the homogeneity-of-variances assumption was examined using the results of Levene's test. If both of these assumptions were met, an ANCOVA was conducted using .05 as the level of significance. If one or both of these assumptions were violated, a  $t$  test was conducted. Where significant results were found, Cohen's  $d$  was used as a measure of effect size. Means and standard deviations for the NLM retells, Renfrew Bus Story (information and sentence length), and the personal stories across preintervention, postintervention, and follow-up assessment times are shown in Table 2.

### Narrative Language Measures: Preschool

**Pretest.** The NLM pretest retell performance was selected as a covariate for NLM analyses. The ANOVA results for the pretest NLM retell were nonsignificant,  $F(1, 19) = 0.257, p = .618$ , indicating that there was no statistical difference between the treatment and control groups at pretest.

**Posttest.** Homogeneity of slopes and homogeneity of variances for NLM retell data were examined to determine whether an ANCOVA was an appropriate analysis. The interaction between group (treatment and control) at pretest was nonsignificant,  $F(1, 17) = 0.188, p = .670$ . Levene's test indicated that the posttest error variances did not differ significantly between groups,  $F(1, 19) = 2.50, p = .130$ . Because the assumptions were met, an ANCOVA was conducted. The results of the ANCOVA were significant,  $F(1, 18) = 6.55$ , mean standard error ( $MSE$ ) = 36.95,  $p = .02$ ,  $d = 1.05$ , indicating that the mean NLM retell score of the treatment group was significantly higher than that of the control group, with a large effect size.

**Follow-up.** Homogeneity of slopes and homogeneity of variances for NLM retell data were examined to determine whether an ANCOVA was an appropriate analysis. The interaction between group and pretest was nonsignificant,  $F(1, 17) = 0.117, p = .736$ . Levene's test indicated that the follow-up test error variances did not differ significantly between groups,  $F(1, 19) = 2.82, p = .109$ . Because the assumptions were met, an ANCOVA was conducted. The results of the ANCOVA were significant,  $F(1, 18) = 4.79$ ,  $MSE = 25.22, p < .05, d = 0.86$ , indicating that the mean retell performance of the treatment group was significantly higher than that of the control group at follow-up, with a moderate to large effect size.

### The Renfrew Bus Story

**Information pretest.** Renfrew Bus Story information pretest performance was selected as the covariate for analyses of Renfrew Bus Story information scores at posttest and follow-up. The ANOVA results for the information analysis of the Renfrew Bus Story were nonsignificant,  $F(1, 19) = 0.491, p = .492$ , indicating that there was no statistical difference between the treatment and control groups at pretest.

**Information posttest.** The interaction between group and pretest information analyses of the Renfrew Bus Story was nonsignificant,  $F(1, 17) = 0.027, p = .872$ . Levene's test

**Table 2.** Means and standard deviations for treatment and control groups at preintervention, postintervention, and follow-up.

Measure	Treatment group ( <i>n</i> = 11)		Control group ( <i>n</i> = 10)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Narrative Language Measures (retell)				
Preintervention	2.27	2.49	2.80	2.25
Postintervention	13.18*	5.56	6.70	6.71
Follow-up	12.27*	4.45	7.80	5.85
The Renfrew Bus Story—Information				
Preintervention	7.36*	6.71	5.70	3.49
Postintervention	14.8*	7.15	8.50	4.79
The Renfrew Bus Story—Sentence length				
Preintervention	4.36	3.14	4.30	2.00
Postintervention	6.45	2.46	5.00	2.10
Personal stories				
Preintervention	7.45	6.53	7.30	5.81
Postintervention	15.27	5.27	11.3	8.14
Follow-up	14.42	7.68	11.0	9.42

\**p* < .05.

indicated that the posttest error variances did not differ significantly between groups,  $F(1, 19) = 0.675$ ,  $p = .422$ . Because the assumptions were met, an ANCOVA was conducted. The results of the ANCOVA were significant,  $F(1, 18) = 7.37$ ,  $MSE = 16.47$ ,  $p = .014$ ,  $d = 1.04$ , indicating that the mean information score of the Renfrew Bus Story for the treatment group was significantly higher than that of the control group, with a large effect size.

**Sentence length pretest.** Sentence length at pretest was selected as a covariate for analyses of Renfrew Bus Story sentence length at posttest and follow-up. ANOVA results for the sentence length analysis of the Renfrew Bus Story were nonsignificant,  $F(1, 19) = 0.003$ ,  $p = .957$ , indicating that there was no statistical difference between the treatment and control groups at pretest.

**Sentence length posttest.** The interaction between group and pretest sentence length analyses of the Renfrew Bus Story was nonsignificant,  $F(1, 17) = 1.13$ ,  $p = .303$ . Levene's test indicated that the posttest error variances did not differ significantly between groups,  $F(1, 19) = 0.561$ ,  $p = .463$ . Because the assumptions were met, an ANCOVA was conducted. The results of the ANCOVA were nonsignificant,  $F(1, 18) = 3.65$ ,  $p = .072$ , indicating that the posttest sentence length means did not differ significantly between groups.

### Personal Stories

**Pretest.** Personal story pretest performance was selected as the covariate for analyses of personal story performance at posttest and follow-up. The ANOVA results for the personal story scores were nonsignificant,  $F(1, 19) = 0.003$ ,  $p = .955$ , indicating that there was no statistical difference between the treatment and control groups at pretest.

**Posttest.** The interaction between group and pretest personal story scores was nonsignificant,  $F(1, 17) = 0.004$ ,  $p = .95$ . Levene's test indicated that the posttest error variances

were significantly different between groups,  $F(1, 19) = 12.22$ ,  $p = .002$ . Because the assumption of homogeneity of variances was violated, a *t* test was conducted. The results of the *t* test were nonsignificant ( $t = 1.34$ ,  $p = .196$ ), indicating that that posttest means for the personal story scores did not differ significantly between groups.

**Follow-up.** The Group  $\times$  Pretest Personal Story Scores interaction was nonsignificant,  $F(1, 17) = 0.254$ ,  $p = .627$ . Levene's test indicated that the posttest error variances were significantly different between groups,  $F(1, 19) = 9.19$ ,  $p = .007$ . Because the assumption of homogeneity of variances was violated, a *t* test was conducted. The results of the *t* test were nonsignificant ( $t = 1.02$ ,  $p = .321$ ), indicating that the follow-up means for the personal story scores did not differ significantly between groups.

### Social Validity

At the end of the study, the researchers asked the three Head Start teachers to answer questions about treatment acceptability and feasibility of *Story Champs*. Statements were rated using a Likert scale on which 1 represents *strongly disagree* and 5 represents *strongly agree*. The items and mean responses were as follows:

1. Storytelling is an appropriate context for learning and practicing language ( $M = 4.7$ ).
2. *Story Champs* activities are appropriate for preschoolers ( $M = 4.0$ ).
3. My students enjoyed *Story Champs* ( $M = 4.0$ ).
4. *Story Champs* procedures appear easy to learn ( $M = 3.7$ ).
5. My students' language improved as a result of *Story Champs* ( $M = 4.7$ ).
6. I am interested in using *Story Champs* in my classroom ( $M = 4.3$ ).

## Discussion

Although it is not always feasible for SLPs to deliver preventative language interventions within early childhood programs, their expertise qualifies them to contribute knowledge about effective language intervention approaches and to help establish language assessment procedures that drive multitiered intervention systems. An evidence base large enough to indicate clear intervention options can facilitate that task. We cannot assume that the identification of children for supplemental interventions works the same in preschools as it does in elementary schools. In fact, there is reason to believe that large portions of children attending Head Start and other low-income preschools would benefit from small-group interventions. The purpose of this study was to examine the feasibility of using a dynamic assessment approach to identify preschoolers who are in need of more intensive language intervention and to investigate whether the children who received the Tier 2 language intervention showed improved narrative language measured through narrative retells and personal stories over children in a control group.

### Phase I

Phase I represents an initial attempt at using a whole-group dynamic assessment approach to help accurately identify children in need of Tier 2 instruction and to reduce the wait time necessary for Tier 2 assignment. Children who are culturally and linguistically diverse are frequently overidentified as needing intensive instruction by static measures (Catts et al., 2009). Furthermore, using a traditional RTI/MTSS approach to identifying children in need of more intensive services can take considerable time. Because children attend preschool for only 1 or 2 years, waiting to establish a pattern of response to instruction using progress-monitoring results could take several months (or longer)—time preschoolers do not have.

Given the urgency of early intervention (Dickinson, McCabe, & Essex, 2006), researchers have been developing general outcome measures that assist in the allocation of tiered interventions as early as 1 or 2 months within children's start in preschool (e.g., individual growth and development indicators; Bradfield et al., 2014). Individual growth and development indicators are fluency-based measures similar to general outcome measures that are used in elementary schools, but they feature early childhood oral language skills such as picture naming. Although general outcome measures for preschoolers show promise for RTI/MTSS models, they are subject to the same bias as all other static tests—they tend to overidentify culturally and linguistically diverse children because they measure current performance, not responsiveness or modifiability (Barrera & Liu, 2010). Through the dynamic assessment process in the current study, we removed children who were able to respond to low-dose, low-intensity language instruction, suggesting they do not have difficulty learning language (responders = 17%). Had we used the

NLM retell as a static measure, the pretest scores would have indicated that 83% of our sample needed language intervention.

Assignment to individualized, tiered intervention services took less than a week to allocate using our dynamic assessment process, identifying a smaller proportion of culturally and linguistically diverse children in need of more intensive instruction than what is typically noted with static measures. This is the most notable finding from Phase I, in which we identified 54% of the initial sample from three Head Start preschool classrooms for Tier 2 language intervention. This included the children who made no progress from pretest to posttest or who made progress but did not score above the cut score of 8. Because our participants were primarily Latino, American Indian, and multiethnic from low-SES households, this figure is not surprising, and it supports findings from Gettinger and Stoiber (2008) and Ball and Trammell (2011). Many have suggested that RTI procedures for school-age children cannot be applied simplistically to early childhood settings (Barnett, VanDerHeyden, & Witt, 2007; Barrera & Lui, 2010; Greenwood et al., 2013), and the pervasiveness and diversity of language needs in Head Start preschools are compelling reasons to approach the identification process differently and respond as needed.

Using a story retelling task and a dynamic assessment process, we found that more than half the children in the current study qualified for Tier 2 instruction. The disproportion of Tier 2 candidates in our sample calls into question whether the small-group intervention provided should be considered a Tier 2 intervention. Perhaps more attention should be given to enhancing Tier 1 instruction (Greenwood et al., 2011) or arranging Tier 1 instruction to include explicit small-group language instruction (Gettinger & Stoiber, 2008; Justice et al., 2009).

Certainly, in a preschool context similar to the one we studied, an SLP's expertise is probably best used to train teachers and paraprofessionals in how to implement small-group language interventions within the classroom and to provide ongoing consultation. SLPs may also provide leadership in the conduct of a whole-class dynamic assessment, including delivering a large-group narrative intervention in classrooms. We recommend that SLPs strive for collaborative partnerships with teachers that result in coteaching and coassessing so that preventative MTSLS are viable.

Some additional interesting findings emerged from Phase I. First, there were two children with developmental disabilities sorted into the Tier 1 group. This is noteworthy because the presence of a developmental disability by itself did not indicate the need for supplemental language support. *Developmental disability* is a broad category often used in preschool settings that is ambiguous and can be unrelated to language problems. These research findings highlight that not all children who have developmental disabilities should be automatically assigned to Tier 2 or Tier 3 intervention services for language. Second, the responsiveness analysis revealed four different patterns of response,



and Latino children were distributed across all of them. Although all seven of the children whose parents reported Spanish as their dominant language qualified for Tier 2 intervention, there were many other Spanish-speaking children who fell within the responder category. This highlights the probability that language learning needs are more complex than a dichotomous risk/no risk classification, which is typically generated when general outcome measures are used in a static fashion. This finding also suggests that professionals cannot assume all English language learners are in need of intensive supplemental language intervention. Depending on the timing, sequence, and duration of English exposure, languages spoken at home, aptitude, and many other factors, children's language learning needs will vary. And whereas it would be difficult to measure all of the contributing factors leading to difficulty with academic language, we can measure each child's individual response to focused narrative language intervention and respond appropriately.

A third finding of interest is the basis on which we separated candidates for Tier 2 and Tier 3. We did not anticipate that some children would be unable to participate in the simple retell assessment procedures and therefore had not preplanned eliminating children from Tier 2 candidacy for that reason. However, it quickly became apparent that the five children who did not have the language or self-management skills to comply with testing or to participate in a group intervention without paraprofessional assistance were in need of something more intense than our Tier 2 instruction. This finding draws attention to the need to attend to more than just test results when assigning children to tiered interventions. For example, information on children's attention, cooperation, and the functional impact of their disability may need to be integrated into the decision-making process. On the basis of two recent *Story Champs* studies with children with autism (Petersen, Brown, et al., 2014) and severe developmental delays (Spencer, Kajian, Petersen, & Bilyk, 2014) as well as other, similar research (Petersen, Gillam, Spencer, & Gillam, 2010), we are confident that narrative intervention can be effectively used with children with more significant disabilities. However, all of these narrative studies featured an intensive, individual arrangement, which we concluded would be more appropriate for the five preschoolers we sorted into Tier 3. Note that we used our clinical judgment to arrive at this conclusion, but it may be important for future research to examine more objective mechanisms for differentiating between Tier 2 and Tier 3 language intervention candidates.

## Phase II

After using a dynamic assessment process to identify children in need of more intensive language instruction, we implemented a small-scale randomized control group design to investigate the efficacy of a Tier 2 language intervention with diverse preschoolers. We examined the effect of the small-group narrative intervention on a proximal retell measure (NLM) and a more distal retell measure (the

Renfrew Bus Story). Even though stories were unfamiliar to the children, the NLM is closely aligned with the *Story Champs* intervention. Because the Renfrew Bus Story is a norm-referenced test that does not closely align with the intervention, it was used as a secondary, distal retell measure. The extent to which the small-group narrative intervention was powerful enough to affect children's language in personal stories was also examined. The personal stories were elicited using a natural conversation context.

### Proximal Retell Measure

Statistically significant differences between the treatment and control children were found for the NLM. Improvements were maintained at follow-up, which occurred 4 weeks following the completion of intervention. Effect sizes at postintervention and at follow-up were moderate to large. With these results, it is likely that the 18 sessions of Tier 2 language intervention delivered in groups of four was responsible for the observed changes. However, a more rigorous design involving an active treatment control, in which commensurate time was spent with the control children, would have allowed for additional confidence in this outcome.

Although the control group mean also improved from preintervention to postintervention, which would be expected because of maturation and participation in preschool, the treatment group mean ( $M = 13.2$ ) at postintervention was almost double that of the control group ( $M = 6.7$ ) and slightly above the mean of the responders group ( $M = 11.0$ ) at the end of the dynamic assessment. Using the initial cut score of 8 as a point of comparison, significantly more children in the treatment group scored above this criterion after the 9-week intervention than children in the control group. We believe that a retell score of 8 or above is socially meaningful because it usually translates into a basic, minimally complete episode, with a problem, action, and consequence. In early childhood learning standards, retelling the beginning, middle, and end of a story is a common, parallel learning objective (Huppenthal, Johnson, & Hrabluk, 2013) and is developmentally appropriate for children age 5 years and older (Hughes, McGillivray, & Schmidek, 1997; Peterson & McCabe, 1983). An average score of 13 suggests that children who received intervention were retelling stories with complete episodes and were including other setting information such as characters and locations, emotions, and temporal or causal referents (e.g., *then*, *because*).

### Distal Retell Measure

The use of the Renfrew Bus Story was selected because the NLM aligns so closely to the intervention. As a more conservative measure of intervention effects, we administered the Renfrew Bus Story preintervention and postintervention. On the basis of the results of the information analysis, there were statistically significant differences between the groups at postintervention and follow-up assessment times. However, no differences were noted for the sentence length analysis. The combination of these two results suggest that children who received the intervention

did not produce longer sentences than children from the control group but included significantly more useful and accurate information about the story.

### Personal Stories

Personal stories are socially important for young children. The majority of narratives that young children produce through natural conversation are personal in nature (Preece, 1987), and it is through personal narratives with parents, teachers, and peers that children receive extensive language practice. Although the small-group narrative intervention procedures included opportunities for children to produce personal stories, at the postintervention and follow-up assessment times, no statistically significant treatment effects were observed. There are a few possible reasons why we did not find an effect on personal stories. First, the intervention emphasized retelling, and children had fewer opportunities to practice personal stories. Even though there may be benefits of working more intensely on personal stories, it is more challenging to elicit them in an intervention context. Many young children, especially those with language limitations or emerging English language, are reticent. Often, children do not have a story or they do not want to share a story, even if they have the language to produce it. As a result, it is easier to use story retells for explicit language intervention than personal stories. Second, our sample size was not large enough to properly power the statistical analyses. Evidence of this problem can be seen in Table 2. For the NLM, the treatment group made an average gain of 11 points from preintervention to postintervention compared with the control group's gain of 4 points. For the personal stories, the treatment group made an average gain of almost 8 points compared with the control group's 4 points, but, unlike the NLM gains, the personal story gains were not statistically significant. It is possible that with a larger sample size, statistical significance could be detected, but without additional research with preschool children, we cannot draw firm conclusions about the effect of small-group narrative intervention on children's personal stories.

### Social Validity

The teachers were asked about the feasibility, acceptability, and perceived progress of their students. All three teachers responded that the children enjoyed *Story Champs* and their language improved. The lowest score was for the item regarding how easy *Story Champs* was to learn. All three, however, reported they were interested in using *Story Champs* in the future. This level of social validity is tangential because none of the teachers actually learned the program or delivered it; they only observed. There is a significant need to examine the extent to which end users (e.g., teachers or SLPs) believe the program is worthwhile after they have actually implemented it themselves. Future research needs to address this through effectiveness research with a full-scale implementation of all the MTSLs components, including dynamic assessment, progress monitoring,

continuation of large-group language instruction, small-group intervention, and individual intervention as needed.

### Contributions, Limitations, and Future Directions

We extended Spencer and Slocum's (2010) multiple baseline study by examining a similar small-group narrative intervention with Head Start preschoolers using a small-scale randomized control group design. This study's favorable results suggest it can be grouped with the recent literature that adds to an evidence base for narrative interventions (cf. S. L. Gillam et al., 2012; S. L. Gillam, Olszewski, Fargo, & Gillam, 2014; Petersen et al., 2010; Petersen, Brown, et al., 2014; Petersen, Thompson, Guiberson, & Spencer, in press; Spencer, Kajian, Petersen, & Bilyk, 2014; Squires et al., 2014). Narrative interventions have traditionally been used to treat children with language disorders (for a review, see Petersen, 2011); however, evidence to support their use with preschoolers without disabilities is accumulating (McGregor, 2000; Spencer, Petersen, et al., 2014; Spencer & Slocum, 2010). A growing evidence base means that SLPs can more confidently use and/or recommend narrative interventions for teachers and paraprofessionals to use in their classrooms. Although the current study represents an important contribution to the literature on narrative interventions for preschoolers with risk factors, more and stronger research is necessary.

In addition to including larger samples in future research, we recommend a few other methodological improvements for this line of inquiry. We feel that more attention can be given to the promotion of personal stories and linguistic targets that transcend narrative formats and genres. It is possible that longer, more potent interventions are needed to have that kind of generalizable impact. In the current study, children received two brief interventions a week for 9 weeks. The intervention can be intensified by increasing the number of weekly sessions, increasing the number of weeks intervention is delivered, increasing the length of sessions (Fuchs & Fuchs, 2006; Warren, Fey, & Yoder, 2007), or by adding large-group narrative activities (e.g., Spencer, Petersen, et al., 2014) to the classroom routine. In school-age RTI/MTSS models, students receive Tier 2 interventions on top of their Tier 1 classroom instruction. To date, the *Story Champs* intervention arrangements have been examined in isolation, but large- and small-group interventions can be sensibly implemented together so that the latter is "extra" intervention. For preschool settings in particular, the double dose of large- and small-group instruction should make a stronger impact than either conducted in isolation. Potentially, large-group narrative intervention can focus on story grammar in a retell format, whereas in small-group interventions preschoolers can receive differentiated and tailored focus on personal stories and individualized linguistic and vocabulary targets, including those that are aimed at improving English language proficiency. In some instances, SLPs may be able to provide push-in services to children on their caseloads and include other children in the small groups who may benefit from intervention

but do not qualify for special education. It is our experience that the more SLPs can be present in classrooms, the more language education can be transferred to the teachers and the benefits of their involvement can be shared with children who are not on their caseloads.

The potential for narrative interventions to facilitate individualization indicates that groups do not need to be homogeneous—that is, children do not need to have equivalent or similar language skills in order to benefit from language intervention embedded in small-group storytelling activities. It also means that skilled interventionists can be efficient and ensure that all of the children in the group are challenged.

Although we assert that differentiated narrative intervention is preferred for SLP clinical practice, it may lack feasibility for less expert language interventionists. Two of our interventionists were college students and one was an SLP; variations among them are probable. We monitored the fidelity of the *Story Champs* intervention according to the steps involving visual scaffolding and retelling versus personal stories, but we did not document the extent to which interventionists emphasized individual targets within sessions. We were unable to determine how much each target was prompted or corrected, which limits our confidence that the intervention was uniformly delivered as intended. Future research could examine the extent to which each interventionist addresses individual targets, and it would be helpful to be able to detect differences in interventionists' ability to differentiate in small-group interventions.

An aspect of this study that is both a contribution and a limitation is the conceptualization of the featured intervention as a Tier 2 language intervention. The primary reason we classified it as a Tier 2 intervention is because the small-group intervention is the middle tier extracted from a multitiered language intervention program. With respect to RTI/MTSS for language, the current study represents a meaningful addition to a very small body of research on manualized language programs appropriate for widespread implementation in classrooms (Ukrainetz, 2006b). However, as mentioned previously, if more than half the sample qualifies for Tier 2 intervention, is a designation of Tier 2 appropriate? To answer this question, the validity of our dynamic assessment approach to identifying participants would need to be examined. Because that was not the purpose of this study, we are limited by the unknown accuracy of our identification process. There is, however, some circularity in the question as to whether a child who demonstrates limited response to an evidence-based intervention is in need of something different (in the form of a Tier 2 intervention) to incite change. Considering that Tier 2 language interventions and candidate selection are a mostly unexplored area of research, there is a great need to address these questions in future research.

Now that there is at least one carefully designed study investigating each of the various aspects of narrative intervention and assessment (i.e., tier assignment, progress monitoring, and all tiers of intervention), we recommend that the findings from these studies be integrated to create

and examine a program-wide MTSS. Although there are numerous research questions still to address, the evidence is sufficient to move toward more stringent investigations of a tiered model of language intervention and to subsequently or simultaneously explore the effectiveness of the system when implemented by end users (e.g., teachers and SLPs). The current climate within communication sciences is such that SLPs must adapt their roles to fulfill the needs of schools and clinical settings, which may involve venturing out of their comfort zone and into prevention of language-related academic problems. Preventative models of language intervention will benefit most from coteaching and coassessing partnerships with teachers and leadership from SLPs.

## References

- Ball, C. R., & Trammell, B. A.** (2011). Response-to-intervention in high-risk preschools: Critical issues for implementation. *Psychology in the Schools, 48*, 502–512.
- Barnett, D. W., VanDerHeyden, A. M., & Witt, J. C.** (2007). Achieving science-based practice through response to intervention: What it might look like in preschools. *Journal of Educational and Psychological Consultation, 17*, 31–54.
- Barrera, M., & Liu, K.** (2010). Challenges of general outcomes measurement in the RTI progress monitoring of linguistically diverse exceptional learners. *Theory Into Practice, 49*, 273–280.
- Bishop, D. V. M., & Edmundson, A.** (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52*, 156–173.
- Bradfield, T. A., Besner, A. C., Wackerle-Hollman, A. K., Albano, A. D., Rodriguez, M. C., & McConnell, S. R.** (2014). Redefining individual growth and development indicators. *Oral Language Assessment for Effective Intervention, 39*, 233–244.
- Catts, H. W.** (2009). The narrow view of reading promotes a broad view of comprehension. *Language, Speech, and Hearing Services in Schools, 40*, 178–183.
- Catts, H. W., Adolf, S. M., & Ellis Weismer, S.** (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*, 278–293.
- Catts, H. W., Petscher, Y., Schatschneider, C., Bridges, M. S., & Mendoza, K.** (2009). Floor effects associated with universal screening and their impact on early identification. *Journal of Learning Disabilities, 42*, 163–176.
- Clarke, P., Snowling, M. J., Truelove, E., & Hulme, C.** (2010). Ameliorating children's reading comprehension difficulties: A randomized controlled trial. *Psychological Science, 21*, 1106–1116.
- Collozzo, P., Gillam, R. B., Wood, M., Schnell, R. D., & Johnston, J. R.** (2011). Content and form in the narratives of children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 54*, 1609–1627.
- Cowley, J., & Glasgow, C.** (1994). *The Renfrew Bus Story*. Centreville, DE: The Centreville School.
- De La Paz, S., & Graham, S.** (1997). Strategy instruction in planning: Effects on the writing performance and behavior of students with learning difficulties. *Exceptional Children, 63*, 167–181.
- Deno, S. L.** (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*, 184–192.
- Dickinson, D., & McCabe, A.** (2001). Bringing it all together: The multiple origins, skills, and environmental supports of early literacy. *Learning Disabilities Research and Practice, 16*, 186–202.
- Dickinson, D. K., McCabe, A., & Essex, M. J.** (2006). A window of opportunity we must open to all: The case for preschool



- with high-quality support for language and literacy. In K. K. Dickinson & S. B. Newman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 11–28). New York, NY: Guilford Press.
- Dodge, D. T., Colker, L. J., & Heroman, C.** (2002). *The creative curriculum for early childhood* (4th ed.). Washington, DC: Teaching Strategies.
- Ehren, B. J., & Nelson, N. W.** (2005). The responsiveness to intervention approach and language impairment. *Topics in Language Disorders, 25*, 120–131.
- Feagans, L., & Short, E. J.** (1984). Developmental differences in the comprehension and production of narratives by reading-disabled and normally achieving children. *Child Development, 55*, 1727–1736.
- Fuchs, D., Compton, D. L., Fuchs, L. S., Bryant, J., & Davis, G.** (2008). Making “secondary intervention” work in a three-tier responsiveness-to-intervention model: Findings from the first-grade longitudinal reading study of the National Research Center on Learning Disabilities. *Reading and Writing, 21*, 413–436.
- Fuchs, D., & Fuchs, L. S.** (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*, 93–99.
- Fuchs, L. S., Fuchs, D., & Speece, D. L.** (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly, 25*, 33–46.
- Fuerstein, R., & Fuerstein, S.** (1991). Mediated learning experience: A theoretical review. In R. Feuerstein, P. S. Klein, & A. J. Tannenbaum (Eds.), *Mediated learning experience (MLE): Theoretical, psychosocial, and learning implications* (pp. 3–51). London, England: Freund.
- Gersten, R., & Dimino, J. A.** (2006). RTI (response to intervention): Rethinking special education for students with reading difficulties (yet again). *Reading Research Quarterly, 41*, 99–108.
- Gettinger, M., & Stoiber, K.** (2008). Applying a response-to-intervention model for early literacy development in low-income children. *Topics in Early Childhood Special Education, 27*, 198–213.
- Gillam, R. B., & Johnston, J. R.** (1992). Spoken and written language relationships in language/learning-impaired and normally achieving school-age children. *Journal of Speech and Hearing Research, 35*, 1303–1315.
- Gillam, S. L., Gillam, R. B., & Reece, K.** (2012). Language outcomes of contextualized and decontextualized language intervention: Results of an early efficacy study. *Language, Speech, and Hearing Services in Schools, 43*, 276–291.
- Gillam, S. L., Olszewski, A., Fargo, J., & Gillam, R. B.** (2014). Classroom-based narrative and vocabulary instruction: Results of an early-stage, nonrandomized comparison study. *Language, Speech, and Hearing Services in Schools, 45*, 204–219.
- Graham, S., & Harris, K. R.** (1993). Self-regulated strategy development: Helping students with learning problems develop as writers. *Elementary School Journal, 94*, 169–181.
- Greenhalgh, K. S., & Strong, C. J.** (2001). Literate language features in spoken narratives of children with typical language and children with language impairments. *Language, Speech, and Hearing Services in Schools, 32*, 114–125.
- Greenwood, C. R., Bradfield, T., Kaminski, R., Linas, M., Carta, J. J., & Nylander, D.** (2011). The response to intervention (RTI) approach in early childhood. *Focus on Exceptional Children, 43*(9), 1–22.
- Greenwood, C. R., Carta, J. J., Atwater, J., Goldstein, H., Kaminski, R., & McConnell, S.** (2013). Is a response to intervention (RTI) approach to preschool language and early literacy instruction needed? *Topics in Early Childhood Special Education, 33*, 48–64.
- Griffin, T. M., Hemphill, L., Camp, L., & Wolf, D. P.** (2004). Oral discourse in the preschool years and later literacy skills. *First Language, 24*, 123–147.
- Grigorenko, E. L.** (2009). Dynamic assessment and response to intervention two sides of one coin. *Journal of Learning Disabilities, 42*, 111–132.
- Hayward, D., & Schneider, P.** (2000). Effectiveness of teaching story grammar knowledge to pre-school children with language impairment. An exploratory study. *Child Language Teaching and Therapy, 16*, 255–284.
- Hudson, J. A., & Shapiro, L. R.** (1991). From knowing to telling: The development of children’s scripts, stories, and personal narratives. *Developing Narrative Structure, 89*–136.
- Hughes, D. L., McGillivray, L., & Schmidek, M.** (1997). *Guide to narrative language: Procedures for assessment*. Eau Claire, WI: Thinking Publications.
- Huppenthal, J., Johnson, J., & Hrabluk, K.** (2013, May 1). Arizona Department of Education Early Learning Standards. Retrieved July 3, 2014, from <http://www.azed.gov/early-childhood/files/2011/11/arizona-early-learning-standards-3rd-edition.pdf>
- Individuals with Disabilities Education Act, 20 U.S.C. § 1400** (2004).
- Jenkins, J. R., & O’Connor, R. E.** (2002). Early identification and intervention for young children with reading/learning disabilities. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 99–149). Mahwah, NJ: Erlbaum.
- Justice, L. M., McGinty, A. S., Cabell, S. Q., Kilday, C. R., Knighton, K., & Huffman, G.** (2010). Language and literacy curriculum supplement for preschoolers who are academically at risk: A feasibility study. *Language, Speech, and Hearing Services in Schools, 41*, 161–178.
- Justice, L. M., McGinty, A., Guo, Y., & Moore, D.** (2009). Implementation of responsiveness to intervention in early education settings. *Seminars in Speech and Language, 30*, 59–74.
- Kaderavek, J. N., & Sulzby, E.** (2000). Narrative production by children with and without specific language impairment: Oral narratives and emergent readings. *Journal of Speech, Language, and Hearing Research, 43*, 34–49.
- Kamhi, A.** (2007). Knowledge deficits: The true crisis in education. *The ASHA Leader, 12*(7), 28–29.
- Larney, R.** (2002). The relationship between early language delay and later difficulties in literacy. *Early Child Development and Care, 172*, 183–193.
- Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., Wood, F. B., . . . Olson, R.** (2001, May). Rethinking learning disabilities. In C. E. Finn Jr., A. J. Rotherham, & C. R. Hokanson Jr. (Eds.), *Rethinking special education for a new century* (pp. 259–287). Washington, DC: Thomas B. Fordham Foundation and Progressive Policy Institute.
- McGregor, K. K.** (2000). The development and enhancement of narrative skills in a preschool classroom towards a solution to clinician–client mismatch. *American Journal of Speech-Language Pathology, 9*, 55–71.
- Mellard, D., McKnight, M., & Jordan, J.** (2010). RTI tier structures and instructional intensity. *Learning Disabilities Research & Practice, 25*, 217–225.
- Nation, K., Clarke, P., Marshall, C. M., & Durand, M.** (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language, and Hearing Research, 47*, 199–211.
- Nation, K., & Snowling, M.** (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology, 67*, 359–370.



- Nation, K., & Snowling, M. J. (1998a). Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child Development*, 69, 996–1011.
- Nation, K., & Snowling, M. (1998b). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, 39, 85–101.
- National Assessment of Educational Progress. (2013). *2013 reading assessment*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.
- Nippold, M. A., & Taylor, C. L. (1995). Idiom understanding in youth: Further examination of familiarity and transparency. *Journal of Speech and Hearing Research*, 38, 426–433.
- Nippold, M. A., Ward-Lonergan, J. M., & Fanning, J. L. (2005). Persuasive writing in children, adolescents, and adults: A study of syntactic, semantic, and pragmatic development. *Language, Speech, and Hearing Services in Schools*, 36, 125–138.
- Peña, E., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology*, 10, 138–154.
- Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics: Three ways of looking at a child's narrative*. New York, NY: Plenum.
- Petersen, D. B. (2011). A systematic review of narrative-based language intervention with children who have language impairment. *Communication Disorders Quarterly*, 32, 207–220.
- Petersen, D. B., Allen, M. M., & Spencer, T. D. (2014). Predicting reading difficulty in first grade using dynamic assessment of decoding in early kindergarten: A large-scale longitudinal study. *Journal of Learning Disabilities*. Advance online publication. doi:10.1177/0022219414538518
- Petersen, D. B., Brown, C., Ukrainetz, T. A., Wise, C., Spencer, T. D., & Zebre, J. (2014). Systematic individualized narrative language intervention on the personal narratives of children with autism. *Language, Speech, and Hearing Services in Schools*, 45, 67–86.
- Petersen, D. B., Gillam, S. L., Spencer, T. D., & Gillam, R. B. (2010). The effects of literate narrative intervention on children with neurologically based language impairments: An early stage study. *Journal of Speech, Language, and Hearing Research*, 53, 961–981.
- Petersen, D. B., & Spencer, T. D. (2012). The Narrative Language Measures: Tools for language screening, progress monitoring, and intervention planning. *Perspectives on Language Learning and Education*, 19, 119–129.
- Petersen, D. B., & Spencer, T. D. (2013). *The Narrative Language Measure Scoring Guide*. Retrieved from <http://www.languagedynamicsgroup.com>
- Petersen, D. B., & Spencer, T. D. (2014). Narrative assessment and intervention: A clinical tutorial on extending explicit language instruction and progress monitoring to all students. *SIG 14 Perspectives on Communication Disorders and Sciences in Culturally and Linguistically Diverse (CLD) Populations*, 21, 5–21.
- Petersen, D. B., Thompson, B. E., Guiberson, M., & Spencer, T. D. (in press). Cross-linguistic interactions from L2 to L1 as the result of individualized narrative language intervention with children with and without language impairment. *Applied Psycholinguistics*.
- Preece, A. (1987). The range of narrative forms conversationally produced by young children. *Journal of Child Language*, 14, 353–373.
- 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.
- Spaulding, T. J., Plante, E., & Farinella, K. A. (2006). Eligibility criteria for language impairment: Is the low end of normal always appropriate? *Language, Speech, and Hearing Services in Schools*, 37, 61–72.
- Spencer, T. D., Kajian, M., Petersen, D. B., & Bilyk, N. (2014). Effects of an individualized narrative intervention on children's storytelling and comprehension skills. *Journal of Early Intervention*. Advance online publication. doi:10.1177/1053815114540002
- Spencer, T. D., & Petersen, D. B. (2012a). *Narrative Language Measures: Preschool*. Available from the Language Dynamics Group at <http://www.languagedynamicsgroup.com>
- Spencer, T. D., & Petersen, D. B. (2012b). *Story Champs: A multi-tiered language intervention program*. Retrieved from <http://www.languagedynamicsgroup.com>
- Spencer, T. D., Petersen, D. B., Slocum, T. A., & Allen, M. M. (2014). Large group narrative intervention in Head Start classrooms: Implications for response to intervention. *Journal of Early Childhood Research*. Advance online publication. doi:10.1177/1476718X13515419
- Spencer, T. D., & Slocum, T. A. (2010). The effect of a narrative intervention on story retelling and personal story generation skills of preschoolers with risk factors and narrative language delays. *Journal of Early Intervention*, 32, 178–199.
- Squires, K. E., Lugo-Neris, M. J., Peña, E. D., Bedore, L. M., Bohman, T. M., & Gillam, R. B. (2014). Story retelling by bilingual children with language impairments and typically-developing controls. *International Journal of Language and Communication Disorders*, 49, 60–74.
- Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children: The role of language comprehension and working memory skills. *Reading and Writing*, 4, 245–256.
- Strong, C. J., & Shaver, J. P. (1991). Stability of cohesion in the spoken narratives of language-impaired and normally developing school-aged children. *Journal of Speech and Hearing Research*, 34, 95–111.
- Tomblin, J. B., & Buckwalter, P. R. (1998). Heritability of poor language achievement among twins. *Journal of Speech, Language, and Hearing Research*, 41, 188–199.
- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, 40, 1245–1260.
- Ukrainetz, T. A. (Ed.). (2006a). *Contextualized language intervention: Scaffolding pre-K–12 literacy achievement*. Eau Claire, WI: Thinking Publications.
- Ukrainetz, T. A. (2006b). EPB, RTI, and the implications for SLPS: Commentary on L. M. Justice. *Language, Speech, and Hearing Services in Schools*, 37, 298–303.
- Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, 13, 70–77.
- Westby, C. E. (1985). Learning to talk—Talking to learn: Oral literate language differences. In C. S. Simon (Ed.), *Communication skills and classroom success: Therapy methodologies for language-learning disabled students* (pp. 181–218). San Diego, CA: College Hill Press.
- Wiig, E., Secord, W., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals—Preschool, Second edition (CELF Preschool-2)*. Toronto, Ontario, Canada: The Psychological Corporation.

---

## Appendix

### Story Champs: Small-Group Procedural Checklist

---

#### 1:Model Story

- \_\_\_\_\_ Lay out 5 pictures
- \_\_\_\_\_ Model the story
- \_\_\_\_\_ Place icons on pictures

#### 2:Group Retell

- \_\_\_\_\_ Leave pictures on table
- \_\_\_\_\_ Allow selection of SG icon
- \_\_\_\_\_ Each part of story retold
- \_\_\_\_\_ Each student places icon
- \_\_\_\_\_ Summarize the story

#### 3:Individual Retell 1

- \_\_\_\_\_ Leave pictures and icons out
- \_\_\_\_\_ Select preassigned student
- \_\_\_\_\_ Distribute game materials
- \_\_\_\_\_ Each part of story retold
- \_\_\_\_\_ Play story game
- \_\_\_\_\_ Summarize the story

- \_\_\_\_\_ Provide corrections immediately
  - \_\_\_\_\_ Provide least restrictive prompting

#### 4:Individual Retell 2

- \_\_\_\_\_ Remove pictures
- \_\_\_\_\_ Select preassigned student
- \_\_\_\_\_ Each part of story retold
- \_\_\_\_\_ Play story game
- \_\_\_\_\_ Summarize the story

#### 5:Individual Generation 1

- \_\_\_\_\_ Leave icons out
- \_\_\_\_\_ Select preassigned student
- \_\_\_\_\_ "Has something like that ever happened to you?"
- \_\_\_\_\_ Each part of story included
- \_\_\_\_\_ Play story game
- \_\_\_\_\_ Summarize the story

#### 6:Individual Generation 2

- \_\_\_\_\_ Remove SG icons
- \_\_\_\_\_ Select preassigned student
- \_\_\_\_\_ "Has something like that ever happened to you?"
- \_\_\_\_\_ Each part of story included
- \_\_\_\_\_ Play story game
- \_\_\_\_\_ Summarize the story

Interventionist \_\_\_\_\_

Percentage Correct \_\_\_\_\_

Observer \_\_\_\_\_

Date \_\_\_\_\_

---

*Note.* SG = small-group. Copyright © 2015 Language Dynamics Group. Reprinted with permission.

---