Bridging Oral and Written Language: An Oral Narrative Language Intervention Study With Writing Outcomes

Trina D. Spencer\textsuperscript{a} and Douglas B. Petersen\textsuperscript{a,b}

**Purpose:** Despite literature showing a correlation between oral language and written language ability, there is little evidence documenting a causal connection between oral and written language skills. The current study examines the extent to which oral language instruction using narratives impacts students’ writing skills.

**Method:** Following multiple baseline design conventions to minimize threats to internal validity, 3 groups of 1st-grade students were exposed to staggered baseline, intervention, and maintenance conditions. During the intervention condition, groups received 6 sessions of small-group oral narrative instruction over 2 weeks. Separated in the school day from the instruction, students wrote their own stories, forming the dependent variable across baseline, intervention, and maintenance conditions. Written stories were analyzed for story structure and language complexity using a narrative scoring flow chart based on current academic standards.

**Results:** Corresponding to the onset of oral narrative instruction, all but 1 student showed meaningful improvements in story writing. All 4 students, for whom improvements were observed and maintenance data were available, continued to produce written narratives above baseline levels once the instruction was withdrawn.

**Conclusions:** Results suggest that narrative instruction delivered exclusively in an oral modality had a positive effect on students’ writing. Implications include the efficiency and inclusiveness of oral language instruction to improve writing quality, especially for young students.

Most elementary school children in the United States are not writing at the level expected of them. In 2002, 72\% of fourth graders fell below grade level in writing, with even higher percentages among minority students and students with disabilities (Persky, Daane, & Jin, 2003). This discouraging pattern remains evident in the most recent National Assessment of Education Progress writing results, indicating that 73\% of eighth graders performed below the proficient level (National Center for Education Statistics, 2012). Although disappointing, these statistics are not surprising in light of how little writing instruction and writing practice students receive (Coker et al., 2016; Puranik, Al Otaiba, Sidler, & Greulich, 2014). In two national surveys of writing instruction in first through sixth grades, teachers reported that students spend less than 30 min a day writing and that writing instruction, other than instruction focused on the mechanics of printing, occurs infrequently in elementary classrooms (Cutler & Graham, 2008; Gilbert & Graham, 2010).

Current academic standards, such as the Common Core State Standards (CCSS), highlight the necessity of strong oral language and writing skills for college and career readiness, and one of the most striking shifts in academic standards is related to writing expectations for kindergarten and first-grade students. Kindergarten writing standards indicate that students should be able to “narrate a single event or several loosely linked events, tell about the events in the order in which they occurred, and provide a reaction to what happened” (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010, p. 19). In first grade, students should be able to “write narratives in which they recount two or more appropriately sequenced events, include some details regarding what happened, use temporal words to signal event order, and provide some sense of closure” (p. 19). The supposition in this shift is that, for students to be college and career ready at the end of 12th grade, they need to write coherent and minimally complete stories in first grade, and that oral narration precedes written narration. The introduction of high expectations and well-defined, ordered teaching objectives is an

*Disclosure:* Trina D. Spencer and Douglas B. Petersen developed the intervention program featured in this study and are entitled to royalties related to its sale.
important step in the right direction for writing, but without increased or improved writing-related instruction, most students will more than likely continue to fail to write at the level expected of them.

**Components of Writing and Review of Writing Interventions**

The cognitive model of the writing process focuses on planning, translating, and reviewing/rewriting (J. Hayes & Flower, 1980; J. R. Hayes, 2009, 2012a, 2012b). For writers across the life span, ideas are translated into written language, and according to this model, the translation process entails two subprocesses, transcription, where written orthographic symbols are used to represent language, and text generation, which entails creating, organizing, and elaborating ideas (V. W. Berninger, Cartwright, Yates, Swanson, & Abbott, 1994; V. Berninger & Swanson, 1994; V. Berninger et al., 2002, 1992). J. Hayes and Flower’s (1980) cognitive model of skilled writing, in conjunction with V. Berninger and Graham’s (1998) simple view of writing, focuses attention on the interplay between transcription, text generation, and working memory. Difficulty with transcription (handwriting and spelling) could lead to interference with the development of text generation, especially when the transcription process demands considerable working memory resources from a limited capacity system. When resources that need to be allocated to high-level composition skills are disproportionately allocated to transcription, text generation suffers (McCutchen, 1996). However, transcription is necessary but not sufficient for successful writing (V. Berninger et al., 2006), and even when compositional fluency is automatized, quality of writing will be heavily dependent on the foundational language skills and working memory capacity used to generate text. On the basis of the current writing expectations for kindergarten and first-grade students, writing instruction needs to begin early, and that instruction should target text generation as well as transcription.

We examined the extent to which writing intervention studies with primary grade students addressed transcription and text generation by reviewing three recent systematic reviews. Graham, McKeown, Kihara, and Harris (2012) conducted a meta-analysis of writing instruction in elementary grades. Thirty writing intervention studies with students in Grades 1–3 were included; no studies with kindergarten students were found. Transcription skills were targeted in 27% of the studies. Text generation was targeted in a small number of studies involving multicomponent or peer-mediated interventions, but only for students in second or third grade. There were no first-grade intervention studies that targeted text generation. Similarly, Datchuk and Kubina (2012) conducted a review of writing interventions for students with learning disabilities or at risk for writing disabilities. Only 10 studies included students in Grades 1–3, and none of them targeted text generation. They also reviewed nine text generation studies (e.g., sentence construction intervention), but none of the studies included students below the fourth grade. Finally, in their review of writing intervention studies for K–3 students, McMaster, Kunkel, Shin, Jung, and Lembke (2017) discovered a similar pattern. Seven transcription studies, two transcription + text generation studies, and 15 text-generation + self-regulation studies were found. Of the text generation studies, none was with students younger than second grade. These authors found one study that exclusively targeted text generation (i.e., story construction) using a computer-based instructional method with a 7-year-old child with autism (Pennington, Stenhoff, Gibson, & Ballou, 2012). There were no studies with kindergarten students, and the only studies with first graders examined handwriting interventions (McMaster et al., 2017). Collectively, these recent reviews suggest that only transcription interventions have been researched with first graders and that text generation interventions have been reserved for older students. Moreover, research with kindergartners has not focused on writing interventions. The focus on prior work with young children has been on transcription skills.

Because the writing standards for K–1 students highlight the need to address text generation in addition to transcription, it is imperative that research advances to fill this gap. Undoubtedly, K–1 writing instruction must address transcription. However, that does not obviate the need to promote text generation as well. Research indicates that, although oral and written narration are strongly associated, the development of oral narration precedes the development of written narration (Fey, Catts, Proctor-Williams, Tomblin, & Zhang, 2004; R. B. Gillam & Johnston, 1992; Scott & Windsor, 2000), suggesting that it is possible that young students can learn to generate language that is transferable to writing without having a simultaneous focus on spelling and handwriting. Therefore, the purpose of the current study was to investigate oral, narrative-based language instruction that aimed to improve first-grade students’ text generation, in the absence of transcription instruction, so that an examination of the causal relation between oral and written language would be possible.

**Narrative Language as a Bridge**

The theoretical framework that motivated this study is derived from schema theory related to narrative structure (Anderson, 1984; Mandler, 1984) and the potential of a cognitive bridge connecting oral and written language via oral narrative instruction. Oral narratives (stories) are a frequently used means of communication across different cultures and languages (McCabe & Bliss, 2003). Narration is often used to construct a common reality during communicative interactions with others. How past events are perceived and then related through narration is closely tied to the cultural models that individuals maintain (Bartlett, 1932; Bidell, Hubbard, & Weaver, 1997; Goffman, 1974; Naremore, Densmore, & Harman, 1995). Cultural models represent a community’s conscious and subconscious organized understanding of how the world works and how to interact with that world (Shore, 1996). Cultural models are
an extension of schema theory. Schema theory proposes that an individual’s perception of the world is organized according to generalized internal representations or mental models (schemata) that facilitate understanding (Anderson, 1984; Mandler, 1984). These schemata are general rules or principles that motivate expectations, and when a narrative follows an expected pattern, with agreed-upon story grammar elements (e.g., Stein & Glenn, 1979), comprehension and production in either oral or written form are facilitated (Rumelhart, 1980). The cultural model for story grammar adopted by U.S. schools is driven by current academic standards. The resultant schema (Short & Ryan, 1984; Thorndike, 1977) is thought to be foundational to the comprehension and production of narration and should therefore scaffold both oral and written narration. With intact story grammar schema, narrative text generation should be facilitated in written form.

In addition to story grammar, oral and written narratives share a number of complex literate language features such as causal and temporal subordination, use of mental-linguistic verbs, dialogue, and elaborated noun phrases (Greenhalgh & Strong, 2001; Roth & Spekman, 1986; C. Westby, 1984; C. E. Westby, 1985). Schools across the United States expect the use of academic language that first develops in students’ oral language repertoire (Ukrainetz, 2006) and in narrative genres before informational genres (Crowhurst, 1987; Langer, 1985). Oral language abilities are foundational to writing. Young students represent ideas in vocal or subvocal form before written form and often talk during the writing process (Dyson, 1983, 2009). Young students cannot write what they cannot say or think. Therefore, interventions that explicitly teach story grammar and literate language in a manner that cannot be inhibited by transcription abilities are theoretically viable.

Studies have documented a concurrent relationship between oral language and writing in kindergarten and first grade (V. W. Berninger & Abbott, 2010; Kim, Al Otaiba, Folsom, Greulich, & Puranik, 2014; Kim, Al Otaiba, Silder, & Greulich, 2013; Shanahan, 2006), and in a recent study, Kim, Al Otaiba, and Wanzek (2015) found that oral language at kindergarten predicted third-grade narrative writing quality. Moreover, Shanahan (2006) documented that verbal IQ and oral language volubility are related to stronger writing skills. It should be noted that correlational studies do not denote a causal relationship between the two modes of communication, and research has indicated that oral and written language are separate but related constructs that depend on shared brain processes (V. Berninger et al., 2006; V. W. Berninger & Abbott, 2010). However, the concurrent and predictive relationships between oral language and writing quality suggest that scaffolded input through oral language could cause improved output in writing, despite transmission through a different modality. Because story grammar serves to organize one’s thoughts and thereby enhance the expression of those thoughts, we consider narrative interventions to be potential text generation interventions.

On the basis of schema theory, we hypothesized that improvement in oral narrative language would result in improvements in written narrative language. Specifically, we were interested to see what the effects of oral narrative instruction would be on narrative writing in the absence of handwriting instruction. A classroom teacher initiated this action research study and conducted the instructional sessions as part of her core instruction with a mixed class of kindergarten and first-grade students. It is common in action research for one or more practitioners to be heavily involved in the planning and execution of the study so that the research study addresses an immediate need in the natural setting (Bradbury-Huang, 2010). In this case, the problem that activated this project was that none of the teacher’s students consistently generated a complete and coherent written story. In the absence of proper preparation for teaching writing, the teacher’s goal was to find an instructional approach that would ensure appropriate inclusion of one student with autism and advance all of her students’ abilities to generate and write their own stories. Importantly, an oral narrative language approach aligns with first-grade reading standards (e.g., CCSS [National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010]).

One oral narrative instructional program that shows promise with a broad array of students and a potential for inclusive classroom instruction is Story Champs (Spencer & Petersen, 2012). On the basis of a story grammar schema for teaching story structure (Stein & Glenn, 1979), Story Champs uses child-friendly stories, brightly colored icons, and illustrations. In addition, explicit instructional strategies with systematic visual scaffolding are included as standardized components. With guidance and support from the teacher, children receive multiple opportunities to retell stories with faded support within and across sessions. Although there are demonstrations of Story Champs delivered to whole classes of students (Spencer, Petersen, Slocum, & Allen, 2015), small group and individual arrangements promote better differentiation and individualization. A number of studies document that Story Champs, delivered in small groups, leads to improved oral storytelling among economically disadvantaged and culturally and linguistically diverse preschoolers (Spencer, Petersen, & Adams, 2015; Spencer & Slocum, 2010; Weddle, Spencer, Kajian, & Petersen, 2016) and, when delivered individually, leads to improvements in oral storytelling of preschoolers with developmental disabilities (Spencer, Kajian, Petersen, & Bilyk, 2014) and school-age students with autism (Petersen et al., 2014). Story Champs has also been used with preschool and school-age Spanish–English bilingual students in small group and individual arrangements (Petersen, Thompson, Guiberson, & Spencer, 2015; Spencer, Petersen, Restrepo, Thompson, & Gutierrez-Arvizu, 2018). With an abundance of research showing that Story Champs improves oral narratives across various groups of children, it is a logical choice for testing our hypothesis about the effect of oral narrative
instruction on written narratives. Thus, the research questions were:

1. To what extent does oral narrative language instruction improve first-graders’ narrative writing quality?
2. To what extent do improvements in narrative writing quality maintain after withdrawal of the instruction?

Method

Participants and Setting

Participants were drawn from a kindergarten/first-grade mixed class. The focus of this study was on the first-grade students, but because this was an action research study and the teacher controlled the instructional groups, the kindergartener students also received the oral narrative instruction. The teacher implemented the same procedures with the students in both grades, but the kindergarten students were not included in the study because they were unable to consistently produce legible written narratives. Although analysis of their transcription skills would have been interesting and possible (see Puranik et al., 2014), it was not the researchers’ objective in the current study. All first graders produced legible writing samples, but two students were excluded from the study because of frequent absences. Therefore, only first-grade students who had regular attendance and produced a sufficient number of legible writing samples were included as research participants (three girls and four boys). One participant (Fran) was diagnosed with autism when she was 3 years old and had an Individualized Education Program related to that diagnosis. All participants were English-speaking students from middle class backgrounds living in a western state. Five students were White, and two were Latino. The study started in January of the participants’ first-grade year. None of the students had prior exposure to Story Champs.

During literacy center rotations within the daily schedule, the classroom teacher delivered the oral narrative instruction to groups of four to five students at a time. The teacher created groups based on teacher-selected variables. The classroom teacher was a first-year dual-certified (elementary education and special education) teacher who had no prior experience with Story Champs. Small group instruction took place within the general education classroom. Writing samples were collected at a time in the regular classroom schedule designated for writing, which occurred at least an hour before the small-group Story Champs sessions.

Research Design and Procedures

To investigate the causal relationship between the independent and dependent variables, a multiple baseline design across the three groups was conducted. The seven research participants were distributed across the three groups: Group 1 had three participants including Fran, and Groups 2 and 3 had two participants each. Other students, including kindergarten students, participated in the intervention in these three groups but were not included as research participants. Each of the three groups experienced three conditions: baseline, intervention, and maintenance. The onset of the intervention phase was staggered for the three groups based on generally stable baseline performances of the research participants in the group. During the baseline condition, students did not receive oral narrative instruction. However, a writing sample was collected from each student on most days. On the days in which students were absent, a writing sample was not collected. Once the intervention condition began for a group, the students participated in six small group (four to five students) oral narrative instructional sessions. Each session lasted 20–30 min and occurred three times a week for 2 weeks. During the intervention condition, students continued to produce writing samples. For students who were available (five of the seven students), two writing samples produced 3–4 weeks after the last intervention session were used to document the extent to which writing improvements maintained.

Measurement of Writing Quality

To collect the writing samples, the classroom teacher gave each student a lined paper and said, “Please write a story. Do the best you can.” No other prompts were given. Students sat at their individual desks to write their stories. Students were allowed to ask how to spell a word, but no other assistance or directions were provided. Students were familiar with this process because it was the same as what occurred before the study; however, before the study, the teacher provided any type of assistance upon request. The classroom teacher collected the writing samples, made copies of them while removing identifying information, and gave them to the researchers. The second author who was blind to group and condition scored the writing samples using a scoring rubric (i.e., Narrative Language Measures [NLM] Flow Chart) that reflects writing quality and that is aligned with academic standards.

NLM Flow Chart

The NLM Flow Chart (Petersen & Spencer, 2016) uses a decision-making tree approach to quantify the extent to which written stories are composed of story grammar elements and to characterize the complexity of the sentences used. Examiners begin at the top of the flow chart for each element measured and then move downward answering yes/no questions until they reach the student’s present level of performance. The flow chart scoring guide is divided into two distinct sections: Story Grammar and Language Complexity. Each section is composed of individual features that are scored along a continuum of complexity, frequency, and/or clarity. This continuum provides sensitivity to subtle changes in oral and written narration over time. Flow chart items were based on their relevance to the CCSS, their relevance to narration, and their relevance to academic language.

Questions of validity of the NLM Flow Chart are soundly evidenced using findings from the oral narrative
research (Curran, Petersen, Spencer, & Heikkila, 2013; Petersen, Gillam, & Gillam, 2008; Petersen & Spencer, 2012, 2014; Pettipiece & Petersen, 2013). That research is summarized in the examiner’s manual, which includes the NLM Flow Chart and is available with a free download of the assessment (Petersen & Spencer, 2016). The NLM Flow Chart also has been used in oral narrative intervention studies (Spencer et al., 2014). Because considerably less research has been conducted on the NLM Flow Chart applied to written language, we took great care to examine the reliability of the flow chart scoring process in this study. The examiner’s manual also provides detailed information on how to score each area of the rubric. The individual who initially scored the stories and an independent scorer who scored some of the stories for interrater reliability had the manual available for reference.

Scoring Story Grammar. The Story Grammar section is modeled after story schema outlined by Stein and Glenn (1979). Although there are several ways in which gross narrative structure can be analyzed, including high-point analysis (e.g., Noel & Westby, 2014; C. Westby & Culatta, 2016; C. E. Westby, 2012), the Stein and Glenn approach is often used in research and aligns well with the oral and written narrative expectations outlined in the CCSS. Some of the story grammar elements included in the NLM Flow Chart (Petersen & Spencer, 2016) are named differently than the Story Grammar elements outlined by Stein and Glenn; however, the elements are conceptually the same. The Story Grammar section includes character, setting, problem (initiating event), plan/attempt, consequence, ending (resolution), and emotion (internal response). Story sequence is also analyzed in the Story Grammar section. When a story is scored, story grammar elements are assigned 0–3 points. The combination of clear and complete story grammar features that contribute to a complete episode is weighted for episodic complexity. Additional points are assigned for the inclusion of two or more of the following elements: problem, plan/attempt, and consequence. Thus, stories with complete or partially complete episodes receive higher scores than more basic, incomplete stories.

Scoring Language Complexity. The Language Complexity section contains language features that align with academic standards and are reflective of oral and written language (academic language) most commonly expected of children attending elementary school in the United States. The elements assessed in the Language Complexity section include prepositions, verb and noun modifiers, vocabulary and rhetoric, temporal ties, causal ties, and dialogue. When a story is scored, all Language Complexity elements are assigned 0–3 points based on their frequency and/or complexity, except dialogue, which is assigned only 0–2 points.

Composite score. A composite score was obtained by summing the Story Grammar (along with episode complexity) and Language Complexity scores. These sections of the NLM Flow Chart correspond to scoring schemes used in several previous oral narrative intervention studies (e.g., Cleave, Girolametto, Chen, & Johnson, 2010; S. L. Gillam, Gillam, & Reece, 2012; S. L. Gillam, Olszewski, Fargo, & Gillam, 2014; Petersen, Chanthongthip, Ukrainetz, Spencer, & Steeve, 2017; Spencer, Petersen, Slocum, & Allen, 2015; Spencer, Weddle, Petersen, & Adams, 2018). Yet, in contrast to many previous narrative intervention studies, the students in this study were asked to generate a written story without any imposed limitations (e.g., picture prompts or a retell context). The NLM Flow Chart scoring allows for scores up to 52, but for the participants in this study, a ceiling of 30 was observed.

Irrater Reliability

A second scorer who was also blind to group and condition rescored a subset of the writing samples to document the extent to which scoring was reliable. For each participant, four samples were randomly selected across baseline, intervention, and maintenance conditions (31% of the total number). There were 11 categories of the NLM Flow Chart that were scored (Story Grammar and Language Complexity sections), each providing an opportunity for agreement between the first and second scorers. The mean scoring agreement was 87%, with a range of 71%–100%. Lower scoring agreements were primarily from the Story Grammar section of the rubric, where the two scorers would occasionally assign scores that differed by 1 point. The authors discussed any discrepancies, yet in each case, the first rater’s scores were retained and graphically displayed. Kappa coefficients were calculated to document interrater reliability for the Story Grammar and Language Complexity sections separately. The mean correlation for Story Grammar scoring was .57 (range = .20–.85); all correlations were statistically significant (p < .05). The mean correlation for Language Complexity scoring was .69 (range = .39–1.00); all correlations were statistically significant (p < .01).

Oral Narrative Instruction

Story Champs, a multitiered language intervention program, was used as the independent variable. It was hypothesized that, through oral narrative language instruction, narrative writing would improve and result in more complete, coherent written stories. In the current study, the small-group Story Champs procedures were implemented with groups of four or five students. For each intervention session, a story was selected from the Story Champs program and was the main focus of the teaching procedures. To maximize relevance to young students, the Story Champs stories used in this study featured realistic themes such as shopping with mom, looking for misplaced shoes, and playing at the park. Stories were approximately 100 words in length and consisted of the primary story structure elements of character, setting, problem, feeling, action, ending, and end feeling. All of the stories contained the same linguistic complexity as well. For example, each story had two causal subordinate clauses (e.g., “...because they needed to buy some food.”) and two temporal subordinate clauses (e.g., “After they got to the store...”) as well as two prenominal adjectives (e.g., little oranges) and two adverbs (e.g., nicely,
mean length of utterance and word frequency were not controlled in this set of intervention stories.

The visual supports consisted of five illustrations and seven story grammar icons. Each illustration depicted one or two main parts of the story. For example, the character and setting were shown in the first illustration, the problem was shown in the second illustration, the feeling was shown in the third illustration, and so on. Story grammar icons were used to represent story structure elements that are common to all stories, whereas the illustrations depicted the content for specific stories. Icons were 1.5 in. x 1.5 in. colored circular representations of story parts, namely, character, setting, problem, feeling, action, ending, and end feeling. Additional materials included sticks, bingo cards, and cubes on which story grammar icons were affixed. These materials were used for active responding games, which are described below.

Multiple exemplar training, frequent opportunities to respond, systematic fading of visual supports, differentiated prompting, and immediate corrective feedback were the core teaching strategies embedded in the implementation of the small group instruction. The first author provided 1 hr of training and demonstration to the teacher how to deliver the instruction. In addition, she observed the teacher’s first session to make sure she was able to deliver the program effectively. To maximize consistency and fidelity to the manualized program, the teacher used a checklist during each session while implementing the procedures. As she completed each step, she referred to the checklist to make sure she delivered it correctly. The checklist was an abbreviated version of the following steps:

**Step 1: Model**

The teacher displayed the five illustrations on a table in front of the students and read the story word for word. The teacher placed the story grammar icons on or near the corresponding illustration as she read each part of the story. Before moving to the retelling steps, the teacher named the story grammar parts and asked the students to name them.

**Step 2: Team Retell**

The teacher picked up the icons and passed them out to the students randomly so that each student had one or two story parts. Beginning with the student who had the character icon and moving through the story parts in order, students took turns retelling the part of the story that corresponded with the icon they had. After each student retold his or her part of the story, the teacher summarized the entire story to present an additional cohesive model.

**Step 3: Individual Retell 1**

For Steps 3–6, an individual student retold the model story while the other students played one of the story games (i.e., story sticks, story bingo, story cubes, and story gestures). Story gestures did not require materials. To play the story games, students listened carefully to their peer retell the story. When they recognized each part of the story, they held up the corresponding stick, touched the corresponding bingo square, rotated the cube to the correct icon, or made the corresponding gesture to indicate that they understood which part the peer told. This constituted active listening to maximize the opportunities to respond even when students were not talking. In the first individual retell (Step 3), the illustrations and story grammar icons were left on the table and the student used them while retelling. At the end of the step, the teacher quickly summarized the story.

**Step 4: Individual Retell 2**

The teacher removed the illustrations from the table. A second student retold the story while the other students played one of the story games. When the student finished retelling the story, the teacher summarized it.

**Step 5: Individual Generation 1**

With the story grammar icons left on the table, the third student was asked, “Has something like that ever happened to you?” and encouraged to tell a personal story. In the event that a student could not or would not generate his or her own personal story, the teacher asked the student to retell the model story in the first person. The other students played a story game as this student recounted a personal experience related to the model story. The teacher paraphrased the student’s story.

**Step 6: Individual Generation 2**

The teacher removed the story grammar icons and asked the fourth student, “Has something like that ever happened to you?” Again, if the student did not generate his or her own personal story, the student was prompted to retell the model story in the first person. As this student told about a personal experience, the others played one of the story games. The teacher summarized the student’s story. This step was repeated if there were five students in the group.

**Step 7: Team Generation**

Once all the students had an opportunity to retell the model story or generate a personal story, the teacher randomly distributed the story grammar icons to the students like in Step 2. Starting with the character icon and moving through the parts in order, each student made up a part of a fictional story. As each part was generated, the teacher created stick figure drawings on the white board to correspond to the story parts. When all the parts had been generated, each student took turns retelling the story they had generated together.

The above steps outline the visual scaffolding and how the illustrations, icons, and active responding games were used; however, additional procedures included teacher prompting that occurred as soon as the help was needed. In other words, the teacher provided support during the
story retelling and telling rather than waiting until the students finished what they were saying. Throughout the implementation steps, the teacher used a form of prompting that was minimally intrusive, if possible. For example, if a student forgot to retell one of the story parts, she asked, “Wait. How does he feel about his problem?” If this type of prompt was sufficient to help the students retell the part, then he or she was allowed to continue retelling the story. If the prompt was ineffective, she provided a more intrusive prompt such as, “Say it like this. He felt sad because his knee hurt.” The use of minimally intrusive prompts before a more intrusive prompt ensured that students did not become prompt dependent but that they received help when necessary. The teacher only prompted the inclusion of story grammar elements and did not require students to use complex language features in their sentences.

Results

The NLM Flow Chart composite scores are displayed in Figure 1. Composite scores include the sum of Story Grammar and Language Complexity scores that ranged from 0 to 30. During baseline, all three groups of students demonstrated variable responding, with four students producing at least one story with composite scores of 14 or higher on the NLM Flow Chart. Aside from these outlying writing samples in baseline, most of the students’ samples were judged to be low and underdeveloped, with NLM Flow Chart composite scores ranging from 3 to 13. Once intervention began, all students (except Mike) quickly produced stories with scores of 20 or above. Although two students’ scores dipped during the intervention phase, most of the scores remained high, showing clear level changes and improved means from baseline to intervention (see Table 1). Although Mike received a score of 18 in baseline, making it difficult to judge effect, his mean responding in baseline was only 8.6, whereas his mean responding in the intervention condition was 19.33, suggesting an overall level change.

Five of the seven participants produced samples after 3–4 weeks of no intervention. All five of these students produced written stories with scores at or above their intervention performance, and four of the five students produced written stories above their baseline performance. Mike’s stories in the maintenance condition did not differ meaningfully from one of his baseline scores, yet his mean performance in the maintenance condition was considerably higher at 17.33.

Most of the growth observed corresponds to improvements in the Story Grammar (and episode) scoring section of the NLM Flow Chart. Although most students had improved Language Complexity scores, improvements were expectedly small with gains of 0–2 points. As a result of the narrow ranges, Story Grammar and Language Complexity scores were not graphed separately.

We calculated percentage of nonoverlapping data (PND) for each participant, with intervention and maintenance phases calculated separately. Although a PND is not a true effect size (Ledford, Wolery, & Gast, 2014), it estimates the confidence we have in the effect for each student. These data are included in Table 1. For all students except Mike, the confidence of an effect is estimated to be moderate to strong (Parker, Vannest, & Davis, 2014) based on the following interpretation bands: PNDs of > 70% indicate effectiveness, PNDs of 50%–70% are moderately effective, and PNDs below 50% are questionable (Scruggs & Mastropieri, 1998). Table 1 also displays means and gain scores.

Discussion

This text generation intervention study with first graders provides preliminary evidence of a causal relation between oral narrative instruction and written narrative outcomes. This study also represents an attempt to identify writing instruction that could accommodate the needs of diverse students, including students like Fran who had a classification of autism. All the participants except one (i.e.,
<table>
<thead>
<tr>
<th>Participants</th>
<th>PND</th>
<th>Baseline M</th>
<th>Baseline range</th>
<th>Intervention M</th>
<th>Intervention range</th>
<th>Maintenance M</th>
<th>Maintenance range</th>
<th>Gain from the highest baseline to the highest intervention points</th>
<th>Gain from the highest baseline to the highest maintenance points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fran</td>
<td>60</td>
<td>100</td>
<td>8.50</td>
<td>100</td>
<td>16</td>
<td>14.80</td>
<td>5</td>
<td>14.80</td>
<td>5</td>
</tr>
<tr>
<td>Connor</td>
<td>100</td>
<td>100</td>
<td>5.75</td>
<td>2</td>
<td>14</td>
<td>19.33</td>
<td>17</td>
<td>19.33</td>
<td>2</td>
</tr>
<tr>
<td>Zane</td>
<td>83</td>
<td>—</td>
<td>4.00</td>
<td>1</td>
<td>7</td>
<td>20.50</td>
<td>18</td>
<td>20.50</td>
<td>2</td>
</tr>
<tr>
<td>Mike</td>
<td>50</td>
<td>0</td>
<td>8.60</td>
<td>3</td>
<td>18</td>
<td>17.33</td>
<td>16</td>
<td>17.33</td>
<td>1</td>
</tr>
<tr>
<td>Bobby</td>
<td>100</td>
<td>100</td>
<td>8.30</td>
<td>2</td>
<td>12</td>
<td>24.00</td>
<td>23</td>
<td>24.00</td>
<td>1</td>
</tr>
<tr>
<td>Amy</td>
<td>100</td>
<td>100</td>
<td>7.77</td>
<td>3</td>
<td>13</td>
<td>24.50</td>
<td>24</td>
<td>24.50</td>
<td>1</td>
</tr>
<tr>
<td>Ellen</td>
<td>100</td>
<td>—</td>
<td>11.50</td>
<td>4</td>
<td>18</td>
<td>25.00</td>
<td>25</td>
<td>25.00</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. PND = percentage of nonoverlapping data; — = missing data.
Mike) made clear and meaningful growth in writing quality related to the instruction, evidenced by their behavior change during the intervention phase. After instruction, students included more story grammar elements in their stories, creating longer stories with complete episodes. These data address the first research question regarding effect of instruction. Given the experimental nature of multiple baseline designs and the adherence to quality procedures throughout the study, the results represent a strong demonstration of a causal relationship. On the basis of the What Works Clearinghouse single-case design standards (Kratochwill et al., 2010), this study meets evidence standards. The independent variable was systematically manipulated, and the researcher determined when and how the intervention was to be implemented. Two independent and blinded scorers judged the writing quality of students’ samples on 31% of the total number of samples collected, which is above the minimum standard of 20%. There were three attempts to demonstrate an effect across the three groups, which were replicated with six of the seven participants. Clear improvements related to level and variability were observed. Finally, all baseline and intervention conditions consisted of five or more data points.

The second research question addressed the extent to which improvements in writing quality maintained after a period of no instruction. Mike’s maintenance data do not help answer this question because we did not observe a clear improvement for him. If he did not show clear improvement during the intervention phase, then improvements cannot be maintained. All of the students for whom intervention effects were demonstrated, which excluded Mike, and maintenance data were obtained, including Fran, Connor, Bobby, and Amy, continued to produce written stories above their baseline performance across two writing samples obtained from each participant 3–4 weeks after the intervention phase. The absence of maintenance data on two students, for whom the instruction was effective, weakens our ability to answer this question with confidence. However, four demonstrations of maintenance were achieved from four possible participants with no overlap of data from baseline to maintenance conditions.

The story grammar outlined by Stein and Glenn (1979) is that very same structure used in the literature that students are expected to read and understand and is consequently the same structure that students are expected to use when composing. Improving students’ use of oral story grammar enhanced their use of written story grammar. This study offers evidence that the alignment between oral narrative structure and written narrative structure is not only relational but also causal.

The results of this study offer evidence of schema theory and suggest that narrative writing is not a particularly distal outcome of oral narrative instruction. An increase in the number and quality of story grammar elements primarily accounted for the growth recorded in the intervention phases of this study. This improvement in story grammar across multiple written stories is likely the result of a more developed narrative schema. Students in first grade are still mastering the transcriptional aspects of writing. However, it appears that text generation can be advanced without simultaneously addressing transcription. Because cognitive schema can be expressed in oral form, oral narrative instruction can facilitate more advanced, organized thinking, which facilitates narrative writing, as was demonstrated in the current study. This cognitive schema that spans oral and written narratives leads to improved comprehension and production of narration. The results of this study indicate that it may be only a small step, once transcription is sufficiently developed, for a student to transfer this story grammar schema from an oral modality to a written modality. Although the teacher in this study did not provide explicit instruction to the students to transfer their newly developed oral narrative skills to their writing, such a transfer appeared to happen. This oral to written transfer occurred presumably because the genre produced in writing matched the genre used for instruction, albeit a different modality. Oral narrative instruction, founded on a cognitive schema, forms a strong theoretical bridge between oral and written language.

Considering that only story grammar was explicitly targeted during the oral narrative instruction, it is not surprising that few improvements were observed in students’ use of language complexity features (e.g., prepositions, verb/noun modifiers, vocabulary/rhetoric, temporal ties, causal ties, and dialogue). It is possible, however, that, if language complexity features were explicitly targeted in the oral language instruction, we might have observed corresponding improvements in students’ writing. This conjecture is based on findings that, when language complexity features are explicitly targeted as part of the Story Champs narrative language intervention, children begin to show meaningful increases in their use of language complexity complements (Petersen et al., 2014; Weddle et al., 2016). The reason that we did not see improvements in language complexity scores could be because we did not target those skills as part of the instruction; instead, we focused on narrative structure. Future research should intentionally target language complexity to examine this relationship.

Oral language instruction to improve writing outcomes for students has many practical implications. Most text generation writing interventions are delivered to students in later elementary grades (Graham et al., 2012). However, given the current writing standards and the need to increase efforts to improve writing so that students are prepared for college and careers, text generation interventions that can begin earlier should be highly valued. Oral language interventions that address text generation can be delivered before students have mastered handwriting mechanics and spelling. Although handwriting is necessary to be able to make the transition from oral to written language, it is possible that students who have received story grammar instruction may make that transition much easier than students without such preparation, which should be examined in future research. For all students, including those with disabilities, beginning with oral language before writing is an important step (as indicated in current academic...
standards) that unfortunately is overlooked in schools because teachers often lack the expertise to explicitly teach oral language skills. In the current study, the classroom teacher was able to work with all the students in her class in small groups, despite their diverse oral and written language skills. Thus, the intervention could be used to increase inclusion of students with disabilities and students who are culturally and linguistically diverse (Spencer et al., 2014; Spencer & Slocum, 2010; Weddle et al., 2016). This study was arranged to establish experimental control across the three groups, which necessitated delivering the intervention in small groups. However, Story Champs is a multitiered curriculum that can be delivered to whole classes at a time (Spencer, Petersen, Slocum, & Allen, 2015). A whole class arrangement may be even more efficient to help all students obtain the cognitive schema necessary to better organize their thoughts for writing. For typically developing students, a low-dose intervention such as whole class may be sufficiently robust to promote oral language that influences written language. On the other hand, differentiation for diverse learners is more challenging during large group instruction, so small group instruction may be necessary for students with greater oral language needs.

Limitations and Future Directions

Notwithstanding the important contributions this study makes to the field of education, there are noteworthy limitations to mention. Because a classroom teacher initiated this study, the sample of students was convenient. The participants were not representative of all students in the United States, and their selection cannot be easily replicated. It is possible that the participants in this study already had average to advanced oral language skills and other populations with greater risk factors would not respond as rapidly to six sessions of instruction. Only one of the students had been identified as having limited language skills secondary to a diagnosis of autism spectrum disorder, with teacher report indicating that the remaining students had typical language. We did not conduct a comprehensive language assessment with the students in this study, so additional information about each participant’s language abilities could not be provided. It would have been helpful to have obtained additional information on the language skills of the participants at the outset of the study, and future research should resolve this shortcoming. The causal relation established in this study should be examined with different populations, especially with those who have more extensive oral language needs. Likewise, the intervention could be examined in a large group arrangement, which would be more characteristic of a class of typical first-grade students. Moreover, longitudinal effects of oral language instruction delivered in kindergarten on later writing should be examined in future research.

The manner in which fidelity of the intervention was monitored in this study constitutes a weakness. As an action research study, the classroom teacher was primarily driving the collection of the writing samples and the delivery of the independent variable. Although she referenced a step-by-step checklist during the intervention sessions, there was not a second independent observer available to document the extent to which she adhered to the steps. Videos of the intervention sessions were not recorded for later examination. After providing some initial training and observations, the first author (researcher) was involved from a distance helping the teacher know when to begin intervention phases and how to collect the writing samples. The first author tracked under what conditions the samples were collected but ensured that the second author did not have that information when he scored the samples. Although there is feasibility and acceptability value in action research studies, future researchers should focus on ensuring that fidelity is observed independently and that the selection process is fair and adequately described.

A third limitation is related to the measure used to judge writing quality. The NLM Flow Chart’s psychometric properties have not been examined independently of intervention studies. It was recently developed using the CCSS as a framework and then field-tested with several groups of teachers. The teachers provided detailed feedback about what to include in the scoring rubric and how to weigh each element. Although there are other writing rubrics available, most are holistic in nature (e.g., McFadden & Gillam, 1996) with poor reliability (Popp, Ryan, Thompson, & Behrens, 2003). Koutsoftas and Gray (2012) selected a scoring rubric commonly used to score state-level writing assessments for their study, but we elected to use the NLM Flow Chart (Petersen & Spencer, 2016) because it was intended to be used by classroom teachers and other educators. In the current study, the mean inter scorer agreement was 87%. The judgment of each element requires some level of subjectivity, especially with young writers who do not always spell or form letters correctly. The individual who conducted the reliability scoring was not as experienced with scoring narratives as the authors. Higher scoring reliability might have been achieved if the first author conducted the reliability scoring instead, but it was not appropriate because she was not blind to the participants’ condition. For this study, we believe that the NLM Flow Chart provided valuable and concrete measurement of the dependent variable, allowing for confident interpretations; however, future research should investigate the tool for additional evidence of validity and reliability.

Conclusion

The results of this study indicate that oral narrative instruction with first-grade students can causally impact writing in a positive direction. With the wide-scale adoption of higher academic standards, there is an increased urgency to push writing-related instruction into the early elementary school grades. By using an oral narrative instructional approach, students can benefit from the cross-modality transfer of improved oral story grammar and literate academic language to written language. The oral narrative instructional
procedures used in this study have been found to have a causal relationship to oral narrative language (e.g., Petersen et al., 2014, 2015; Spencer, Petersen, & Adams, 2015; Spencer & Slocum, 2010; Weddle et al., 2016). Writing can now be tentatively added to the list of consequences derived from the improvement of oral narration. A focus on a systemic change in the conceptualization of story grammar appears to have a pervasive effect across multiple modalities. When working with young students who are in the beginning stages of transcription, writing instruction can be unnecessarily stifled by the mechanics of writing or underdeveloped spelling skills. By addressing writing instruction in the absence of transcription instruction, students are free to improve their written text generation via oral language. Once transcription is improved, such richly developed story grammar and established organizational framework could be transferred from oral to written form.

Acknowledgments

The authors would like to thank Ms. Heidi Smith for her contributions to this research.

References


