

Volume 3, Issue 4 December 2008

EBP *briefs*

A scholarly forum for guiding evidence-based
practices in speech-language pathology

A CLINICAL LANGUAGE/LITERACY DECISION:
EVIDENCE-BASED STORY GRAMMAR INSTRUCTION

KIMBERLY A. MURZA AND CHAD NYE

DEPARTMENT OF COMMUNICATION
SCIENCES AND DISORDERS

UNIVERSITY OF CENTRAL FLORIDA, ORLANDO

EBP Briefs

A publication of Pearson

Editor

Chad Nye

University of Central Florida

Editorial Review Board

Frank Bender

Private Practice

Bertha Clark

Middle Tennessee State University

Gayle Daly

Longwood University

Donna Geffner

St. John's University

Joan Kaderavek

University of Toledo

Cheryl Lang

Detroit Public Schools

Anita McGinty

University of Virginia

Judy Montgomery

Chapman University

Barbara Moore-Brown

Anaheim Union High School District

Jamie Schwartz

University of Central Florida

Sheila Ward

Detroit Public Schools

Managing Director

Tina Eichstadt

Pearson

5601 Green Valley Drive

Bloomington, MN 55437

Cite this document as:

Murza, Kimberly A. & Nye, Chad (2008).

A Clinical Language/Literacy Decision: Evidence-Based Story Grammar Instruction

EBP Briefs 3(4), 1-14.

A Clinical Language/Literacy Decision: Evidence-Based Story Grammar Instruction

Kimberly A. Murza and Chad Nye
Department of Communication Sciences and Disorders
University of Central Florida, Orlando

Problem

Bryan is an SLP working at a small urban elementary school in a large local district. Most of his students come from lower socio-economic homes; 90% of them receive free or reduced lunch. Bryan's school is feeling a lot of pressure to make annual yearly progress as stated in the No Child Left Behind Act of 2001 (NCLB). On last year's statewide test, the majority of students in the 3rd, 4th, and 5th grades at Bryan's school were not performing at grade level. While the statewide test evaluates students' comprehension of both narrative and expository texts, younger elementary grades primarily focus on narrative text. Bryan understands that language plays an important role, especially in early reading skills, and wants to ensure that he is having an impact on his students' academic performance. In particular, one group of his language-impaired second graders is struggling with reading comprehension. They are reading fluently and accurately but their reading comprehension is not commensurate with their fluency. They are not able to answer questions based on narrative text and are especially struggling with story prediction, main idea, and story vocabulary. The statewide test requires students to read a passage, interpret it, and demonstrate understanding by answering a variety of questions on these areas.

Several months before the statewide testing, the school reading specialist discovered through the school interim testing program that several students on Bryan's caseload demonstrated high-risk performance in the area of reading comprehension. After consulting with the reading specialist, Bryan realized that he needed to learn more about how to provide appropriate literacy intervention to support the reading comprehension skills of his students.

Bryan's school district has provided a series of professional development programs for the SLPs that address many areas of language and literacy. One of the programs dealt specifically with reading comprehension and strategies the SLP might use to support the classroom curriculum. The type of strategy that seemed to best fit Bryan's situation was the story grammar approach. The question Bryan needs to answer is: Does explicit instruction in story grammar positively impact elementary school students' comprehension abilities in reading narrative text?

Background

Understanding the common attributes of language and reading comprehension is a valuable skill that SLPs contribute to the educational process and an important asset of the profession that is relevant to addressing current mandates such as NCLB. The American Speech-Language Hearing Association's (ASHA, 2001) technical report on the roles and responsibilities of SLPs in regard to literacy suggests that there is a significant overlap between reading comprehension and spoken-language comprehension. This overlap suggests that readers and listeners use similar linguistic and higher order processes, which has been extensively supported by research (Olofsson & Niedersoe, 1999; Wise, Sevcik, Morris, Lovett, Wolf, 2007; Scarborough, 1991). Additionally, reading comprehension research (Gersten et al. 1998; Roberts, Torgesen, Boardman, Scammacca, 2008) has identified major areas of common weakness for students with disabilities, including:

- deficits in decoding;
- lack of necessary background knowledge/vocabulary;
- lack of story and expository text structure; and
- difficulties with strategically processing text.

In addition, evidence has been presented that suggests that typically developing preschool-age children are able to use story grammar to organize narratives for listening comprehension (Shapiro & Hudson, 1989; Potts, 1989). Students with disabilities are slower to develop this skill and often struggle with certain aspects of story grammar, such as the ability to identify important story information and themes, and often lack the ability to make inferences (Williams, 2000).

Students with disabilities are slower to develop this skill and often struggle with certain aspects of story grammar.

Like Bryan's students, many children from lower socioeconomic status or minority backgrounds come to school exhibiting delays in the broad areas of pre-reading skills such as general oral language, vocabulary, and phonological skills (Hecht, Burgess, Torgesen, Wagner, & Rashotte, 2000; Robertson, 1998; Aikens, Barbarin,

2008). Children at risk for reading failure require a different kind of instruction than standard classroom curriculum and this alternative instruction may require the involvement of the SLP. Foorman and Torgesen (2001) have argued that reading interventions for children who are at risk will need to be more *explicit and comprehensive*, *more intensive*, and *more supportive* than programs for typically developing children. In addition, Foorman and

Reading interventions for children who are at risk will need to be more explicit and comprehensive, more intensive, and more supportive.

Torgesen suggest that children at risk may need more cognitive support in the form of scaffolding, which involves teachers guiding students in completing a task the students would not have been able to accomplish on their own. SLPs commonly use scaffolding as a means of support for their speech- and language-impaired students and are in a unique position to use these teaching/intervention skills to help children at risk for reading failure.

The focus of the remainder of this paper is to illustrate how Bryan can find an answer to his question by using evidence-based guidelines and procedures that have become a standard for identifying and implementing best practices in the schools. ASHA advocates the clinician's use of the best scientific evidence available, integrated with both the SLP's clinical expertise and the values and needs of the client (ASHA, 2005). Ehren (2008) provides a rationale for SLPs such as Bryan to engage in Evidence-Based Practice (EBP): "SLPs want to be confident that they are doing the best job they can to help struggling students...Students you serve don't have time to waste with practices that may be ineffective" (p. 2).

Method

While we recognize that access to a large and diverse collection of research literature is not readily available outside a university setting, we also recognize that the process of identifying and implementing best practices often requires the SLP to venture beyond the convenience of general public access tools such as Google™. The approach we present below is intended to reflect a typical process that a school-based SLP such as Bryan might use to answer an EBP question, including: inclusion criteria, information retrieval, research quality assessment, data analysis, and clinical application.

Scenario

Inclusion criteria.

As a first step, Bryan determined that any study included in his EBP review and synthesis would meet the following standards:

1. Implements an intervention focusing on story grammar.
2. Includes a control or comparison group condition (e.g., experimental or quasi-experimental design).
3. Provides a post-intervention comparison of outcomes measured for both treatment group and control or comparison groups.

Information retrieval.

Bryan began his search with the ASHA website (www.asha.org), since access to ASHA journals is free to members. He decided to start his search by using the key term "story grammar" with the criteria of searching "all words anywhere in article" in all of the ASHA journals. This strategy yielded 441 citations, which was promising but not specific enough to be useful at this early stage of the retrieval process. Bryan next decided to narrow his search by changing the search term criteria from "all words anywhere in article" to "exact phrase anywhere in article." This search yielded 95 citations, which was a more manageable number. A review of the titles and abstracts revealed to Bryan that many of these citations involved the evaluation of children's narrative abilities as opposed to interventions to improve reading comprehension. Of the 95 citations, three studies appeared to meet his inclusion criteria and warranted obtaining the full text (Swanson, 2005; Hoggan, 1994; Ukrainetz, 1998; McGregor, 2000). Because Bryan was able to identify so few potential studies to answer his question, he expanded his search terms to include keywords cited in the four full text studies he retrieved (e.g., "story retell," "story elements," "narrative intervention," "text structure," "text elements," "story comprehension," "narrative comprehension"). He then followed the same procedure as above in determining the relevance of the retrieved citations. Table 1 provides a description of his searches and the outcomes. Although Bryan's additional search effort ultimately provided one more citation that was relevant to his clinical question, he was confident at this point in the search that he had located all of the relevant articles in the ASHA journals.

Bryan's next step in the EBP process was to review the full text of the articles and determine whether or not they met his inclusion criteria, as described above. Three of the four articles were descriptive in nature and did not

provide statistical information regarding the use of story grammar to improve reading comprehension (Hoggan & Strong, 1994; Swanson, Fey, Mills & Hood, 2005; Ukrainetz, 1998). McGregor's article (2000) described three studies, the first two of which were evaluation studies of preschoolers' narrative abilities without treatment. McGregor's third study used a single-subject experimental design and was described as a preliminary test of an intervention in which peers facilitated narrative retellings.

While the studies provided Bryan with interesting and useful information, he recognized that the conclusions he might draw from them would not provide a scientific basis for clinical decision or implementation. This is not to say that these studies are of poor quality, but simply that they did not meet Bryan's inclusion criteria for his EBP process. He would not be able to confidently use the reported intervention strategies to treat his students' reading comprehension deficits. Since Bryan wasn't able to locate an adequate research base to potentially make a clinical judgment regarding an intervention strategy, he next went to the publicly accessible ERIC database (www.eric.ed.gov <<http://www.eric.ed.gov>>) and used similar keywords for his search as found in Table 2.

Research quality assessment.

Through his search of the ERIC database, Bryan found six additional studies. He then needed to determine whether or not these studies specifically addressed his question. One of the key considerations in a clinician's decision to accept or reject evidence is the quality of the data driving the decision. In EBP, the term "gold standard" is typically applied to studies that report a randomization procedure of participant selection or assignment to a group (treatment or control). This type of research is known as

a randomized controlled trial (RCT).

In EBP, the term "gold standard" is typically applied to studies that report a randomization procedure.

More than 10 years ago, a model was adopted in the medical arena called the CONSORT statement (Consolidated Standards of Reporting Trials) (Moher, Schulz, Altman, 2001). This model has been used to guide the reporting of

RCTs. Law and Plunkett (2006) presented an application of the CONSORT statement to assess research quality of the literature base in communication disorders. Bryan applied this model to the six RCTs he found in his search of the ERIC database in order to summarize the important elements of each study (see Appendix A).

While the information yielded by this process was important, it did not connect directly to Bryan's situation. For example, there were variations for the ages of students treated, the structure of the research design, or the measures used to assess reading comprehension. So, Bryan's next step was to decide how to use those elements of the evidence that he thought would be potentially the most critical to story grammar intervention to improve reading comprehension. After reviewing the complete CONSORT summary, he concluded that the elements most important to his situation were three variables (see Appendix A): age of participants, intervention features, and outcome measures (i.e., how the groups were evaluated post-intervention).

Decision analysis and clinical application.

Having identified the three variable elements in the CONSORT summary that were relevant to his question, Bryan could begin to evaluate the evidence and make an informed decision about implementing a story grammar intervention program for his 2nd-grade students. After reviewing the six RCTs, Bryan realized that some of the studies might not be applicable to his students. Both the Fagella-Luby, Schumaker, and Deschler (2007) study and the Dimino, Gersten, Carnine, and Blake (1990) study used 9th-grade students as participants. Although the principles of both studies might be pertinent to Bryan's second graders, the fact that the participants were ninth graders posed a potential problem. Bryan could not assume that what worked for the study's older participants would work for his students. A closer look at the specifics of the interventions revealed that they were incompatible with Bryan's caseload because both interventions required reading and writing abilities that were more advanced than the skills of those students in Bryan's group who still were learning to read.

Bryan could not assume that what worked for the study's older participants would work for his students.

The remaining four studies (Garner & Bochna, 2004; Paris & Paris, 2007; Short & Ryan 1984; Westerveld & Gullot, 2008) all used elementary students as their study participants. These studies appeared to Bryan to be more relevant to his second graders. However, a closer look at the Garner and Bochna (2004) study revealed some significant gaps in the reporting of the intervention program. For example, the frequency and duration of the control group's intervention was unclear and Bryan wondered if what the researcher reported

about the control group really made the comparison with the treated group a fair one. In addition, the study did not report where the intervention took place (e.g., in the classroom, therapy room). This missing information was problematic for Bryan and he did not feel confident in interpreting the results for use with his caseload.

Bryan then reviewed Short and Ryan (1984). Questions arose about the methodology used in this study. First, Bryan noticed a discrepancy among the number of participants reported throughout the study. Initially, Short and Ryan report that 42 less skilled readers make up the treatment groups but later in the report they refer only to 39 students. It is unclear what happened to the three missing students. In addition, the authors report using random assignment with the constraint that the three treatment groups be racially balanced. It is not explicitly stated how the authors ensure the groups are racially balanced. It seemed to Bryan that this procedure should not be called randomized due to these constraints in assignment. These discrepancies concerned Bryan. As in his review of the Garner and Bochna (2004) study,

Bryan did not feel confident in interpreting the results of the Short and Ryan study for use with his students.

Bryan moved on to the Paris and Paris (2007) study. He learned that the study's participants were first graders but that the

intervention procedures were more complete than the procedures presented in the other studies and were described in detail. The authors provide extensive explanation of the intervention program, the narrative strategy instruction (NSI), and the comparison condition, along with all of the pre-testing and posttesting measures and results. The NSI consists of one unit taught twice each week for 5 weeks for a total of 10 lessons, each 45 minutes long (total time ~ 450 minutes). Although the NSI lessons are taught in the classrooms using whole-group instruction, Bryan felt that the frequency and duration of the intervention lessons could be duplicated in his therapy room with his target students.

Bryan was particularly interested in the Paris and Paris (2007) study for two major reasons: (1) they reported significantly positive effects of their program for implicit comprehension ($d=.43$) and explicit comprehension ($d=.70$), and (2) the specifics of the NSI instructional program were laid out in the paper so that Bryan felt confident that he had enough information to duplicate the

intervention with relative fidelity. The NSI intervention:

- uses picture books to teach students strategies for understanding story grammar;
- teaches students how to make inferences about story characters' feelings and thoughts;
- teaches students how to make predictions and story inferences; and
- teaches students how to retell the story using summarization and sequencing skills.

All of these tasks were part of the 2nd-grade curriculum for Bryan's students and the state standards for reading instruction assessed in the Spring by his school and district.

Bryan not only felt confident that he could duplicate the NSI intervention, he also believed that he had both the time and resources to do so. He concluded that with a little planning he would be able to use stories from his school's library to teach his students the same strategies as described in the study. One of the main factors that Bryan considered is that he believed that a total instructional time of 450 minutes was a reasonable amount of time to realize the benefit to his students.

He next reviewed Westerveld and Gillon (2008) and found it to be a high-quality research study that used similar intervention methods to the Paris and Paris study (2007) (i.e., teaching students specific story grammar components, using visual strategies to enhance comprehension). The study also reports significantly positive effects of their program for a story comprehension task ($d=1.55$) and a story comprehension probe ($d=1.89$). Comparing Paris and Paris to Westerveld and Gillon (2008), Bryan noticed a large difference in the number of participants in each study (Paris & Paris: experimental group $n=83$, control group $n=40$; Westerveld & Gillon: experimental group $n=5$, control group $n=5$). As a rule of thumb, the greater the number of participants in a study, the more confident one can be in the generalization of the reported effects. For this reason, Bryan's final decision was to use the NSI for five weeks and then evaluate the effectiveness of the program with his students.

Bryan succeeded in his attempt to use current research evidence in his clinical decision-making but he is far from finished. He must continue to use evidence to guide his therapy. Just as the authors in the Paris and

A total instructional time of 450 minutes was a reasonable amount of time to realize the benefit to his students.

The greater the number of participants in a study, the more confident one can be in the generalization of the reported effects.

Paris (2007) study did, Bryan decided to create similar pre-tests and posttests for the skills he will be teaching his students. He also will continue to use therapy data to gauge his students' comprehension of his lessons and their performance throughout the NSI units. Since schoolwide standardized testing in both reading and math is conducted quarterly in Bryan's school, he also has the opportunity to see if these therapy skills are indeed generalizing out of the therapy room.

Follow-up Questions

In spite of the fact that in our scenario Bryan has demonstrated one way to deal with research in an EBP context, we all need to recognize that there are limitations to Bryan's evidence-based process. How will he know if any changes in his students' behavior can be directly attributed to what he has done with them? Bryan's students will be working on some of the very same skills in their classrooms that they are working on in the therapy room. If there is success, one could argue

Bryan must collaborate with the teachers at his school to ensure he is having an impact on his students' academic performance.

that it is due to the collaborative and duplicate nature of the intervention. However, if the student fails, the cause may be difficult to determine. It could be because Bryan's program was inappropriately delivered, or because the classroom teacher was inadequately trained, or due to any number of other reasons. The fact remains that ultimately Bryan must collaborate with the teachers at his school to ensure he is having an impact on his students' academic performance. He will need to continually evaluate the intervention programs, assess student performance, and gauge the appropriateness of the intervention strategies for his students' language and literacy development.

References

- Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: the contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology, 100*(2), 235-251.
- American Speech-Language-Hearing Association. (2005). *Evidence-based practice in communication disorders* [Position Statement]. Available from www.asha.org/policy.
- American Speech-Language-Hearing Association. (2001). *Roles and responsibilities of speech-language pathologists with respect to reading and writing in children and adolescents* (position statement, guidelines, technical report and knowledge and skills required). Rockville, MD: Author.
- Dimino, J., Gersten, R., Carnine, D., & Blake, G. (1990). Story grammar: An approach to promoting at-risk students' comprehension of literature. *The Elementary School Journal, 91*(1), 19-32.
- Ehren, B. J. (2008). Making informed decisions about literacy interventions in schools: an adolescent literacy example. *EBP Briefs, 3*(1), 1-11.
- Faggella-Luby, M., Schumaker, J. S., & Deshler, D.D. (2007). Embedded learning strategy instruction: story-structure pedagogy in heterogeneous secondary literature classes. *Learning Disability Quarterly, 30*, 131-147.
- Foorman, B.R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research & Practice, 16*(4), 203-212.
- Garner, J. K., & Bochna, C. R. (2004). Transfer of a listening comprehension strategy to independent reading in first-grade students. *Early Childhood Education Journal, 32*(2), 69-74.
- Gersten, R., Williams, J. P., Fuchs, L., Baker, S., Koppenhaver, D., Spadorcia, S., et al. (1998). *Improving reading comprehension for children with disabilities. A review of research. Final report* (Contract No. HS92017001). Washington, DC: U.S. Office of Special Education Programs.
- Hecht, S. A., Burgess, S. R., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2000). Explaining social class differences in growth of reading skills from beginning kindergarten through fourth-grade: the role of phonological awareness, rate of access, and print knowledge. *Reading and Writing, 12*(1-2), 99-127.
- Hoggan, K. C., & Strong, C. J. (1994). The magic of "Once upon a time": narrative teaching strategies. *Language, Speech, and Hearing Services in the Schools, 25*, 76-89.
- Law, J., & Plunkett, C. (2006). Grading study quality in systematic reviews. *Contemporary Issues in Communication Science and Disorders, 33*, 28-36.
- 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.
- Moher, D., Schulz, K. F., & Altman, D. G. for the CONSORT Group. (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet, 35*, 1191-1194.
- No Child Left Behind Act of 2001, Pub. L. No 107-110 20 U.S. C. §§ 6301 *et seq.*
- Olofsson, A., & Niedersøe, J. (1999). Early language development and kindergarten phonological awareness as predictors of reading problems: from 3 to 11 years of age. *Journal of Learning Disabilities, 32*(5), 464-472.
- Paris, A. H., & Paris, S. G. (2007). Teaching narrative comprehension strategies to first graders. *Cognition and Instruction, 25*(1), 1-44.
- Potts, R. (1989). *West side stories: children's conversational narratives in a black community*. Paper presented at the National Biennial Meeting of the Society for Research in Child Development, Kansas City, MO.
- Roberts, G., Torgesen, J. K., Boardman, A., & Scammacca, N. (2008). Evidence-based strategies for reading instruction of older students with learning disabilities. *Learning Disabilities Research & Practice, 23*(2), 63-69.

- Robertson, K. L. (1998). Phonological awareness and reading achievement of children from differing socio-economic status backgrounds (Doctoral dissertation, University of Rhode Island, 1998). *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 58(8-A), 3066.
- Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development*, 61(6), 1728-1743.
- Shapiro, L. R., & Hudson, J. A. (1989). *Cohesion and coherence in preschool children's picture-elicited narratives*. Paper presented at the National Biennial Meeting of the Society for Research in Child Development, Kansas City, MO.
- Short, E. J., & Ryan, E. B. (1984). Metacognitive differences between skilled and less skilled readers: Remediating deficits through story grammar and attribution training. *Journal of Educational Psychology*, 72(2), 225-235.
- Swanson, L. A., Fey, M. E., Mills, C. E., & Hood, L. S. (2005). Use of narrative-based language intervention with children who have specific language impairment. *American Journal of Speech-Language Pathology*, 14, 131-143.
- Ukrainetz, T. A. (1998). Stickwriting stories: a quick and easy narrative representation strategy. *Language, Speech, and Hearing Services in the Schools*, 29, 197-206.
- Westerveld, M. F., & Gillon, G. T. (2008). Oral narrative intervention for children with mixed reading ability. *Child Language and Teaching Therapy*, 24(1), 31-54.
- Williams, J. P. (2000). *Strategic processing of text: improving reading comprehension of students with learning disabilities* (ERIC/OESP Digest #599). Arlington, VA: ERIC Clearinghouse on Disabilities and Gifted Education, Council for Exceptional Children.
- Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among expressive and receptive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech-Language and Hearing Research*. 50, 1093-1109.

Table 1. Search Strategy

Keyword(s)	Number of citations	Number of citations remaining after review	Citations
story grammar	95	3	Swanson, Fey, Mills, & Hood, 2005; Ukrainetz, 1998; Hoggan & Strong, 1994
story retell	43	0*	
story elements	28	1*	McGregor, 2000
narrative intervention	14	0*	
text structure	21	0*	
text elements	4	0*	
story comprehension	89	0*	
narrative comprehension	25	0*	

*Number of new citations not found in the previous searches

Table 2. ERIC Search Strategy

Keyword(s)	Number of citations	Number of citations remaining after review	Citations
story grammar AND reading comprehension AND intervention	12	5	Westerveld & Gillon, 2008; Fagella-Luby, Schumaker, & Deschler, 2007; Dimino, Gersten, Carnine, & Blake, 1990; Short & Ryan, 1984
story retell AND reading comprehension AND intervention	3	0	
story elements AND reading comprehension AND intervention	11	2	Paris & Paris, 2007; Garner & Bochna, 2004
narrative intervention AND reading comprehension AND intervention	32	0	
text structure AND reading comprehension AND intervention	36	0	
text elements AND reading comprehension AND intervention	5	0	
story comprehension AND reading comprehension AND intervention	87	0	
narrative comprehension AND reading comprehension AND intervention	32	0	

Appendix A: CONSORT Criteria (Law & Plunkett, 2006)

Headings	Subheadings	Garner & Bochna, 2004	Paris & Paris, 2007	Short & Ryan, 1984	Westerveld & Gillon, 2008
Title and abstract		The title does not provide information about study design, but the abstract indicates that a comparison group was used.	The title does not provide information about study design, but the abstract indicates that a comparison group was used.	The title does not provide information about study design, but the abstract indicates that a comparison group was used.	The title does not provide information about study design, but the abstract indicates that a comparison group was used.
Introduction		The paper's introduction provides the research base on story grammar development and instruction and a rationale for the current research.	The introduction provides a detailed discussion of the research in narrative instruction and story grammar. It also provides a clear rationale for the current study.	The introduction provides a detailed discussion on meta-cognitive strategies including information on the two types of interventions: attribution training and story grammar training.	The introduction provides a detailed discussion of the research in narrative structure comprehension and production. It also provides a clear rationale for the current study.
Methods	Participants	Selection criteria are not clearly presented in the paper. Treatment group n=37, comparison group n not clear. Authors reports n=16 girls but not total n. The setting of the intervention was not stated in the paper.	One hundred and twenty-three students in six first-grade classrooms in one elementary school. Demographic data is provided in Table 1. All students whose parents gave consent for the study participated.	Forty-two 4th-grade, less skilled readers were randomly assigned to one of three treatment groups. Fourteen 4th-grade skilled readers did not receive treatment and served as a comparison group using posttest measures.	Ten children identified as poor readers between the ages of 7:11 and 9:2 from the original longitudinal study who were available for intervention. Five of the children formed the treatment group. The remaining five children served as the poor reading control group and received delayed treatment. Ten typically developing children from the longitudinal study served as the control group.

Appendix A: CONSORT Criteria (Law & Plunkett, 2006)

Headings	Subheadings	Garner & Bochna, 2004	Paris & Paris, 2007	Short & Ryan, 1984	Westerveld & Gillon, 2008
	Interventions	<p>A brief explanation of the intervention in both groups was presented. The intervention condition utilized direct explanation and guided practice to teach students: main character, setting, problem, attempted solution, and solution. The comparison group used their language arts basal series to read and listen to stories in comparable frequency to the intervention group. No references to story structure were used.</p> <p>The intervention program received 15-20 minutes of instruction, including 8-10 teacher-read stories per element, twice daily for 16 weeks.</p> <p>The authors did not specify the intensity of the comparison program.</p>	<p>The NSI (Narrative Strategy Instruction) condition consisted of five specific NSI units that were created for the study. Each unit was taught over two lessons in a one-week period. Each class in the NSI condition received 10 lessons, each 45-minutes long, for a total instructional time of 450 minutes. The same instructor (one of the researchers) provided all lessons. A detailed explanation of the instructional sessions is also presented by the authors.</p> <p>The comparison condition utilized the same amount of instructional time in the same manner using the poetry genre instead of the narratives used in the intervention condition.</p>	<p>The total training group received story grammar strategy training and attribution training (designed to increase awareness of effort in efficient reading). The strategy group received only story grammar strategy training. The control group of less skilled readers received only attribution training. A detailed explanation of the instructional sessions is presented by the authors.</p> <p>All treatment groups received three sessions occurring over one week, each 30-35 minutes long. The experimenter provided the intervention.</p>	<p>The treatment group received 12 one-hour group sessions over a six-week period. Each intervention session consisted of instruction of a particular story element. The students listen to part of a story, identify the story element using a story map, and retell the story with feedback provided by other group members.</p>
Objectives		<p>The objectives are stated clearly as the following research questions: (a) What, if any, pre-existing story grammar knowledge exists? (b) What is the benefit of direct instruction over and above mere exposure to narrative stories, and on what tasks might new knowledge be brought to bear?</p>	<p>The objectives are stated explicitly in the form of five research questions: (a) Did NSI have positive effects on children's narrative comprehension skills? (b) Did NSI generalize to expository text? (c) Did the effects of the treatment generalize to listening comprehension? (d) Did the effects of the treatment generalize to oral production? (e) Did NSI benefit all students?</p>	<p>The objectives are stated clearly as the following research questions: (a) Do strategy-trained, less skilled readers differ from skilled readers in their ability to utilize story schemata to aid their comprehension of new information? (b) Do less skilled readers benefit from training with a self-instructional story grammar strategy? (c) Are the benefits reaped from strategy training enhanced by attribution training?</p>	<p>The objectives are stated clearly as the following research questions: a) Does oral narrative intervention enhance the narrative production and comprehension skills of children with a mixed reading ability? (b) Does oral narrative intervention improve these children's reading comprehension performance?</p>

Appendix A: CONSORT Criteria (Law & Plunkett, 2006)

Headings	Subheadings	Garner & Bochma, 2004	Paris & Paris, 2007	Short & Ryan, 1984	Westerveld & Gillon, 2008
	Sample size	Treatment group n=37, comparison group n not clear. Authors reports n=16 girls but not total n.	Treatment group n=83 in four classrooms, comparison group n=40 in two classrooms.	Authors initially report that 14 skilled and 42 less skilled readers participated, but later reported 14 skilled and 39 less skilled readers.	Treatment group n=5, control group of poor readers n=5, control group of typically developing students n=10.
	Randomization	Students were randomly allocated to either oral (n=8) or silent reading (n=9) treatment groups or a comparison group (14) Discrepancy from sample size not described by authors.	Six classrooms were randomly assigned to either the intervention or comparison condition (four classrooms in the experimental NSI condition, two classrooms in the comparison condition).	The less skilled readers were randomly assigned to one of three treatment groups. However, the authors report that they ensured the groups were racially balanced. It is unclear how they did this.	Ten poor readers were allocated to two groups based on the geographical location of their school. The two groups were randomly assigned to the treatment or delayed treatment groups.
	Blinding	No information was provided on blinding in the paper.	No information was provided on blinding in the paper.	No information was provided on blinding in the paper.	No information was provided on blinding in the paper.
	Statistical methods	ANOVA conducted for Reading Comprehension and Listening Comprehension. Subgroup analysis for correct responders to story element measure.	The study's data were analyzed using a 2 (time) x 2 (condition) repeated measures analyses of covariance (ANCOVA & MANCOVA) to test for interaction effects. There was a mismatch of assignment and analysis.	The study used a multivariate analysis of variance with group as a factor to analyze the data.	The study used a non-equivalent pre-test/posttest control group design in which one group received delayed treatment and served as the control group. There was a mismatch of assignment and analysis.
Results	Participant flow	Discrepancy between sample size and analysis n values not described.	It appears that all participants were present during all stages of the study.		It appears that all participants were present during all stages of the study.
	Recruitment	Recruitment of the participants was not specified in the paper.	Recruitment of the participants was not specified in the paper other than that they were all first graders at a single elementary school whose parents gave consent for participation in the study.	Recruitment of the participants was not specified in the paper other than that they were 4th-grade boys from four public and six private schools selected based on their reading comprehension scores from the Stanford Diagnostic Reading Test. All treatment participants scored at or below the 50th percentile on this measure.	The participants were originally recruited for participation in a longitudinal study being conducted by the authors. The 10 poor readers were those who were available for intervention. The 10 typically developing readers also were recruited originally for participation in the longitudinal study and matched the 10 poor readers on age, gender, ethnicity, and grade level.

Appendix A: CONSORT Criteria (Law & Plunkett, 2006)

Headings	Subheadings	Garner & Bochna, 2004	Paris & Paris, 2007	Short & Ryan, 1984	Westerveld & Gillon, 2008
	Baseline data	Baseline demographic and performance data were not provided in the paper.	Baseline demographic data were provided in Table 1 for all students. Pre-test data were provided on all participants in all evaluation measures (Narrative Comprehension task, Expository Comprehension task, Narrative Production task, Expository Production task, Listening Comprehension tasks, Woodcock-Johnson III Picture Vocabulary test, Michigan Literacy Progress Profile Rhyming, Blending, and Phoneme Segmentation, Teacher Ratings, and Literacy Habits Interview).	Some baseline demographic data were provided (i.e., Stanford Diagnostic Reading Test, Comprehension Score) were used for all children to determine whether they were considered "skilled" or "less skilled" readers. Additional pre-test measures included for only the less skilled readers included free recall and probed recall.	Baseline demographic and pre-test data were provided for all participants. Pre-test data were provided in the following evaluation measures: Neale Analysis of Reading Ability, Clinical Evaluation of Language Fundamentals, 3rd Edition, Test of Nonverbal Intelligence, 3rd Edition, Peabody Picture Vocabulary Test, 3rd Edition.
	Numbers analyzed	The study reported that 35 students were assigned to the intervention group and 31 students to the comparison group for Reading Comprehension and Listening Comprehension analysis at the posttest. Discrepancy between participants at the pre-test and posttest not described.	The study reported that 83 students in four classrooms were assigned to the intervention group and 40 students in two classrooms were assigned to the comparison condition. It appears that all students were present for pre-testing and posttesting.	Though there was initially a discrepancy in the number of participants in the less skilled group, it appears that ultimately 39 students in the treatment group and 14 students in the control group were present for pre-testing and posttesting. The authors did note that one child was excluded from the recall analyses due to equipment failure.	The study reported that five students were assigned to the treatment group, five students were assigned to the poor reader control group, and 10 children were assigned to the typically developing control group. It appears that all students were present for pre-testing and posttesting.
Outcomes and estimation	Outcome measures included: Reading Comprehension (reading ability, story structure, free recall and prompted recall) and Listening Comprehension (free recall and prompted recall). The means, standard deviations, and p-values for composite scores are provided.	Outcome measures included: Narrative Comprehension task, Expository Comprehension task, Narrative Production task, Expository Production task, Listening Comprehension tasks, Memory Questionnaire, and Effort and Enjoyment Questionnaire. Some variables were excluded from analysis due to ceiling effects, inter-correlation, and leptokurtic distribution of participant performances.	Outcome measures included: free recall, probed recall, note-taking, Reading Concept Inventory, test of error detection and correction.	Outcome measures included: Neale Analysis of Reading Ability, Oral Narrative Production, Oral Narrative Comprehension.	Outcome measures included: Neale Analysis of Reading Ability, Oral Narrative Production, Oral Narrative Comprehension.

Appendix A: CONSORT Criteria (Law & Plunkett, 2006)

Headings	Subheadings	Garner & Bochma, 2004	Paris & Paris, 2007	Short & Ryan, 1984	Westerveld & Gillon, 2008
	Ancillary analyses	t-test for story element in Tx group	Teacher Ratings and a Literacy Habits Interview were used at pre-testing. A Memory Questionnaire and an Effort and Enjoyment Questionnaire were used at posttesting.	The information subtest of the Wechsler Intelligence Scale for Children-Revised Edition was used to control for any discrepancy between groups in cognitive ability. Also, the comprehension subtest of the Stanford Diagnostic Reading Test was used as pre-testing only.	
	Adverse events	No details for adverse effects.	No details for adverse effects.	No details for adverse effects.	No details for adverse effects.
Discussion	Interpretation	The authors interpret the results in reference to their hypotheses and research questions. They also acknowledge additional areas of research that remain unanswered by their study.	The authors interpret the results in detail, addressing their findings' relevance to the research base they discussed in the introduction. They address and answer each study question in depth and acknowledge the limitations of the study.	The authors interpret the results in reference to their research questions. They also acknowledge the limitations of the study.	The authors interpret the results in reference to their research questions. They also acknowledge the limitations of the study.
	Generalizability	The paper implies generalization of the story grammar instruction to beginning readers when using the same instructional techniques.	The study addresses the question of whether the instructional skills could be generalized. The paper discusses the results and also how to generalize effects in other classroom settings.	The authors call one of the outcome measures "Strategy Generalization" but it appears to be more of a measure of maintenance as it occurs three days later. The authors discuss maintenance and the anecdotal evidence for generalization but also limitations of the interpretation.	The study addresses the question of whether the instructional skills could be generalized to reading comprehension. The authors discuss these results, including the lack of transfer of the skills, and offer some explanation.
	Overall evidence	The findings of the study are discussed relative to the research question that they answered. Suggestions for future research are made in relation to questions that the authors were not able to answer.	The findings of the study are discussed relative to the research question that they answered. Suggestions for future research are made in relation to other populations, ages, and other areas that are lacking a research base.	The findings of the study are discussed relative to the research question that they answered. Suggestions for future research are made in relation to questions that the authors were not able to answer.	The findings of the study are discussed relative to the research question that they answered. However, suggestions for future research are not made.