

INTERVENTION, EVALUATION, AND POLICY STUDIES

Enhancing Social Studies Vocabulary and Comprehension for Seventh-Grade English Language Learners: Findings From Two Experimental Studies

**Sharon Vaughn, Leticia R. Martinez, Sylvia Linan-Thompson,
and Colleen K. Reutebuch**

The Meadows Center for Preventing Educational Risk, The University of Texas at
Austin, Austin, Texas, USA

Coleen D. Carlson and David J. Francis

Texas Institute for Measurement, Evaluation, and Statistics, University of Houston,
Houston, Texas, USA

Abstract: Two experimental studies to improve vocabulary knowledge and comprehension were conducted in 7th-grade social studies classes with English language learners (ELLs). Two different nonoverlapping samples of classes of 7th-grade students ($N = 381$ and $N = 507$) were randomly assigned at the classroom (i.e., section) level to a social studies intervention or to business as usual comparison groups. The number of sections assigned to treatment was 7 and 9 in Experiments 1 and 2, respectively. Eight sections were assigned to comparison in each experiment. In addition, students were randomly assigned to sections prior to assignment of sections to treatment and control. Treatment students received a multicomponent social studies instruction including explicit vocabulary instruction, use of structured pairing, strategic use of video to build concepts and promote discussion, and use of graphic organizers for approximately 12 weeks daily during social studies class. Findings indicated significant differences in favor of the treatment students on curriculum-based vocabulary and comprehension measures for both experimental studies for all students including students who were ELLs.

Keywords: Adolescent English Language Learners, middle school, academic language development, vocabulary and content knowledge, social studies instruction

Address correspondence to Sharon Vaughn, The University of Texas at Austin, College of Education, One University Station, SZB D4900, Austin, TX 78712-0365, USA.
E-mail: SRVaughn@teachnet.edb.utexas.edu

Adolescent English language learners (ELLs) who lack academic English language knowledge and demonstrate low literacy levels are at risk for academic failure in content area classes (Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006). We conducted two experimental studies involving two nonoverlapping samples of seventh graders who were targeted because they included large numbers of ELLs. The goal was to evaluate the effects of instructional practices implemented by social studies teachers to improve vocabulary and comprehension knowledge of their middle-school students in classes that included ELLs.

BACKGROUND ON ELLs

In the United States, ELLs, estimated to include more than 10.8 million students, represent the fastest growing segment among the school-age population (NCES, 2008). Although students who are ELLs are heterogeneous and speak more than 400 different languages, the majority (80%) are Spanish speakers (Kindler, 2002). The academic achievement of ELLs on the whole is low (Lee, Grigg, & Donahue, 2007) because of a variety of individual factors that influence overall academic success including educational history, school placement and instruction, and English language literacy ability and sociocultural background (August & Shanahan, 2006b). Only 4% of eighth-grade ELLs scored at or above the proficient level on national achievement tests in reading compared with 31% of native English speakers in 2007 (Lee et al., 2007), and the graduation rate of ELLs is far lower than their native English-speaking peers (Laird, DeBell, & Chapman, 2006). Although reported statistics often misestimate educational outcomes for ELLs by failing to take into account the performance of students who have been reclassified as fluent English Proficient following gains in English proficiency, there is little doubt that ELLs are at risk educationally (Francis & Rivera, 2007).

In the upper grades, ELLs face challenges because of struggles with academic text, lack of content-area knowledge, and underdeveloped oral language and vocabulary levels that can hamper their academic achievement. A lack of proficiency in academic language, necessary for success with content-area learning, often impedes older second language learners in their abilities to comprehend and analyze middle and high school texts, as well as to express themselves proficiently in academic English oral and written tasks (Francis et al., 2007). According to Short and Fitzsimmons (2007), ELLs must perform twice the work of their monolingual peers because they are learning English while learning content, and all the while held to the same accountability standards as their English-only counterparts.

RESEARCH ON INSTRUCTIONAL PRACTICES

Efforts to boost the academic achievement of older ELLs have become a focal point for researchers, policymakers, and school districts leaders. However, research regarding evidenced-based instruction for ELL adolescents is

lacking. Our goal was to address the need for research in this area by examining the effects of an enhanced social studies instruction designed specifically for students who are ELLs that would benefit all students. We identified instructional practices associated with improved outcomes for ELLs that were feasible for implementation by classroom teachers and recommended as part of high-quality instruction for all learners: (a) research-based vocabulary and concept instruction, (b) the use of media to build comprehension and concept knowledge, (c) the use of graphic organizers, and (d) structured peer-pairings.

Explicit Vocabulary and Concept Instruction

The body of research available on English-only students has identified effective instructional vocabulary strategies such as providing definitional and contextual information about each word's meaning and actively involving students in word learning through talking about, comparing, analyzing, and using targeted words (Beck & McKeown, 2001; Beck, McKeown, & Kucan, 2002; Stahl, 1999). Further, research has shown that students' ability to acquire textbook vocabulary is enhanced when explicit vocabulary instruction is integrated into existing content-area curriculum (Baumann, Edwards, Boland, Olejnik, & Kame'enui, 2003; Bos & Anders, 1990). A meta-analysis on struggling adolescent readers found that older students with reading difficulties benefited from improved knowledge of word meanings and concepts (Scammacca et al., 2007). Knowing the meaning of words relates strongly to comprehension and overall academic success (Baumann, Kame'enui, & Ash, 2003; National Institute of Child Health and Human Development [NICHD], 2000).

One instructional practice that may be especially important for developing the vocabulary of ELLs includes taking advantage of a student's first language knowledge by increasing students' awareness of cognates to extend their vocabulary, thereby also improving their comprehension (August, Carlo, Dressler, & Snow, 2005). Words that are common in Spanish are often cognates of less familiar words in English and appear in academic texts (e.g., "infirm"—"enfermo"; August et al., 2005). Studies of vocabulary instruction also indicate that ELLs are more likely to learn words that are directly taught and embedded in meaningful contexts with many opportunities for repetition and use. Syntheses conducted by the National Literacy Panel and Center for Research on Education, Diversity and Excellence (Genesee, Lindholm-Leary, Saunders, & Christian, 2006) concluded that a student's primary language could be used to preview or introduce new vocabulary and concepts prior to a lesson in English.

Strategic Use of Video and Purposeful Discussion to Build Concepts

The anchored instruction approach is a strategy to help students become more actively engaged in learning by situating or anchoring instruction around

meaningful context (Cognition and Technology Group at Vanderbilt, 1990). Incorporating short video clips into lessons is one way to anchor instruction for students, that is, to provide students who struggle with meaningful access to the curriculum. Video segments serve as another content source to accompany text and discussions and to provide additional background information on topics that are unfamiliar to students (Gersten, Baker, Smith-Johnson, Dimino, & Peterson, 2006). The use of video can also help generate discussion, which in turn supports students' active involvement in learning the content. In their study on teaching historical content to middle school students with learning disabilities, Gersten and colleagues (2006), found that providing instruction that supports active involvement in the learning process (i.e., structured peer-pairings) and presenting materials that supplement traditional textbook reading (e.g., presenting both a magazine article and a short video clip on a key figure or event) enhanced participants understanding of complex grade-level material.

Use of Writing With Graphic Organizers

The use of graphic organizers (i.e., semantic maps, advanced organizers, Venn diagrams, story maps, concept diagrams, etc.) to assist students to organize information and their thoughts has led to positive effects on student reading comprehension outcomes, especially when graphic organizers are constructed by students and used after reading text. The NICHD (2000) identified instruction using graphic and semantic organizers as having a solid scientific base for improving reading comprehension in nonimpaired readers. The use of graphic organizers is also recommended to assist students with learning disabilities in learning from expository text, which is more information driven and often contains unfamiliar technical vocabulary (Kim, Vaughn, Wanzek, & Wei, 2004). The use of graphic organizers may also help EL students capture the *big ideas* (i.e., concepts or principles that facilitate the most efficient and broadest acquisition of knowledge (Kame'enui & Carnine, 1998), better understand the discipline, and make connections among and between concepts (Deshler & Schumaker, 2005). The use of graphic organizers was found to improve reading-related outcomes for adolescents with reading-related difficulties (Kim et al., 2004; Swanson & Deshler, 2003). Hall and Strangman (2002) reviewed 12 studies investigating the effects of graphic organizer use on comprehension and vocabulary and found that interventions that included a teacher introduction describing the purpose of the graphic organizer as well as the purpose of the text were the most successful.

Use of Peer-Pairing

Collaborative/cooperative learning with heterogeneous groups, along with opportunities for students to engage in extended English discourse, has been

reported as effective instructional features for ELLs (Arreaga-Mayer, 1998; August & Shanahan, 2006a; Genesee et al., 2006; Gersten & Jimenez, 1994). In addition, cooperative learning activities were identified as successful for increasing the reading comprehension of ELLs (August & Shanahan, 2006a). Studies suggest that peer-mediated instruction may enhance the learning of more complex content for students who struggle as well as for all learners in the classroom (Elbaum, Vaughn, Hughes, & Moody, 1999; Gersten et al., 2006; Graham, Harris, MacArthur, & Schwartz, 1991). Slavin, Cheung, Groff, and Lake (2008) describe Peer-Assisted Learning Strategies as a form of collaborative learning where students work in pairs taking turn reading aloud while working on prediction and summarization activities. Although success has been noted with elementary students (Fuchs, Fuchs, Mathes, & Simmons, 1997), Peer-Assisted Learning Strategies has also been implemented with varying degrees of success in remedial and special education upper-elementary and secondary settings. Peer-assistance in the form of the Class Wide Peer Tutoring-model, adopted for our study, provides an interactive and motivating structure for peer instructional interactions that promotes academic and linguistic performance and is reported to be superior to conventional forms of teacher-mediated instruction for accelerating reading comprehension and mastery of other basic academic skills (Greenwood, Arreaga-Mayer, Utley, Gavin, & Terry, 2001; Maheady, Harper, & Malette, 2001). In Class Wide Peer Tutoring, peer tutoring occurs with children of one particular classroom and roles are reciprocal. Tutor dyads benefit from immediate error correction, the chance to engage in both teacher and learner roles, and the opportunity to discuss classroom topics (Greenwood et al., 2001).

Although the research base for effective literacy instruction for adolescent second language learners is only beginning to emerge, there is more than 30 years of reading research that has identified effective instruction which should serve as the foundation for all learners and is considered promising for ELLs (i.e., strategy instruction, direct, explicit teaching of vocabulary and comprehension, use of graphic organizers, active engagement, multiple practice opportunities with corrective feedback, peer-pairing; Biancarosa & Snow, 2006; Francis et al., 2006; Kamil et al., 2008; NICHD, 2000; RAND, 2002). Whereas “good instruction for students in general tends to be good instruction for ELLs in particular” (Goldenberg, 2008, p. 8), application of research on monolingual adolescents to ELLs is relevant because many monolingual English-speaking adolescents share similar struggles with literacy including weaknesses in academic language and vocabulary (Torgesen et al., 2007).

A review of reading research on adolescent struggling readers indicates that reading-related interventions have produced positive outcomes for older students who struggle (Roberts, Torgesen, Boardman, & Scammacca, 2008). Students learning English as a second language, as well as native English speakers, benefit from explicit teaching of the components of literacy (i.e., phonemic/phonological awareness, phonics, fluency, vocabulary,

comprehension, and writing), with fluency, vocabulary, and comprehension being crucial to ELLs' academic success (August & Shanahan, 2006a; Francis et al., 2006). Syntheses focused on research specific to second language learners (August & Shanahan, 2006a; Genesee et al., 2006) confirm that ELLs learn in much the same way as non-ELLs, although readers were advised to ensure that interventions for those learning English as a second language are developmentally and linguistically appropriate, as well as adapted to the proficiency levels of the ELL students. Findings of the National Literacy Panel (August & Shanahan, 2006a) indicate that effective second language instruction includes explicit teaching that helps students directly and explicitly learn features of a second language (i.e., syntax, grammar, vocabulary, pronunciation, and norms of social use, along with multiple opportunities to use the second language in meaningful and motivating ways). These syntheses reveal that extensive teacher knowledge about integrating these instructional practices into teaching routines are a necessary part of effective interventions.

PURPOSE

The purpose of the two studies was to examine the efficacy of incorporating instructional practices associated with improved outcomes into middle-school social studies instruction as a means of enhancing vocabulary knowledge and comprehension for ELLs. Structured pairing was incorporated to improve instruction in the following ways: (a) increasing students' access to and practice with the language associated with content area instruction, (b) providing an interactive and motivating structure for reading and discussing ideas and concepts, and (c) engaging English learners in discussions by providing a scaffold and practice. Vocabulary instruction included (a) selecting words to improve students' academic language, (b) providing students with opportunities to encounter new words in texts and/or video clips, and (c) using graphic organizers to reinforce word meanings and show associations between Spanish and English words. In addition, graphic organizers and written responses were used to assist students in organizing information in meaningful conceptual groupings.

We conducted two experimental studies in two successive school years with nonoverlapping samples. As explained next in the Method section, middle school social studies teachers' classes were randomly assigned to treatment and control conditions. The advantage to this design is that teachers were the same for both of our conditions and students in both the treatment and business as usual conditions covered the same material over the same period using the same textbook providing students in each condition equal access to learning content and key vocabulary. Our primary research question was, How does a multicomponent instructional routine developed to enhance effective outcomes for ELLs and provided by classroom social studies teachers influence students' outcomes in vocabulary and comprehension?

METHOD

We report on the method from two experimental studies using the same treatment during two sequential school years, 2006–2007 and 2007–2008. The participants in each study were specially selected from a similar population base and represent two nonoverlapping samples. We conducted two similar studies to determine replicability of findings before extending the intervention to a larger cohort of schools, teachers, and students. Because the treatment in both studies was the same, we describe it only once.

Participants and Setting: Experiment 1

Student Participants. Participants were drawn from two middle schools in the same central Texas school district. Both schools were considered to have a substantial number of ELLs who were designated by the school as “Limited English Proficient” (LEP). At least 65% of the population at both schools was Latino, with 11.45% of the students at one school designated as LEP and 13.80% of students at the other school as LEP. The proportion of students who qualified for the free or reduced-price lunch program ranged from 70 to 82%.

The seventh-grade students at both middle schools were randomly assigned to 15 sections of seventh-grade social studies classrooms at their school ($N = 381$). The 15 sections were randomly assigned within-teacher to seven treatment ($n = 176$ students) classes and eight comparison classes ($n = 205$ students). Thus, students were randomly assigned to course sections and the sections were randomly assigned to treatments within teacher. Of the original 381 students, 97 (25%) were designated as ELLs (50 in the treatment and 47 in the control condition). Because students may require instructional support even after they no longer have the LEP identification, we included students in the LEP status group if they had had the LEP designation at any point in the previous 3 years. Two LEP students (1 in the treatment group and 1 in the control group) were designated as Monitored LEP, meaning they had transitioned from LEP to non-LEP status within the past 2 years. All other LEP students were currently designated as LEP.

Teachers. All social studies teachers were identified by the school principal and selected to participate in the study. The four participant teachers provided seventh-grade Texas History instruction to all the students in this study. Of the four teachers (two female, two male), two were 1st-year teachers who were certified to teach Social Studies (4–8) in the state of Texas. One of the male teachers had 8 years of experience teaching social studies in secondary school settings. In addition, one of the female teachers was Texas certified as a Generalist (4–8) and self-contained Bilingual/ESL teacher (1–8) and had 6 years of teaching experience. These seventh-grade social studies teachers, with

support from research staff, implemented treatment conditions in intervention classes and continued with their typical instruction in comparison classes.

Participants and Setting: Experiment 2

Students. In the year after the completion of Experiment 1, two middle schools from two districts in central Texas with large numbers of ELLs participated in Experiment 2. Only one of the schools had also participated in Experiment 1. That school's LEP student population grew from 14% in Year 1 to 20% in Year 2. The second school was new to the study and had 51% Latino students and 14% with a LEP status. The percentage of students who were eligible for free or reduced-price lunch was 68% at one school and 85% at the other school.

For Experiment 2, 507 students were randomly assigned to 17 sections of seventh-grade social studies classrooms within the two participating junior high schools. Replicating the Year 1 design, these students were randomly assigned to course sections and the sections were randomly assigned to treatments within teacher. There were 273 students assigned to nine treatment sections and 234 students assigned to eight comparison sections. Of the 507 students, 106 (21%) were ELLs (67 in the treatment and 39 in the control condition). The rest of the students were not identified as ELLs and included 206 in the treatment and 195 in the control condition.

Teachers. During Experiment 2, four teachers were identified by the principals as teaching social studies and participated in the study. All four teachers were male and certified to teach Social Studies (4–8) in the state of Texas. Two of the teachers were 2nd-year teachers, one was a 1st-year teacher, and one had 3 years of teaching experience.

Procedures

Professional Development and Teacher Support. For both experiments, teachers were trained by the researchers to implement the instructional practices only in their treatment classes and were taught to continue with “business as usual” in their control classes. They also received materials and research modeling and support to assist them in delivering the intervention with fidelity to treatment class sections.

To begin, the researchers provided teachers with a 1-day (8 hr long) professional development on implementing the treatment practices. The professional development focused on (a) an overview of the study, (b) a careful description of an experimental study and the importance of adhering to “business as usual” in control classes and implementing instructional practices in treatment classes, (c) critical features of the intervention practices, and (d) the lesson framework

and how the researchers expected it to be implemented. Each teacher was provided with all necessary materials to implement the treatment including lesson plans, overheads, and videos. Texas history lesson plans that were based on their school's scope and sequence, as well as state standards, were provided teachers to improve fidelity and to provide a framework for how the instructional practices would be embedded in their content instruction. Throughout the professional development, teachers had opportunities for application and questions.

In addition, the research team provided in-class support and coaching. One research support person was assigned to each participating teacher and was initially present daily in the teacher's treatment classes (2 weeks) and then less so over the course of the intervention (once a week or as needed). On the 1st day that the treatment condition was implemented, the research support person together with the teacher explained the purpose and overview of the study to students and demonstrated how students were expected to work in pairs. During the 1st week of the treatment, the research support person also conducted demonstration lessons to model the lesson framework for the teacher in the classroom. Teachers participated in brief visits with their research support person during the weeks of implementation. They were provided feedback about their instruction based on observations and they were able to ask questions, problem-solve, and plan for lessons with their research support. Furthermore, teachers could request onsite modeling of lesson components throughout the duration of the intervention depending on their needs.

Description of the Treatment Intervention. The same treatment intervention was used by teachers in all treatment classes for both experimental studies. The treatment intervention was composed of (a) overview and vocabulary instruction, (b) the use of brief videos and purposeful discussion to build concepts, (c) the use of graphic organizers and other writing activities to build comprehension and vocabulary through writing, and (d) structure paired grouping. Students in the treatment classes received the intervention during their regularly scheduled seventh-grade social studies class. The intervention was implemented for 50 min a day, 5 days a week for approximately 9 to 12 weeks. The number of lessons was the same across teachers and studies, but interruptions in school schedule extended the number of weeks it took to complete the intervention. The researcher-designed lessons were used by teachers and included all the aforementioned intervention components.

Typical instruction was provided for students who were randomly assigned to control sections in seventh-grade social studies. Typical instruction included teachers reading selected passages from the textbook, instruction on key ideas, using videos to illustrate ideas (usually longer than those in treatment, e.g., 20–50 min once or twice per week).

One of the critical elements of the design that we believe makes the findings compelling is that students in both the treatment and business as usual

conditions covered the same material over the same period using the same textbook. Thus all students had opportunities to learn the words and content taught and tested.

Treatment Intervention Lessons. Lesson plans identified the core subject matter and the “big ideas” that the students needed to learn in their social studies course as well as guided the teachers on the use of specific instructional practices to convey the subject matter. These practices were designed to enhance students’ understanding of social studies content and of expository text by giving all their students opportunities to learn and use the vocabulary, concepts, big ideas, and issues associated with social studies. The lesson plans were not meant to be a script for teachers, but rather a guide for how to build in the study’s strategies and materials.

The unit lessons were designed around one or two central ideas that served as organizing concepts to help the teacher focus the events and ideas in each unit. Every lesson was organized similarly to encourage the teacher to develop a routine for the intervention. The daily instructional routine consisted of the following exercises: (a) a brief overview of the “big idea,” (b) explicit vocabulary instruction that integrated paired students’ discussion of the word, (c) discussion built around a short video clip (2–4 min) that complemented the day’s reading, (d) a teacher-led or paired student reading assignment followed by generating and answering questions, and (e) a wrap-up activity in the form of a graphic organizer or other writing exercise. A description of each of the key components of the lessons follows (sample lessons are available on our Web site: <http://www.texasreading.org>).

Overview and Vocabulary Instruction. A typical lesson began with teachers giving students an overview of the day’s lesson by connecting it with information that had been previously taught. The overview was also used to revisit and build on the unit’s big idea, which was an organizing concept used throughout the week’s unit. Following the overview, teachers explicitly taught four new preselected vocabulary words that were drawn from readings, video clips, and a teacher’s scope and sequence. For each vocabulary term, teachers pronounced the word, identified a Spanish cognate or Spanish translation, gave a student friendly definition, and used a visual representation to help students understand the meaning of the word with language that was clear to them. Furthermore, the teachers shared two sentences using each vocabulary word. One sentence used the word in a historical context and was taken from class text, whereas the second sentence used the word in a clarifying sentence that was relevant to students’ experiences. Finally, after each word was defined and discussed, students were given an opportunity to use the word and apply its meaning by discussing a prompt with their student partner.

Strategic Use of Video. When a brief video clip was available to accompany a reading, teachers used it to help students develop their understanding of the lesson's big idea. The teacher introduced the video clip either before or after students read a passage by previewing what students were going to watch in the video. Teachers had students focus on one or two key questions prior to video use. After students watched the video they contributed their answers to the focus question(s) as the teacher led a brief discussion.

Teacher-Led or Paired Student Reading. The next step in every lesson was to do a read aloud of a carefully selected reading. Before reading, teachers had students think about two or three questions that typically asked them to focus on the most important ideas of the lesson and that they would answer after reading. Teachers and students alternated leading the reading. When the teachers read aloud, they modeled fluent reading while clarifying vocabulary and periodically checking for students' comprehension. When students read aloud they took turns reading. As one student read, their partner was asked to follow along and give corrective feedback as needed. After reading, students worked on answering the questions in their pairs. Teachers then discussed students' answers to the focus questions with the whole class.

Use of Writing With Graphic Organizers. Finally, to wrap up a lesson, students were asked to work with their partners to complete a graphic organizer or some other brief writing activity. Teachers introduced the graphic organizer or activity and explained the information that students were summarizing from their lesson. Students were then asked to write the most important information in their organizers. After students were given an opportunity to write, their teacher reviewed student responses and gave them feedback.

Structured Paired Grouping. Paired instruction procedures were included in intervention lessons to help teachers follow a set of routines for paired reading, paired writing as well as paired vocabulary discussion. Within each intervention classroom, all students were paired according to language ability. Pairing of students was based on LEP status and their sixth-grade reading scores on the Texas Assessment of Knowledge and Skills (state high-stakes test). Each of the intervention classes was first divided into LEP and non-LEP students. The LEP group included students in the monitored LEP category. Students in each category were rank ordered according to their reading test scores from highest to lowest. They were then assigned a partner, pairing the highest scoring LEP student with the highest scoring non-LEP student and proceeding through the list until all students had a partner. In a few instances, LEP students outscored their non-LEP counterparts. These students were paired with the highest scoring non-LEP student in the class and adjustments were made for the rest of the students to follow the pairing procedures as closely as possible. Students were paired in this way to ensure that when partners worked on activities and

discussed ideas they would have another student with similar language background who was also able to challenge and support them with their language development. Students worked with their partners for approximately 12 to 20 min during the 50-min lesson.

Observation of Treatment and Comparison Sections

Throughout the implementation of both treatments, four researchers conducted observations in treatment and control (business as usual) sections to ensure fidelity of treatment and to determine if there was any contamination of instruction in the control classes. Implementation fidelity was monitored by three observations over the course of the treatment intervention. In addition, control sections were observed four times to ensure that teachers were not implementing the intervention in control sections. The fidelity measure was developed to correspond with the critical elements of the treatment and consisted of the following five categories: (a) vocabulary instruction and review, (b) partner reading, (c) strategic use of video, (d) graphic organizers/writing, and (e) structured paired grouping. Each section consisted of two components: quality of instruction and organization. For quality of instruction, observers assigned a 0 to 2 or “not applicable” rating to indicate if the teacher implemented the activity required with sufficient instruction. Not applicable was coded when the element of treatment, for example, graphic organizer, was not part of the designated lesson. For organization, the observers noted how often teachers followed their lesson plan, if they used the necessary materials and if students had access to the required materials, and if the teacher facilitated student partnering. Organization items were assigned values from 1 to 4 with 1 representing no implementation at all, and 4 representing implementation “most of the time.”

One of the issues we were interested in determining is whether there was contamination between treatment and control sections. Observations and teacher reports confirmed that materials and practices designed for treatment classes were not used in the control classes.

In addition, observers rated classroom management using items assessing the degree to which the teachers redirected off-task student behavior, included a variety of students in class discussions, provided positive feedback, and organized students into pairs in an efficient manner. Finally, the observers also provided a global rating of the teacher’s overall instruction on a 3-point scale from *less than average* to *above average*.

Scores were calculated by taking the average ratings of each teacher by class type (treatment or control) across the quality and organization portions of the fidelity measure, as well as across the observation time points. Teachers were not observed using treatment practices during control classes and did not implement any portion of the lessons designed for the treatment group, thus differential instruction for treatment students and control students with

Table 1. Average observation category ratings by group and experiment

Observation Rating Category	Experiment 1				Experiment 2			
	Treatment		Control		Treatment		Control	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Quality	1.39	0.23	1.28	0.27	1.39	0.51	1.71	0.38
Organization	3.37	0.32	2.44	0.39	3.60	0.50	3.84	0.14
Classroom management	3.43	0.19	3.17	0.53	3.20	0.83	3.40	0.68
Global rating of instruction	1.81	0.26	1.54	0.43	1.90	0.74	1.88	0.64

respect to the target instructional practices was achieved. The overall quality and organization ratings for control classes were based primarily on ratings of the “Other” section of the fidelity measures. Classroom management and the overall global ratings were present for all teachers and are presented separately. In total, there are four derived fidelity scores—average quality, organization, classroom management, and global ratings of instruction.

Table 1 presents the average scores from the fidelity observation for the treatment and control groups in Studies 1 and 2. Analyses of group differences on these fidelity ratings indicated a significant difference on ratings of organization in Study 1 between treatment and control sections ($F = 24.78, p \leq .01$). In treatment sections in Study 1, on average, teachers were rated as implementing more aspects of classroom management than in the control sections. No significant differences in average ratings were found between treatment and control sections in Study 2. Overall, in both studies, the teachers were rated fairly similarly during treatment and control instruction.

Individual ratings of teachers during instruction in treatment sections ranged from 0.50 to 1.94 on quality, 2.45 to 3.96 on organization, 1.80 to 3.83 on classroom management, and 1.00 to 2.67 on global ratings of instruction. Ratings during instruction in control sections ranged from 1.00 to 2.00 on quality, 2.05 to 4.00 on organization, 2.50 to 4.00 on classroom management, and 1.00 to 2.67 on global ratings of instruction.

Outcome Measures

Prior to the intervention and after its completion all students were assessed with a researcher-developed content-based measure. The measure was designed to cover students’ understanding of the content taught during a 9- to 12-week period and was meant to serve as an indicator of growth in social studies learning. It resembled traditional assessments of content area classes in that it consisted of vocabulary matching items and comprehension questions. The items were

developed based on content in textbook and weekly quizzes. Students in both treatment and controls covered this same content. The vocabulary section had 20 items that included definitions that had to be matched with vocabulary terms used within the context of a sentence that contained social studies information. For example, the definition, to officially give up power or territory, had to be matched to the target word *cede* in the sentence Mexico agreed to *cede* much of its northern territory to the United States. The second part of the assessment included 10 questions asking students to identify and explain the big ideas of the social studies units taught during the instruction. For example, one comprehension item required students to explain two ways in which slaves' human rights were violated. The content represented in these big ideas was part of the instructional materials and state standards and thus the content was part of the instructional materials for both treatment and control students. Analyses of pre- and postvocabulary and comprehension performance were examined separately for each study. The first step of the analyses examined differences in pretest scores as a function of group (treatment or control). The second step examined group differences in postintervention performance as a function of treatment group controlling for pretest measures of the outcome variable. All analyses were conducted using three-level, hierarchical linear models in HLM 6.06 (Raudenbush, Bryk, & Congdon, 2008). For all HLM analyses, we report results for fixed effects of treatment based on robust standard errors. The three-level models included variability because of students within section, sections within teachers, and teachers. In all models, treatment was entered at Level 2 (i.e., at the section level). We also tested for heterogeneity of regression between Treatment and Control sections and examined models that allowed for pretest regressions that varied by teacher, as well as the possibility that treatment effects differed across teachers. However, because of the small number of teachers in the study, we focus here on results from models where pretest regressions and treatment effects were constrained to be fixed across teachers. Thus, in all reported models, random effects due to teachers were limited to effects on the intercept, that is, the average value across all sections, both treatment and control, for that teacher.

RESULTS

Results are reported separately for Study 1 and 2.

Results for Study 1

Table 2 presents the means and standard deviations for vocabulary and comprehension scores, pre- and postintervention, for Study 1. Table 2 is organized so as to show results separately for ELL and non-ELL students. Sample size

Table 2. Means and standard deviations for Study 1

Measure	Group	No. of Sections	No. of Students (Pre/Post)	Pretest		Posttest	
				M	SD	M	SD
Comprehension	Control	8	136 / 137	1.63	1.21	2.33	1.75
	ELL		46 / 45	1.17	1.15	1.93	1.73
Vocabulary	Treatment	7	114 / 105	1.48	1.31	3.73	2.14
	ELL		50 / 47	1.18	1.17	3.32	2.12
Vocabulary	Control	8	136 / 137	8.68	4.98	10.49	5.05
	ELL		46 / 45	6.54	4.33	7.27	5.28
Vocabulary	Treatment	7	114 / 105	9.13	4.85	12.90	4.97
	ELL		50 / 47	6.88	3.68	10.57	5.11

Note. ELL = English language learner.

information is provided separately for pre- and postassessments. Not surprisingly, ELL and non-ELL students differed at the pretest, although these differences were comparable for Treatment and Control sections. More important, pretest scores were not different between Treatment and Control sections. A three-level HLM analysis of pretest scores showed no differences between Treatment and Control sections for either Comprehension, $t(13) = -0.970$, $p = .350$, or Vocabulary, $t(13) = 0.552$, $p = .590$.

Analysis of posttest scores using a three-level analysis of covariance revealed statistically significant differences between students in Treatment and Control sections for both Comprehension and Vocabulary. Fixed and random effects for the three-level analysis of covariance model for Study 1 are presented in Table 3 for both outcome measures. For comprehension there was a significant effect for treatment condition indicating that students in treatment sections were performing at significantly higher levels than students in control sections postintervention, $t(13) = 14.31$, $p \leq .001$. The estimated difference in comprehension scores between treatment and control sections was 1.57 with a standard error of 0.109. Similar results were found for vocabulary, although the absolute difference between treatment and control groups was found to be somewhat larger (Treatment Effect Estimate = 2.53, $SE = 0.629$), $t(13) = 4.026$, $p = .002$. For both vocabulary and comprehension, treatment effects did not interact with student status as ELL or non-ELL indicating that ELL and non-ELL students benefitted equally from participation in treatment sections. It is instructive to note in Table 2 that posttest means for ELL students in treatment sections were at least as large or larger, in an absolute sense, than means for non-ELL students in the comparison condition for both vocabulary and comprehension.

Translating these differences into effect sizes shows that the effects of the intervention were large for both vocabulary and comprehension. For comprehension, we estimated the effect size to be $g = 1.12$ based on the data presented in Table 2. For vocabulary, we estimated $g = 0.53$. We illustrate the computation of g for the comprehension data in Table 2; computations for vocabulary are directly analogous. To compute g , we computed the difference in the overall posttest mean for treatment and control (Treatment $M = (105 \times 3.73 + 47 \times 3.32)/(105 + 47) = 3.61$; Control $M = (137 \times 2.33 + 45 \times 1.93)/(137 + 45) = 2.23$; difference = 1.38) and divided this difference by the pooled within-groups standard deviation at the pretest ($s_{\text{pooled}} = (136-1) \times 1.21^2 + (46-1) \times 1.15^2 + (114-1) \times 1.31^2 + (50-1) \times 1.18^2 / (136 + 46 + 114 + 50-4)^{1/2} = 1.23$). If instead one uses the pooled within-group standard deviation for the posttest of 1.93, then $g = .71$. Both of these effect sizes are estimates of the unadjusted treatment effect divided by the total, pooled within-groups standard deviation, which includes variability due to students within sections, variability due to sections, and variability due to teachers. If, instead, we base our estimate of effect size on the treatment difference adjusted for the pretest covariate (Treatment Effect estimate from three-level HLM model for comprehension of 1.57), then

Table 3. Fixed and random effects for models of vocabulary and comprehension posttests in Experiment 1

Fixed Effects						
CBM Measure	Predictor	Estimate	SE	<i>t</i>	<i>df</i>	<i>p</i>
Comprehension	Intercept	2.28	0.385	5.93	3	<.001
	Pretest	0.661	0.063	10.52	306	<.001
	Treatment	1.57	0.109	14.31	13	<.001
Vocabulary	Intercept	9.75	0.856	11.39	3	<.001
	Pretest	0.747	0.009	86.49	306	<.001
	Treatment	2.53	0.629	4.026	13	.002
Random Effects						
CBM Measure	Source		Variance	χ^2	<i>df</i>	<i>p</i>
Comprehension	Students within section		2.654			
	Sections within teacher		0.001	9.67	10	>.500
	Teachers		0.513	70.95	3	<.001
Vocabulary	Students within section		12.675			
	Sections within teachers		0.993	29.54	10	.001
	Teachers		0.810	12.21	3	.007

Note. Tests of differences between treatment and control conditions on pretest measures of vocabulary and comprehension were not statistically significant. (See text for details on these tests.)

the estimates of g for comprehension are 1.28 and .81, respectively, depending on whether the pretest or posttest standard deviation as just computed is used in the denominator. These effect size calculations coincide with the standards adopted by the What Works Clearinghouse in their Standards Volume 2, dated December 2008, which recommends use of the pooled within-group standard deviation at the posttest as the measure of standard deviation unless there is evidence that the treatment affects variability in the posttest. For vocabulary, estimates of g for the unadjusted means were .53 and .49 based on the pretest and posttest standard deviations, respectively, whereas the corresponding estimates of g for the analysis of covariance adjusted means from the HLM model were .54 and .50. We have not applied Hedges' correction for small sample bias because the adjustment factor computes to .998 in this instance, which would result in minimal adjustment to the estimates.

Results for Study 2

Table 4 presents the means and standard deviations for vocabulary and comprehension scores pre- and postintervention, for Study 2. Analysis of pretest performance scores indicated no significant differences between treatment and control sections for either vocabulary, $t(15) = 1.450$, $p = .168$, or comprehension, $t(15) = 0.934$, $p = .366$. Similar to Study 1, students performed more poorly on the comprehension measure at both time points than on the vocabulary assessment. However, one must be cautious in interpreting these differences between domains as the assessments were not developed in a manner that would guarantee that they are equivalent in difficulty. Results from the three-level HLM analysis with student pretest scores as a covariate yielded results highly similar to those from Study 1. As in Study 1, performance at the pretest significantly predicted student performance at the posttest for both vocabulary ($\beta = 0.515$, $SE = 0.107$), $t(363) = 4.82$, $p < .001$, and comprehension ($\beta = 0.596$, $SE = 0.030$), $t(361) = 19.46$, $p < .001$. More important, students in treatment sections again outperformed those in control sections on both the comprehension and vocabulary measures (Comprehension: $\beta = 1.09$, $SE = 0.403$), $t(15) = 2.71$, $p = .016$; Vocabulary: $\beta = 1.94$, $SE = 0.550$), $t(15) = 3.53$, $p = .003$. As in Study 1, treatment effects did not interact with students' ELL status for either outcome, indicating that both ELL and non-ELL students benefited equally from being assigned to sections that were randomly assigned to the treatment condition. Expressing the treatment-control differences as effect sizes using g as in Study 1, we find that effect sizes are overall somewhat smaller, but again are classified as large or moderate, and are somewhat larger for comprehension than for vocabulary. Effect sizes for unadjusted posttest means (see Table 4) were .989 and .468 for comprehension, and .452 and .355 for vocabulary, based on the pooled within groups pretest and posttest standard deviations, respectively. Analogous effect sizes for the adjusted posttest

Table 4. Means and standard deviations for Study 2

Measure	Group	No. of Sections	No. of Students (Pre/Post)	Pretest		Posttest	
				<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Comprehension	Control	8	154 / 175	0.72	1.00	2.23	2.15
	ELL		39 / 39	0.49	0.96	1.36	1.32
Treatment	Non-ELL	9	155 / 173	0.83	1.20	3.11	2.45
	ELL		66 / 66	0.80	0.99	3.18	2.54
Vocabulary	Control	8	154 / 175	7.37	4.97	9.86	5.75
	ELL		39 / 39	5.00	3.24	9.47	5.20
Treatment	Non-ELL	9	154 / 172	7.71	4.79	11.78	6.42
	ELL		67 / 67	7.61	4.58	12.25	5.77

means were 1.01 and .480 for comprehension, and .413 and .325 for vocabulary depending on whether the pooled within groups pretest or posttest standard deviation in Table 4 was used in the denominator of g .

The test of homogeneity of pretest regression between treatment and control sections was not statistically significant for either vocabulary or comprehension, justifying the use of a common regression slope in predicting posttest scores. In the case of vocabulary, there was some suggestion that pretest regressions might differ across teachers and that treatment effects might differ across teachers. However, because significance tests for random effects may be misleading when the number of sampling units is small (in this case $n_{\text{teachers}} = 4$), we have focused on the average treatment difference between treatment and control sections, averaged across teachers from the three-level HLM model that constrained the regression effect for the pretest to be the same for all teachers (see Table 5).

DISCUSSION

Many ELLs, like those in this study, attend schools with high percentages of students living in poverty. These schools are less likely than other schools to have adequate funds and resources and to provide appropriate support for academic achievement (Donovan & Cross, 2002). Students in these schools may also have fewer opportunities to engage in academic discussions, to be exposed to rich content instruction, and to have good language models among their peers. This study was designed to provide teachers instructional routines that would improve academic outcomes for ELLs by focusing on improved vocabulary and content knowledge through engaging students in interactions during their social studies classes. Findings from these two experimental studies indicate that students improved in their word knowledge and their comprehension related to the content area of instruction—social studies.

The multicomponent intervention used in this study incorporated practices that have been found to be effective in the development of vocabulary and comprehension of adolescents. However, the efficacy of these practices to improve vocabulary and comprehension outcomes, combined or in isolation, with adolescent ELLs had not been explored. Of the four practices used, three focused on enhancing content instruction and the fourth, peer-pairing, altered the interaction patterns typically used in the classroom. The combination of explicit vocabulary and concept instruction with a conceptual instructional focus shifted the instructional emphasis from the acquisition of historical facts to one in which the *big ideas* provided context for promoting students' using language and understanding the content. Further, the use of structured discussions and graphic organizers by pairs of students provided additional support to students in their use of expressive language. Lessons were explicitly designed to provide students with activities to orally and in writing express their understanding of

Table 5. Fixed and random effects for models of vocabulary and comprehension posttests in Experiment 2

Fixed effects						
CBM Measure	Predictor	Estimate	SE	<i>t</i>	<i>df</i>	<i>p</i>
Comprehension	Intercept	2.18	0.431	5.05	3	.008
	Pretest	0.596	0.030	19.46	361	<.001
	Treatment	1.09	0.403	2.71	15	.016
Vocabulary	Intercept	10.10	0.771	13.10	3	<.001
	Pretest	0.515	0.107	4.82	363	<.001
	Treatment	1.94	0.550	3.53	15	<.003
Random Effects						
CBM Measure	Variance Component	Estimate	Estimate	χ^2	<i>df</i>	<i>p</i>
Comprehension	Students within-section	3.586				
	Sections within teacher	0.275		35.29	12	.001
	Teachers	0.900		39.00	3	<.001
Vocabulary	Students within-section	22.056				
	Sections within teacher	2.269		40.266	12	<.001
	Teachers	3.679		23.61	3	<.001

Note. Tests of differences between treatment and control conditions on pretest measures of Vocabulary and Comprehension were not statistically significant. (See text for details on these tests.)

the concepts and to use the language associated with the topic. It may be that the additional opportunities to engage with the target vocabulary may have had a positive effect on student retention of the vocabulary and to a lesser degree, as measured by the comprehension measures, their learning of the content.

In middle school, ELLs' academic success depends on their ability to acquire the content and vocabulary knowledge associated with each of the content areas. The findings from these two studies provide educators practices that are effective in improving ELLs' vocabulary. The components used in this intervention can be easily adapted for use in other content areas such as science, health, geography, and economics because it does not require a change in the curriculum. It does, however, require a shift in the focus of instruction to an emphasis on the big ideas, attention to vocabulary and background knowledge development, and altering interaction patterns in the classroom between teacher and students and between students.

Each of the components in this intervention addressed a different instructional challenge. In addition to the general instructional effectiveness of the components, research with ELLs indicates that the objective met with each of the components has been identified as important in improving the academic outcomes of ELLs (August & Hakuta, 1997). For example, the use of structured pairs increased students' opportunity for active engagement whereas the explicit and deep teaching of vocabulary and the use of video was included to address ELLs' limited content area vocabulary and background knowledge. Thus, each component had a unique and, we think, necessary role in the intervention. However, given the multicomponent nature of the intervention it is not possible to determine whether any one of these components had a greater effect on student learning or whether each was necessary. In the future, it would be beneficial to explore the relative benefit of the various components.

Another critical finding from this study is that ELLs who participated in the intervention condition in either study benefited from the instruction they received. They outperformed the ELLs in the comparison group on the researcher-developed vocabulary and comprehension measures. ELLs who participated in the intervention gained words at the same rate as students who were not limited English proficient.

Although this intervention was developed to address the instructional and language needs of ELLs, the students who were not LEP in the intervention classes also benefited. Students who were LEP outperformed their counterparts in the comparison condition on both the vocabulary and comprehension measures. When both the target group (ELLs) and their classmates benefit from an intervention or practice, it meets the criteria for universal design. This finding is particularly relevant for teachers who have both ELLs and non-LEP students in their classrooms and who may be concerned about the possible detrimental effect for other students of instruction that targets ELLs. If effective instructional practices for ELLs also benefit non-ELLs, teachers have a strong rationale for implementing the instructional practice. Furthermore, ELLs in the comparison

condition made the least gains and lagged behind all other groups on both the vocabulary and content comprehension measures, providing further support for interventions such as the one in this study to alter the course for ELLs.

We recognize that this finding may be limited to schools and students with similar demographics as those in this study. Many students in these schools were performing below grade level and may have also needed the additional practice and interaction provided by the intervention. In a school in which more students are performing at or above grade level, universal growth across groups of students may not be evident.

Teacher factors that may impact the effect of an intervention include number of years teaching, teacher preparation, and teacher self-efficacy. To control for the many school and teacher effects, we randomly assigned each class section within teacher to either treatment or comparison conditions. A concern with this type of design is that teachers will implement the intervention in the comparison classrooms. To ensure that this did not happen, we carefully trained teachers in the importance of adherence to condition and the value of participating in an experimental study that would allow for causal inferences. To confirm that teachers would adhere to treatment and comparison condition, we observed both treatment and comparison classes and based on both these observations and teacher feedback, we are confident that there was minimal or no use of treatment practices during the comparison condition. However, if they had, we would have expected to see smaller differences between the treatment and comparison groups because students in both conditions would have received the intervention or conversely neither would have. The findings from these two studies suggest that differences in students' outcomes are attributable to the treatment.

The ELLs in this study were provided instruction that engaged them in interactive instructional routines that provided them with opportunities to use the words orally and in writing resulting in increases in their vocabulary and understanding of social studies. Although the growth from this treatment is modest, it does have implications for instruction in general. If students were to receive this instruction beginning in kindergarten, we would expect cumulative impact. For example, if students learned the meaning of the word *reservation* in fourth grade instead of seventh, then vocabulary instruction at seventh grade could focus on more advanced vocabulary.

Limitations

Students who participated in the treatment improved significantly on vocabulary and comprehension, but there are two issues to consider. First, like most vocabulary measures, the one we used to assess vocabulary did not require students to demonstrate use and deep understanding of the words. Students matched a sentence with the target word to the definition they chose from a

list of definitions. This type of response does not determine whether students can produce a definition or whether they can use the word when speaking or writing. Second, students did not generalize this knowledge to general vocabulary development as measured by a standardized vocabulary assessment. We think it is reasonable to assume that a more extensive treatment addressing academic vocabulary across content areas and contexts would demonstrate a greater impact on standardized measures. For this reason, our subsequent study will integrate the vocabulary and comprehension instruction across content areas including science and math.

Students' improvement on the comprehension measure was statistically significant. However, the correct number of responses increased from less than one at pretest to an average of three to four. Unlike the vocabulary measure, the comprehension measure may have underrepresented students' comprehension of the content taught because it required that they provide written responses. In addition to the lack of vocabulary, difficulty spelling and composing sentences may account for the apparent lack of representation of the content taught. A format with options may have provided a better indication of students understanding of the big ideas. One consideration in designing measures to assess ELLs is the purpose. If the goal is content area knowledge, then the assessment should provide students the accommodations needed to demonstrate they understand concepts even if their language skills are limited (Francis et al., 2007). However, if in addition to content knowledge, one is interested in how well students can use language to communicate that understanding, then an open-ended measure would be more appropriate.

Finally, significant differences were found on curriculum-based measures. The long-range implications of these findings such as whether students maintained their comprehension knowledge and vocabulary were not addressed. The finding from this study are consistent with conclusions reached in a recent meta-analysis (Scammacca et al., 2007) that reported positive outcomes for monolingual students from vocabulary interventions in which gains on vocabulary were tested on the words students were taught. The authors also note that few studies use standardized measures of vocabulary and that there are few standardized measures that measure vocabulary adequately. Although there is a link between vocabulary and comprehension (Beck et al., 2002; Biemiller, 1999; NICHD, 2000; RAND, 2002), the positive relationship between vocabulary and comprehension seems to be limited to research-developed, near measures with a weaker relationship with standardized measures (Elleman, Lindo, Morphy, & Compton, 2009).

There were limitations to some of the teachers' implementation of the intervention. Many of the teachers across the two studies needed support to make appropriate modifications to meet the language needs of students. When students exhibited confusion or failed to answer a question, the teachers just repeated the same prompt or moved to another student or pair. A second factor was changing the interaction pattern between the teacher and the students.

Based on the observations in the comparison sections, after asking a question, teachers often provided the answer. Students often just waited until the answer was provided to complete the worksheet or other graphic organizer. Our experiences indicate that some teachers may require considerable coaching in addition to initial professional development to implement interactive instructional practices.

In summary, students in the treatment condition had a better grasp of the vocabulary associated with the content compared to the students in the comparison condition. Although scores on the comprehension questions were higher for the treatment students, there was still room for growth in demonstrating students' understanding of what they were learning. In terms of the delivery of instruction, the use of a different interaction pattern requiring student pairing was challenging for both the students and the teachers. More time needs to be dedicated to teaching students to work effectively in pairs (Stenhoff & Lignugaris/Kraft, 2007) and for teachers to learn appropriate instructional modifications that facilitate learning for students who are ELLs.

REFERENCES

- Arreaga-Mayer, C. (1998). Language sensitive peer-mediated instruction for culturally and linguistically diverse learners in the intermediate elementary grades. In R. M. Gersten & R. T. Jimenez (Eds.), *Promoting learning for culturally and linguistically diverse students* (pp. 73–90). Belmont, CA: Wadsworth.
- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*, 50–57.
- August, D., & Hakuta, K. (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Academy Press.
- August, D. L., & Shanahan, T. (2006a). *Developing literacy in second-language learners: A report of the national literacy panel on language minority children and youth*. Mahwah, NJ: Erlbaum.
- August, D. L., & Shanahan, T. (2006b). Synthesis: Instruction and professional development. In D. L. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 351–364). New York: Routledge.
- Baumann, J. F., Edwards, E. C., Boland, E., M. Olejnik, S., & Kame'enui, E. J. (2003). Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meanings. *American Educational Research Journal, 40*, 447–494.
- Baumann, J. F., Kame'enui, E. J., & Ash, G. E. (2003). Research on vocabulary instruction: Voltaire redux. In J. Flood, D. Lapp, J. R. Squire, & J. M. Jensen (Eds.), *Handbook of research on teaching the English language arts* (2nd ed., pp. 752–785). Mahwah, NJ: Erlbaum.
- Beck, I. L., & McKeown, M. G. (2001). Text talk: Capturing the benefits of read-aloud experiences for young children. *The Reading Teacher, 55*, 10–20.

- Beck, I., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York: Guilford.
- Biancarosa, G., & Snow, C. E. (2006). *Reading next—A vision for action and research in middle and high school literacy: A report from the Carnegie Corporation of New York* (2nd ed). Washington, DC: Alliance for Excellence in Education.
- Biemiller, A. (1999). *Language and reading success*. Cambridge, MA: Brookline.
- Bos, C. S., & Anders, P. L. (1990). Effects of interactive vocabulary instruction on the vocabulary learning and reading comprehension of junior-high learning disabled students. *Learning Disability Quarterly*, 13, 31–42.
- Cognition and Technology Group at Vanderbilt. (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher*, 19, 2–10.
- Deshler, D. D., & Schumaker, J. B. (2005). *Teaching adolescents with disabilities: Accessing the general education curriculum*. Thousand Oaks, CA: Corwin.
- Donovan, M. S., & Cross, C. T. (2002). *Minority students in special and gifted education* (National Research Council, Committee on Minority Representation in Special Education, Division for Behavioral and Social Sciences and Education). Washington, DC: National Academies Press.
- Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (1999). Grouping practices and reading outcomes for students with disabilities. *Exceptional Children*, 65, 399–415.
- Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, 2, 1–44.
- Francis, D. J., August, D. L., Snow, C. E., Vaughn, S., Linan-Thompson, S., Hiebert, E. H., et al (2007). *Oracy/literacy development in Spanish-speaking children* (Developed with funds provided by the National Institute of Child Health and Human Development, #HD-39-521, Development of literacy in Spanish Speaking Children).
- Francis, D. J., & Rivera, M. O. (2007). Principles underlying English language proficiency tests and academic accountability for ELLs. In J. Abedi (Ed.), *English language proficiency assessment in the nation: Current status and future practice* (pp. 13–32). Davis: University of California.
- Francis, D. J., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: Research-based recommendations for instruction and academic interventions* (Under cooperative agreement grant S283B050034 for U.S. Department of Education). Portsmouth, NH: RMC Corporation, Center on Instruction, Center on Instruction. Available from <http://www.centeroninstruction.org/files/ELL1-interventions.pdf>
- Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34, 174–206.
- Genesee, F., Lindholm-Leary, K., Saunders, W. M., & Christian, D. (2006). *Educating English language learners: A synthesis of research evidence*. New York: Cambridge University Press.
- Gersten, R., Baker, S. K., Smith-Johnson, J., Dimino, J., & Peterson, A. (2006). Eye on the prize: Teaching complex historical content to middle school students with learning disabilities. *Exceptional Children*, 72, 264–280.

- Gersten, R., & Jimenez, R. T. (1994). A delicate balance: Enhancing literature instruction for students of English as a second language. *The Reading Teacher*, 47, 438–440.
- Goldenberg, C. (2008). Teaching English language learners: What the research does—and does not—say. *American Educator*, 32, 7-23, 42–44.
- Graham, S., Harris, K. R., MacArthur, C. A., & Schwartz, S. (1991). Writing and writing instruction for students with learning disabilities: Review of a research program. *Learning Disabilities Quarterly*, 14, 89–114.
- Greenwood, C. R., Arreaga-Mayer, C., Utley, C. A., Gavin, K. A., & Terry, B. J. (2001). ClassWide peer tutoring learning management system: Applications with elementary-level English language learners. *Remedial and Special Education*, 22, 34–47.
- Hall, T., & Strangman, N. (2002). *Graphic organizers*. Wakefield, MA: National Center on Accessing the General Curriculum. Retrieved November 7, 2008, from www.cast.org/publications/ncac/ncac_go.html
- Kame'enui, E. J., & Carnine, D. W. (Eds.). (1998). *Effective teaching strategies that accommodate diverse learners*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices* (No. NCEE 2008–4027). Washington, DC: U.S. Department of Education. Retrieved from <http://ies.ed.gov/ncee/www>
- Kim, A.-H., Vaughn, S., Wanzek, J., & Wei, S. (2004). Graphic organizers and their effects on the reading comprehension of students with LD: A synthesis of research. *Journal of Learning Disabilities*, 37, 105–118.
- Kindler, A. L. (2002). *Survey of the states' limited English proficient students and available educational programs and services 2000–2001 summary report* (Office of English Language Acquisition, Language Enhancement, and Academic Achievement for Limited English Proficient Students). Washington, DC: National Clearinghouse for English Language Acquisition & Language Instruction Education Programs. Available from http://www.ncela.gwu.edu/stats/2_nationl.htm
- Laird, J., DeBell, M., & Chapman, C. (2006). *Dropout rates in the United States: 2004* (NCES 2007–024). Washington, DC: National Center for Education Statistics, U.S. Department of Education. Retrieved January 27, 2009, from <http://nces.ed.gov/pubsearch>
- Lee, J., Grigg, W., & Donahue, P. (2007). *The Nation's report card: Reading 2007* (NCES 2007–496). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Maheady, L., Harper, G. F., & Mallette, B. (2001). Peer-mediated instruction and interventions and students with mild disabilities. *Remedial and Special Education*, 22, 4–14.
- National Center for Education Statistics. (2008). *The condition of education 2008* (NCES 1008–031). Washington, DC: U.S. Government Printing Office.
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00–4769). Washington, DC: U.S. Government Printing Office.

- RAND Reading Study Group. (2002). *Reading for understanding: Toward a research and development program in reading comprehension*. Santa Monica, CA: RAND Corporation.
- Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2008). *HLM 6* (Version 6.06). Lincolnwood, IL: Scientific Software International.
- 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*, 63–69.
- Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., et al. (2007). *Interventions for adolescent struggling readers: A meta-analysis with implications for practice*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Short, D., & Fitzsimmons, S. (2007). *Double the work: Challenges and solutions to acquiring language and academic literacy for adolescent English language learners—A report to the Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.
- Slavin, R., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A Best-evidence synthesis. *Reading Research Quarterly, 43*, 290–322.
- Stahl, S. (1999). *Vocabulary development*. Cambridge, MA: Brookline Books.
- Stenhoff, D. M., & Lignugaris/Kraft, B. (2007). A review of the effects of peer tutoring on students with mild disabilities in secondary settings. *Exceptional Children, 74*, 8–31.
- Swanson, H. L., & Deshler, D. D. (2003). Instructing adolescents with learning disabilities: Converting a meta-analysis to practice. *Journal of Learning Disabilities, 36*, 124–135.
- Torgesen, J. K., Houston, D. D., Rissman, L. M., Decker, S. M., Roberts, G., Vaughn, S., et al. (2007). *Academic literacy instruction for adolescents: A guidance document from the Center on Instruction*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.