

A Comparison of Multiple-Strategy Methods: Effects on Fourth-Grade Students' General and Content-Specific Reading Comprehension and Vocabulary Development

Deborah Simmons

Texas A&M University, College Station, Texas, USA

Angela Hairrell, Meghan Edmonds, and Sharon Vaughn

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

Ross Larsen, Victor Willson, William Rupley, and Glenda Byrns

Texas A&M University, College Station, Texas, USA

Abstract: In this study, we compared the effects of two experimental multiple-strategy approaches (content-area comprehension and vocabulary) to typical fourth-grade social studies instructional practices. An 18-week, cluster-randomized study was conducted to estimate effects measured by normative-referenced reading comprehension and vocabulary measures and researcher- and district-developed measures of social studies vocabulary and content. Forty-eight teachers and their respective 903 students from 15 schools were randomly assigned by school to one of three conditions: content vocabulary, content reading comprehension, or typical practice. Experimental teachers participated in 6 professional development sessions over 21 weeks. Structural equation modeling results indicated reliable differences favoring both experimental conditions over typical practice on the social studies content measure and substantively important effects on content and standardized vocabulary measures. Students in the vocabulary intervention also outperformed typical practice peers on the curriculum-based vocabulary assessment. Effects of the comprehension and vocabulary conditions were comparable except for the significant effect of vocabulary on the curriculum-based vocabulary measure. Effect sizes for teaching quality on the standardized comprehension measure ranged from $d = .26$ to $.32$; however, these effects were not statistically significant.

Keywords: Comprehension, vocabulary, content-area text, strategy instruction, social studies comprehension

Address correspondence to Deborah Simmons, Texas A&M University, Educational Psychology, 704 Harrington, College Station, TX 77843, USA. E-mail: dsimmons@tamu.edu

In recent years, reading practitioners and researchers have focused attention on the prevalent and persistent challenges facing students as they engage in the process of reading-to-learn from informational text (Biancarosa & Snow, 2004; Pearson, Hiebert, & Kamil, 2007; Rand Reading Study Group, 2002). The consistent message in reports, research syntheses, and handbooks of research stresses the multiple challenges of expository text and the benefit and need for instruction that supports students' access to and comprehension of this form of text—a challenge that begins around fourth grade (Baxter & Reddy, 2007; Gersten, Fuchs, Williams, & Baker, 2001; Kamil et al., 2008).

Fourth grade is considered a “watershed” grade because expectations for comprehension presume that students will be able to read increasingly complex text proficiently and independently. Specifically, when students enter fourth grade, the distribution of text type they read shifts to one that is predominantly expository and content-laden, that is, text about science, geography, social studies, and history (Willson & Rupley, 1997). Developmentally, many readers entering fourth grade have just acquired the capability of *learning-to-read*, whereas the tasks required in content-area classes expect them to be proficient at *reading-to-learn* from expository text, a skill not emphasized during earlier stages of reading development (Chall, 1996).

Thus, many students have not received sufficient instruction to adequately prepare them for the tasks that reading expository text requires (Biancarosa & Snow, 2004; Duke, 2000). Consequently, many students who were previously reading within grade-level expectations begin to fall behind and demonstrate difficulty understanding and interpreting what they read, often referred to as the “fourth-grade slump” (Chall, Jacobs, & Baldwin, 1990). This slump may be the result of many factors; however, there is general agreement that comprehending expository texts that are conceptually dense and organizationally complex requires adequate knowledge and use of vocabulary and self-regulation strategies to access and make sense of the content (Cromley & Azevedo, 2007; Gersten et al., 2001; Kemple et al., 2008).

SOCIAL STUDIES TEXT AND THE NEED FOR MULTIPLE STRATEGIES

Facilitating students' understanding of content-area texts is particularly challenging because multiple factors create barriers to comprehension. When comparing the content-area subjects, social studies is particularly textbook driven (Okolo, Englert, Bouck, & Heutsche, 2007) as teachers and students are presented with “inconsiderate” text (Armbruster & Anderson, 1984) as a primary source of information.

It is generally accepted that effective comprehension of expository text involves awareness and coordination of multiple teaching and learning strategies

(Kamil et al., 2008; Torgesen et al., 2007). Although single-strategy instruction promotes students' proficiency on specific skills (e.g., main idea, question generation), accruing evidence suggests that complex text comprehension requires strategy combinations (Baxter & Reddy, 2007; Scammacca et al., 2007). Compared to narrative text, however, few experimental studies have investigated multiple-strategy approaches to improving struggling readers' comprehension of social studies text in the middle and secondary grades (Edmonds et al., 2009; James-Burdumy et al., 2009; Okolo et al., 2007). Further, few contemporary multiple-strategy interventions in social studies have addressed both comprehension and vocabulary.

In a multistrategy comprehension intervention conducted in social studies, Klingner, Vaughn, and Schumm (1998) integrated cognitive strategies, including previewing, generating and answering questions, monitoring understanding, finding the main idea, and summarizing. Findings indicated that compared to a control group, students in the multiple-strategy condition demonstrated significantly greater growth on a standardized measure of comprehension and comparable performance on a measure of content learning.

In an examination of a multiple-strategy vocabulary intervention, Baumann, Edwards, Boland, Olejnik, and Kame'enui (2003) contrasted a textbook vocabulary condition with a combined morphological and contextual analysis condition. Students in the taught vocabulary condition learned more textbook vocabulary, whereas students in the combination condition were more successful at inferring untaught word meanings. No differences were detected on a measure of social studies learning.

Perhaps most noteworthy among multiple-strategy interventions is the large-scale experimental comparison of four reading comprehension curricula on fifth-grade students' comprehension on general, social studies, and science measures (James-Burdumy et al., 2009). The four curricula (Project CRISS, ReadAbout, Read for Real, and Reading for Knowledge) shared many common components focused on preparing students to strategically read and comprehend text (e.g., summarization, question generation). The interventions were designed to supplement the existing core curriculum, and teachers were asked to incorporate their assigned curriculum into their daily schedules and their core reading instruction. Moreover, according to the methods reported, teachers were asked to determine how best to incorporate strategies in their instructional practices that focused on general and content-area comprehension.

The primary finding of this impact study was that the reading comprehension test scores of the four reading comprehension curricula were not statistically significantly higher than scores in control schools on general or content-specific measures. These results were unanticipated; nonetheless, they reveal the complexity of improving expository text comprehension and shed light on areas that may strengthen future research, particularly the

need for professional development that situates practices into the contexts of classrooms.

Although the extant research offers some guidance on factors that influence learning and comprehension from text, the range of strategies is broad, the application to social studies text limited, and the most effective combinations elusive. In content-area classrooms where teachers are being asked to assume responsibility for content learning and strategy instruction, effective, parsimonious strategy combinations are of the essence. In a standards-based system where educators are accountable for multiple subjects in a fixed period, identifying effective methods that can be taught in the structural and temporal conditions of schools is essential.

In the present study, we compared two distinct instructional approaches that combined multiple teacher-directed and student-regulated components. One approach focused on teaching students cognitive strategies to comprehend and learn from text independently; the second combined teacher-directed and student-regulated strategies to promote content vocabulary acquisition and application. Both approaches used an explicit instructional progression that moved from teacher-directed to guided practice to independent practice. The contrast focused on different approaches using the same text. Our goal was to better understand the effects of these contrasting approaches, recognizing the dual need for parsimony and efficacy. In the following, we summarize the rationale and dimensions of each approach.

A MULTIPLE COGNITIVE STRATEGY APPROACH TO CONTENT-AREA COMPREHENSION INSTRUCTION

Developing new knowledge and comprehension in content areas involves the coordinated application of multiple reading strategies (Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006). To meet this lofty goal, instruction must go beyond the commonly observed techniques of students reading from text and answering a set of questions (Graves, 2006; Shanahan & Beck, 2006) and move toward more robust instruction. Evidence suggests that struggling readers experience difficulty in cognitive processes such as making inferences, drawing conclusions, and predicting outcomes. Indeed, when struggling readers are taught how to utilize a combination of metacognitive strategies that guide and support them in their self-monitoring of comprehension and teachers facilitate the process through explicit instruction, comprehension levels increase (Sencibaugh, 2007; Swanson, 1999).

Ideally, we want students to self-monitor their understanding while reading, to think about what they are learning and what does and does not make sense, and to be prepared to ask and answer questions about the text read while integrating recent and previous text reading. This is a multifaceted task and one that few fourth graders can do without a highly skilled teacher as their guide.

Thus, our goal was also to develop highly skilled teachers who understood and used research-based strategies that focused on the process of comprehension within the context of social studies learning. According to Swanson, Hoskyn, and Lee (1999), the essence of effective comprehension strategy instruction is parsimony. Yet there seem to be several essential strategies that are associated with improved comprehension that can be initially taught independently and then integrated into a flexible model of learning from text. At issue is which strategy combinations can be both effective and instructionally economical?

Based on several syntheses of practices associated with improved comprehension (Edmonds et al., 2009; Gersten et al., 2001; Kim, Vaughn, Wanzek, & Wei, 2004; Sencibaugh, 2007; Torgesen et al., 2007), as well as a large-scale observational study designed to identify effective instructional practices for improving reading outcomes (Taylor, Pearson, Peterson, & Rodriguez, 2003), we designed the following as an essential but parsimonious approach to improving reading comprehension:

1. *Main idea as a building block for summarization*: During early instruction, students were taught to identify the main idea of short passages. As students became facile at this skill, they were taught to use multiple main-idea statements to create summaries for longer text passages.
2. *Asking and answering questions with emphasis on higher-level questioning*: The goal of this practice was for students to ask questions periodically during the reading task to monitor their own understanding. To reach this goal, students were taught the characteristics of different question types and how to answer each. After building skills in question answering, students were taught to write and answer their own questions before, during, and after reading.
3. *Using graphic organizers to represent key learning*: Graphic organizers were used to organize the development of main idea, summarization, and questioning skills. Graphic organizers were provided to students during each lesson to cue students to recognize key learning opportunities where they could employ main idea and questioning strategies. They were also used to organize students' thinking and writing processes required during summarization.

A MULTIPLE-STRATEGY APPROACH TO CONTENT-AREA VOCABULARY INSTRUCTION

Our second approach to enhancing content-area comprehension builds on the knowledge base in vocabulary research and the recent "renaissance" of vocabulary instruction and research (Pearson et al., 2007, p. 282). Correlational, experimental, and meta-analytic research has documented the association between

vocabulary knowledge and reading comprehension (e.g., Baumann, 2009; Elleman, Lindo, Murphy, & Compton, 2009; Pearson et al., 2007). It is generally agreed that when students have knowledge of vocabulary that occurs in expository text and the ability to derive word meanings independently, they are more likely to understand the content of what they are reading (Baumann & Kame'enui, 1991; Boardman et al., 2008). This commonly held belief notwithstanding, the causal relation between vocabulary learning and improved comprehension is largely restricted to text-specific tasks and fails to generalize to general standardized measures, a finding that Pearson and colleagues attributed to insensitive and inadequate vocabulary measurement.

Vocabulary serves a critical role in content-area textbooks such as history, social studies, and science, and for many students it is a consequential obstacle to comprehension. Research evidence indicates that struggling readers experience difficulty comprehending content-area text because of the prevalence of unfamiliar vocabulary and an absence of domain knowledge or conceptual understanding (Biemiller, 2003). Researchers have noted (Kamil et al., 2008) that content-area text vocabulary “carries a large share of the meaning” (p. 11); therefore, readers cannot easily apprehend the meaning of connected text if they lack understanding of content-bearing words.

Recent guidance documents on how to improve adolescent literacy from the Institute of Education Sciences (Kamil et al., 2008) and the National Institute for Literacy (Baxter & Reddy, 2007) affirmed the importance of vocabulary instruction to content area instruction and learning. The level of evidence supporting the roles of explicit vocabulary instruction and explicit vocabulary strategy instruction was considered strong, as determined by a sufficient number of experimental and quasi-experimental studies (Kamil et al., 2008). Similar to comprehension, the most effective strategy, or combination of strategies, remains elusive.

In a recent meta-analysis, Elleman and colleagues (2009) found that “no matter what type of vocabulary instruction was used, it produced the same effect on comprehension as any other type of vocabulary instruction” (p. 36). On measures of vocabulary, interventions that involved greater amounts of discussion were associated with greater vocabulary growth. Overall, tests of mean weighted effect sizes (ESs) indicated that students who received vocabulary instruction significantly increased their performance on researcher-developed comprehension measures as well as standardized and researcher-developed measures of vocabulary over students in control or comparison conditions. Kamil et al. (2008) noted that although only some studies showed direct impact of vocabulary instruction on standardized measures of comprehension, vocabulary is a component construct of comprehension and important in its own right.

Based on research syntheses and primary studies of practices associated with enhanced vocabulary learning, we incorporated the following components in our content vocabulary development model. The strategies include those

explicitly taught by teachers as well as student-regulated generative vocabulary learning strategies.

1. *Prioritized sets of content vocabulary*: Prior to implementation of the intervention, teachers and researchers convened to identify, review, and select a defined set of vocabulary considered critical to text understanding and later learning. These vocabulary words were presented in sets of four to six per week (Beck, McKeown, & Kucan, 2002; Johnson, Gersten, & Carnine, 1986).
2. *Explicit teaching of multiple strategies to learn and apply vocabulary*: The goals of these practices were to provide initial instruction and to extend students' depth of knowledge through activities that included definitions, illustrations, word associations, and use of vocabulary in oral and written sentences (Beck et al., 2002; Kamil et al., 2008; Stahl & Fairbanks, 1986).
3. *Activation and building of background knowledge related to vocabulary*: Understanding and learning about the complex issues in social studies demands adequate background and world knowledge related to the ideas taught (Chen & Graves, 1995; Dole, Valencia, Greer, & Wardrop, 1991; Kintsch, 1998). Specific strategies to build and activate background knowledge included anticipation guides, text previews, and advanced organizers designed to help students relate or activate what they know with the content to be learned.
4. *Contextual clues to derive word meanings from text independently*: A variation of Baumann et al.'s (2003) contextual clue strategy was used to teach students an independent strategy to learn words from context. This strategy was first introduced explicitly with teacher modeling and gradually transferred to students (Fukink & de Glopper, 1998; Kuhn & Stahl, 1998; Stahl & Fairbanks, 1986).
5. *Semantic organizers to graphically organize and reinforce vocabulary learning*: Unit and chapter organizers were created to integrate target vocabulary within the larger units of information to be studied. Organizers were used to preview information and as postreading activities for students to summarize content using vocabulary (Bryant, Goodwin, Bryant, & Higgins, 2003; Carnine, Crawford, Harniss, Hollenbeck, & Miller, 2002; Jitendra, Edwards, & Sacks, 2004).

OVERVIEW AND RESEARCH QUESTIONS

Though multiple-strategy instruction is recommended to improve reading comprehension, few experimental studies have compared the efficacy of different multiple-strategy interventions with social studies text in the middle and secondary grades. In this study, we investigated the relative effects of two strategy combinations that explicitly prepared teachers to integrate content vocabulary,

reading comprehension instruction, and independent learning strategies in their teaching of social studies. We were interested in learning the overall effects of different combinations of strategies to typical practice and whether interventions were differentially effective by outcome measure. Further, we were interested in the influence of teachers' quality of instruction on student learning. Achievement was assessed through general measures of comprehension and vocabulary as well as specific measures of social studies vocabulary and content. Professional development for both experimental interventions was situated in the fourth-grade social studies classroom content and curriculum and based on best practices gleaned from professional development research. Fourth-grade social studies teachers were randomly assigned at the school level to one of three conditions (two treatment, one comparison): (a) multiple-strategy approach to content comprehension, (b) multiple-strategy approach to content vocabulary, and (c) typical instructional practice. In all three conditions, teachers taught the same units of social studies and used the same text for approximately the same amount of time. The following research questions guiding this investigation:

1. What are the effects of multiple-strategy interventions (i.e., cognitive comprehension and content vocabulary) compared to typical social studies practice on fourth-grade students' performance on general and social studies vocabulary and reading comprehension assessments?
2. Is there a differential effect of multiple-strategy interventions (i.e., cognitive comprehension versus content vocabulary) on student performance on general and social studies vocabulary and reading comprehension assessments?
3. How are the effects of the strategy approach influenced by teaching quality and fidelity?

METHOD

Participants

School Sites

The study was conducted in two medium-sized school districts in central Texas, both serving high proportions of children from low-socioeconomic households. District 1 had 7 elementary schools and all participated. District 2 had 15 elementary schools. Based on Title I designation and fourth-grade reading performance on state achievement tests, 8 schools in District 2 were identified by district administrators and agreed to participate.

In District 1, 77% of the students qualified for the free or reduced-price lunch program; in District 2, 65% qualified. Schools in Districts 1 and 2 were matched on the previous year's reading achievement performance on the Texas Assessment of Knowledge and Skills (TAKS; Texas Education Agency,

2005–2006) and assigned through stratified random assignment to one of three experimental conditions of social studies instruction: comprehension (C), content vocabulary (V), or typical practice (TP). Each condition was represented in multiple schools in each district. To maximize comparability between conditions and sites, all teachers who taught fourth-grade social studies and all students enrolled in their classrooms were invited to participate.

Fourth-Grade Social Studies Teachers

Forty-eight elementary teachers participated. All taught fourth-grade social studies during the 2006–2007 school year, either in self-contained ($n = 21$) or departmentalized ($n = 27$) classrooms. All teachers in departmentalized settings taught two to three sections of social studies per day. Number of years teaching experience averaged 9.36 years ($SD = 9.58$) in the comprehension condition, 5.72 years ($SD = 6.38$) in the content vocabulary condition, and 7.50 years ($SD = 9.83$) in the typical practice condition. Mean differences in years teaching did not differ statistically across groups, $F(2, 44) = 0.57$, $p > .40$, nor did years teaching fourth grade, $F(2, 44) = 0.06$, $p > .90$.

Table 1. Teacher demographic information by condition

Variable	Comprehension		Vocabulary		Typical Practice	
	<i>N</i>	(%)	<i>N</i>	(%)	<i>N</i>	(%)
Number	17		17		14	
Gender						
Male	1	(5.9)	0	(0.0)	4	(28.6)
Female	16	(94.1)	17	(100.0)	10	(71.4)
Ethnicity						
African American	5	(29.4)	0	(0.0)	0	(0.0)
Hispanic	4	(23.5)	4	(23.5)	5	(35.7)
European American	6	(35.3)	13	(76.5)	7	(50.0)
Other	2	(11.7)	0	(0.0)	2	(14.3)
Teaching experience						
Total years teaching	7.50	($SD = 9.8$)	5.72	($SD = 6.4$)	9.36	($SD = 9.6$)
Years teaching fourth grade	4.33	($SD = 7.6$)	3.22	($SD = 5.1$)	4.93	($SD = 5.7$)
Highest degree earned						
BS	16	(94.1)	14	(82.4)	12	(85.7)
MS	1	(5.9)	3	(17.7)	2	(14.3)
Classroom organization						
Self-contained	3	(17.6)	13	(76.5)	5	(35.7)
Departmentalized	14	(82.4)	4	(23.5)	9	(64.3)

Table 1 presents teacher demographic information, including gender, race/ethnicity, teaching experience, degrees earned, and certifications held.

Chi-square analyses conducted on categorical variables and one-way analyses of variance conducted on continuous variables indicated no reliable differences among groups on gender, years teaching, years teaching social studies, or highest degree earned. There was a significant relation among groups on ethnicity, $\chi^2(4) = 12.93$, $p = .01$, with higher percentages of African American teachers in the comprehension condition. During the school year, one teacher moved and one transferred to a different grade, and neither completed the study. No significant differences were noted between bilingual and general education teachers on gender, years teaching, years teaching social studies, or highest degree earned. Likewise, comparisons of self-contained versus departmentalized classrooms revealed no statistically significant differences among groups on these same variables. Results revealed a significant difference in the percentage of departmentalized and self-contained classrooms between conditions, $\chi^2(2) = 8.13$, $p = .02$. We addressed this difference by treating classroom organization as a covariate in subsequent analyses.

Students

Informed consent was obtained for 911 fourth-grade students from 61 social studies classes: 324 in the C condition, 355 in the V condition, and 232 in the TP comparison condition. The percentage of students who consented to participate was 71% from District 1 and 73% from District 2. Students represented diverse ethnic and racial groups, with average percentages across all students of 16% African American (C = 20%, V = 14%, and TP = 11%), 70% Hispanic (C = 66%, V = 68%, and TP = 80%), and 13% European American (C = 13%, V = 16%, and TP = 9%). One percent of the students reported other races/ethnicities. Students were, on average, 9.7 years of age (SD = .60).

Group equivalence was evaluated using analyses of variance for continuous variables (e.g., age) and chi-square tests for categorical variables (e.g., gender, ethnicity, special education services). There was a larger percentage of Hispanic students in the TP condition than the other conditions, $\chi^2(4) = 13.88$, $p < .01$. To address this difference, we treated bilingual classrooms as a covariate in subsequent analyses.

Of the 911 students, 624 (68%) participated in all pretest and posttest assessments; 287 (32%) did not complete at least one of the measures. Eight students' data were so incomplete that they were dropped, resulting in 903 students in the database. The primary reason for attrition was students moving to another school. Chi-square analysis confirmed no significant relation between treatment and attrition ($\chi^2 = .61$, $p = .74$); therefore, we assumed that the changes were random. Imputation was employed with the 903 students' variables to construct complete data for analysis. Details of the imputation process are provided in the Results section.

Measures

All pretest and posttest measures were group administered by research personnel prior to intervention (September) and again at intervention completion (March). Measures of teaching quality and treatment fidelity were collected at three different points during the study.

Comprehension Measures

Gates-MacGinitie Reading Test—4th Edition—Passage Comprehension Subtest (GMRT-4). The GMRT-4 (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) is a group-administered test that uses both narrative and expository passages, ranging from 3 to 15 sentences. Students read passages silently and answer three to six multiple-choice questions. Across the test levels, internal consistency reliability ranges from .91 to .93, and alternate form reliability ranges from .80 to .87 in national samples.

Vocabulary Measures

Test of Reading Comprehension, Social Studies Vocabulary subtest (TORC-3 SS). The TORC-3 SS (Brown, Hammill, & Wiederholt, 1995) was used as the general outcome measure of vocabulary and was administered pre- and postintervention. The TORC-3 SS is an untimed, group-administered test of 25 items measuring social studies vocabulary and knowledge. Each item stem presents three words that are similar in meaning, and the student selects two words from a list of four word choices that have similar meaning to the three stem words. The TORC-3 was normed on 1,962 students in 1994 across 19 states. Reliability of the measure is in the .90 range.

Curriculum-Based Vocabulary Assessment (CBA-V). CBA-V is a fluency-based measure constructed by the researchers to assess knowledge of social studies vocabulary identified by school personnel as critical to success in fourth-grade social studies. The CBA-V measure was administered pre- and postintervention to assess cumulative growth in vocabulary knowledge. Items were from district-identified vocabulary selected in consultation with a team of expert social studies teachers. Ten words for every 6-week unit were randomly selected for a total of 30 items. The 5-min, timed test was presented on an 8.5 × 11 page with the list of vocabulary words in the first column and a list of definitions for the words in the second column. Words and definitions were randomly ordered. Example vocabulary and definitions included *agriculture*—the planting of seeds to grow food; *descendants*—a person's children, grandchildren, and so on; *irrigate*—to water. The CBA-V was modeled after commonly used curriculum-based vocabulary measures reported in previous research (Espín, Busch, Shin, & Kruschwitz, 2001). Correlations between the

CBA-V and the TORC-3 SS vocabulary and GMRT-4 reading comprehension posttests were significant ($r = .34$ and $.44$, respectively).

Social Studies Content Test. This measure was designed by one of the participating school districts to measure student knowledge of concepts taught during each 6-week unit of study. For this measure, we used the tests that corresponded with the three units taught during the intervention period. The test protocol included 10 multiple-choice questions comprised of a short question stem followed by a set of four answer choices. Students completed the tests during a group-administered testing session at the conclusion of each of three units. Students earned 1 point for each correct answer, with a maximum score of 10 on each test. For the purpose of analysis, results of the three administrations were summed for a total of 30 items. Items primarily sampled factual information such as “What did the Mayan, Aztec, and Incan civilizations have in common?” and “The first explorer to reach Texas was ____.” Scores from this measure were in the low-moderate range ($r = .40$ and $.36$, respectively), and significantly correlated with TORC-3 SS and GMRT-4 measures.

Covariate Measures of Reading

The TAKS (Texas Education Agency, 2005–2006) Grade 3 reading comprehension score was obtained for all students and used as a covariate in statistical analyses. Administered to all Grade 3 students, this is a measure of general reading achievement in which students read a series of narrative and expository passages, each followed by a series of multiple-choice questions. Passages are approximately 500 to 700 words long. Scaled scores range from 1,357 to 2,616, with a score of 2,100 designated as passing. Internal consistency reliability coefficients range from the high .80s to the low .90s (Texas Education Agency, 2005–2006).

Test of Silent Contextual Reading Fluency (TOSCRF). The TOSCRF (Brown et al., 1995) is a 3-min, group-administered measure of silent contextual reading fluency. The measure broadly assesses word recognition, syntax and morphology, and students’ fluency in applying word knowledge and grammatical knowledge to determine the meaning of connected text. Passages from the Gray Oral Reading Test are printed in all caps with no punctuation or spaces between words. Students place slash marks between words. There are four alternate forms in each grade, Grades 2 to 12. The measure has documented reliability (range = $.82$ – $.88$). Reported scores are highly correlated with measures of fluency ($r = .89$ criterion validity with the Gray Oral Reading Test) and moderately correlated with comprehension ($r = .61$ Woodcock Johnson Passage Comprehension). Scores from this measure also correlated with measures

of word identification/fluency (e.g., for the Test of Word Reading Efficiency, $r = .70-.74$ across subtests).

Intervention Fidelity and Quality of Instruction Measures

Teachers' fidelity of implementation and quality of instruction were ascertained by coding audio-recorded lessons of intervention implementation. Teachers selected and audio-taped a representative lesson from each 6-week unit using a digital recorder for a total of three recorded lessons. Lesson files were coded by members of the research team using a fidelity protocol. Researchers collaborated to develop the protocol, which was adapted from similar instruments used in other research studies (Vaughn, 2002; Vaughn, Klingner, & Bryant, 2001).

The adapted instrument included instructional variables representing the research-supported comprehension and vocabulary components (e.g., generating and answering questions, introducing content vocabulary, providing multiple exposures to vocabulary words) and variables describing the elements of explicit strategy instruction (e.g., modeling, guided practice, application). Three fidelity protocols were developed, one for each condition. Each observation protocol included six items representing implementation of the major intervention components. Items were rated on a 4-point scale ranging from 0 (*not at all*) to 3 (*exemplary implementation*). In addition to specific components, two summary items were included. These items were adapted from existing observation systems (Foorman & Schatschneider, 2003; Taylor et al., 2003). Coders rated the overall program implementation on a 7-point scale ranging from 1 (*less than adequate*) to 7 (*above expectations*). Ratings on this item were substantiated by ratings on the major intervention components. An item of overall instructional quality was used to obtain a summative assessment of a teacher's instructional quality. Teaching quality was rated on a 7-point scale, ranging from 1 (*not at all effective*) to 7 (*highly effective*). Several general features of effective instruction were considered such as lesson pacing, corrective feedback, and explicitness of instruction when rating this item. Global ratings of teaching quality, although moderately inferential, have been found to be reliable (e.g., $> .80$; Foorman & Schatschneider) and valid. For example, Foorman and Schatschneider found moderate to high correlations between a single-item rating scale and (a) components of teacher competency (.44-.80) and (b) student comprehension outcomes in first grade (.62-.70), and small to moderate correlations with second-grade comprehension (.18-.41).

The protocol for the typical practice comparison condition consisted of a checklist designed to document the presence of practices featured in the intervention conditions. As mentioned, the item for rating overall teaching quality was also included on the typical practice protocol. Team members coded 2 randomly sampled recordings of instruction per teacher (94 recordings). All 3 recordings were coded for three teachers. The random selection was counterbalanced such that for one third of the teachers, the first 2 recordings were

coded, for another third, the second 2 recordings were coded, and for the last third, the first and last recordings were coded. Coders demonstrated a .88 interrater reliability prior to coding independently. Agreement was calculated as the number of agreements divided by the number of agreements plus disagreements across all eight items. In addition, 20% of recorded lessons were double-coded and assessed for rater agreement. In the event of a discrepancy between coders, items in question were reviewed and rated by an expert coder. Reliability on double-coded tapes ranged from 0.75 to 1.0, representing an acceptable range for a moderate-inference instrument.

Instructional Content, Procedures, and Professional Development

The study was conducted from October through January. Figure 1 provides a schedule of content instruction used by all groups. Specific strategies for comprehension and vocabulary are also outlined by condition. All conditions (content comprehension, vocabulary, or typical practice) were taught during regularly scheduled social studies periods. Teachers who taught self-contained classes ($n = 21$) taught one section of social studies per day; teachers in departmentalized structures ($n = 27$) taught two sections per day. To situate the experimental interventions in social studies content, we designed both the comprehension and vocabulary conditions using the following content and instructional elements.

Content. Comprehension and vocabulary instruction and strategies were integrated into the fourth-grade social studies curriculum, which was based on state-designated social studies and reading/language arts standards. The content was organized around district-developed curricular maps that paced content coverage in 6-week units with weekly instructional targets. The study was conducted using the district-adopted social studies textbook (*Social Studies Texas*, 2003) as a primary instructional source. Across the three, 6-week units, the following topics were covered: The First Texans, Europeans Come to Texas, Texas Colonies and Conflicts, The Republic of Texas, and The Lone Star State.

Prior to the study, teachers reported that they typically allocated 225 min to social studies instruction each week. This time was most commonly distributed evenly over 5 days. Teachers in the comprehension and vocabulary conditions were asked to embed 90 min of strategy instruction in their regular social studies instruction. Depending on the teachers' schedules, this time was distributed 30 min three times weekly or 45 min twice weekly. Time allocated to intervention strategies was designed to replace rather than supplement part of the time focused on social studies.

A common structure was used across conditions that simulated the lesson plan format used by the districts. Strategies to develop vocabulary and comprehension were drawn from the extant evidence base reviewed previously. The lesson plan organized strategies according to when they would occur in

Condition	Unit 1	Unit 2	Unit 3
Comprehension	Preteach proper nouns and preview		
	Questioning strategy		
		Main idea strategy	
			Summary writing using a graphic organizer
Vocabulary	Identify critical vocabulary		
	Explicit vocabulary instruction using vocabulary maps		
	Connect content and vocabulary through chapter organizers		
	Practice activities		
		Build background knowledge through anticipation guides	
			Context clue strategy

Figure 1. Instructional sequence for comprehension and vocabulary interventions.

the instructional sequence. In some instances, a strategy (e.g., semantic map or graphic organizer) was used before, during, and after instruction. A small number of strategies was introduced in each 6-week unit and designed to cumulatively build on one another. That is, students were taught in the second 6 weeks to generate main ideas for designated sections of text. In the third 6-week period, they were taught how to synthesize the main-idea statements

to construct a summary statement for the entire text. In initial instruction, teachers and students were provided scaffolds in the form of specific questions and language to facilitate implementation. These supports were systematically withdrawn over the three units.

Strategies and content were taught using an explicit instructional model that consisted of the following elements: clearly explaining the instructional task, modeling the instructional task or strategy, breaking the task or strategy learning into small steps, providing sufficient practice at each step, and allowing ample opportunities for feedback and extended learning. To promote active engagement, students worked in collaborative study pairs to complete designated tasks such as generating questions, constructing main ideas, or completing vocabulary maps. Lessons for both comprehension and vocabulary were designed around common social studies content and scheduled for similar amounts of time. In the following, we describe the unique elements of each condition.

Content Comprehension Procedures

The comprehension intervention was developed to be responsive to understanding and learning information text and was organized into before-, during-, and after-reading phases. Specifically, strategies emphasized the content, text structure, and purpose for reading. In particular, the importance of proper nouns (e.g., key persons and places) was emphasized. Teachers previewed important proper nouns (e.g., the *Alamo*), pronouncing the word and providing a brief definition or explanation prior to reading. Teachers also lead a brief preview that guided students in identifying key concepts and content (headings, titles, bold print) and making connections to prior learning. As part of the preview, teachers provided students with the “big idea” of the section being read. During reading, students learned and practiced three core comprehension strategies that were taught within the context of information text and phased in across the 18-week intervention—asking and answering questions, writing a main-idea statement, and summarizing larger pieces of text.

Strategies were taught using an explicit instructional model that included teacher modeling followed by guided and independent practice. Strategy instruction was distributed and cumulatively integrated with a new strategy added to students’ routine every 6 weeks. This process allowed students time to practice applying one strategy during reading before being introduced to the next strategy. Strategies were sequenced such that students gradually gained the skills needed to successfully apply the final, most challenging strategy: summarization. In other words, students first learned to ask important questions about the text to guide their thinking and then used questions to help them develop a main-idea statement for a short section of text. They then learned to combine several main idea statements into a longer summary with the aid of a graphic organizer.

The questioning strategy, adapted from Question and Answer Relationships (Raphael, 1986; Raphael, Highfield, & Au, 2006), taught students to ask and answer questions ranging in difficulty from easy (answers found explicitly in text) to more difficult (answers requires students to make connections with prior learning). The main-idea strategy, Get the Gist (Klingner, Vaughn, Dimino, Schumm, & Bryant, 2001), consisted of cues that prompted students to first identify the most important thing about the paragraph and, second, to tell the most important information about that thing in 10 or fewer words. After students had multiple practice opportunities writing gist statements, they learned to organize their gist statements on a content web and use the relationships depicted on the Web to write a summary.

Content Vocabulary Procedures

The objective of the content vocabulary condition was to support teachers' use of multiple strategies to promote students' learning of social studies vocabulary and use of vocabulary to comprehend and learn from social studies text. Toward that end, we combined teacher-directed strategies and independent word-learning strategies integrating opportunities for contextualized and decontextualized vocabulary use. Introduction and application of the strategies was scaffolded across the 18 weeks of the study.

In the first 6-week unit, the focus was on explicit teaching of content vocabulary. Vocabulary were selected from district curriculum by teacher experts and included terms from state standards and the adopted social studies textbook. From this list, teachers selected 4 to 6 vocabulary per week based on their importance to understanding content and overall learning and alignment with chapter content for a total 62 words. For example, in Week 1 target vocabulary included *descendant*, *culture*, *artifact*, and *agriculture*. See the appendix for a list of vocabulary. Vocabulary maps were constructed by researchers for each target word. Maps integrated multiple strategies including reviewing a definition for critical terms, viewing illustrations, identifying root words, using the word in context, associating new vocabulary with related vocabulary, and using the word in an original sentence or definition. Teachers explicitly modeled the connection between content knowledge and new vocabulary using a semantic features analysis grid. Practice activities were built into the intervention to provide students with multiple exposures to new words. Practice activities included word association games, use of multiple vocabulary words in sentences, and summary activities.

In the second 6-week unit, background knowledge was activated through anticipation guides (Ausubel, 1990). Students were provided statements about the content, which included vocabulary terms that were then used to stimulate discussions of critical background knowledge. Example statements included, "The *boundary* between Texas and Mexico is the Rio Grande." In the final 6-week unit, an independent word-learning strategy was added to the instructional

routine. Students were taught to use contextual and morphological clues to determine the meaning of new vocabulary based on the work of Baumann et al. (2003).

Professional Development

Professional development sessions were distributed over 21 weeks and included face-to-face researcher-directed sessions followed by application opportunities and small-group teacher study teams for a total of 18 hr of professional development. Specifically, at the beginning of each 6-week unit, teachers engaged in 4-hr professional development sessions. At the midpoint in each unit, teachers participated in 2-hr teacher study teams. Practices were introduced gradually over the three 6-week units to allow teachers the opportunity to learn, practice, and discuss each strategy. Teachers received no additional in-classroom professional development or coaching.

Typical Practice Condition

Teachers who participated in the typical practice condition were asked to maintain their customary instructional practices and were not apprised of the purpose of the research. Teachers reported using the textbook and district curriculum guides to inform instruction. The most commonly observed instructional activity was reading the textbook and answering questions. Teachers often read the text to the students or used audiotapes of the text due to the difficulties students experienced with independent reading of the text. Other observed activities included content-focused projects, such as building models of Native American housing or researching current events in newspapers, or completing worksheets.

DATA ANALYSIS

Structural equation modeling (SEM) with a hierarchical linear model framework was used to analyze the data. SEM was chosen because of its capacity to estimate and test models with multiple outcome measures. In addition, SEM helps control for the family-wise error that may occur when performing multiple statistical tests (Kline, 2005). Because of non-independent observations due to the nested structure in our data, multilevel modeling was applied to analyze the data, as it takes nonindependence into account (Kline, 2005).

Two multilevel models were created to answer the research questions regarding the impact of comprehension and vocabulary interventions compared to typical practice and the differential impact of the experimental conditions to one another. In each model, the observed measure for teaching quality was

used to examine and test its relationship to the four dependent measures in the context of interventions.

In the within-level models (student level), a latent variable (Reading) was created using TAKS and TOSCRF. This latent variable was then used to predict pretest scores, with the Reading variable significantly predicting pretest outcomes. The proportion of variance explained (R^2) is reported for the within-level variables. Covariates for the between-level models (i.e., bilingual and class structure) were used to address disproportional distributions identified in the first stages of data analysis. Models were run with random effects for generalizability purposes and with fixed effects to compute standardized effects. Effects sizes for between-level models were computed as standardized path coefficient differences and are interpreted as “substantively important” if values exceeded 0.25 (What Works Clearinghouse, 2008, p. 22).

First, the normality of the endogenous variables (i.e., GMRT-4, CBA-V, TORC-3 SS, and social studies content) was checked through graphical means and found to be acceptable. In addition, the linearity of the endogenous variables with their respective covariates was found to be linear. Because one covariate was non-normal (TAKS), it was transformed using the log transformation. When the models were run with and without transformations, results were very similar; thus, the untransformed model was used to assist in ease of interpretation. To address missing data, we used the multiple-imputation technique PROC MI in SAS. Ten imputed data sets were created. Each dataset was analyzed individually and averaged results reported.

RESULTS

Descriptive Statistics

Descriptive statistics for all outcome variables by condition are shown in Table 2. All represent raw scores except TAKS, which is a scaled score. Zero-order correlations are reported in Table 3. All correlations were significant at the .01 level.

Model 1: Comprehension and Vocabulary Compared to Typical Practice

Model 1 (see Figure 2) compares the comprehension and vocabulary groups to typical practice instruction in the presence of classroom structure (self-contained or departmentalized), teaching quality, and bilingual classroom covariates. Pretest measures, TAKS, and TOSCRF were included to control for initial student ability. Table 4 displays the results of the structural equation model for the between-level analysis comparing comprehension and vocabulary to typical practice. As illustrated, for each outcome variable and predictor

Table 2. Pre- and posttest means and standard deviations for student achievement measures

Measure	Comprehension		Vocabulary		Typical Practice	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
CBA-V Pretest	0.88	(1.71)	0.81	(1.01)	0.76	(1.18)
CBA-V Posttest	6.03	(5.73)	11.43	(7.81)	5.47	(5.51)
GMRT-4 Pretest	19.26	(9.64)	19.48	(9.89)	15.55	(8.97)
GMRT-4 Posttest	24.84	(10.36)	24.64	(10.72)	20.89	(10.09)
TORC-3 SS pretest	37.73	(6.89)	37.94	(6.29)	36.42	(6.32)
TORC-3 SS Posttest	41.02	(5.30)	40.94	(5.07)	39.57	(5.37)
Social Studies Content	16.48	(4.89)	17.50	(4.08)	15.55	(8.97)
TAKS	2275.75	(161.36)	2263.36	(225.26)	2244.99	(154.75)

Note. CBA-V = Curriculum-Based Vocabulary Assessment; GMRT-4 = Gates-MacGinitie Reading Test–4th Edition–Passage Comprehension Subtest; TORC-3 SS = Test of Reading Comprehension, Social Studies Vocabulary subtest; TAKS = Texas Assessment of Knowledge and Skills.

variable, four numbers are reported: the unstandardized path weight (*B*), the standard error, the standardized path weight (*B*) reported as an ES in standard deviation units, and the significance level.

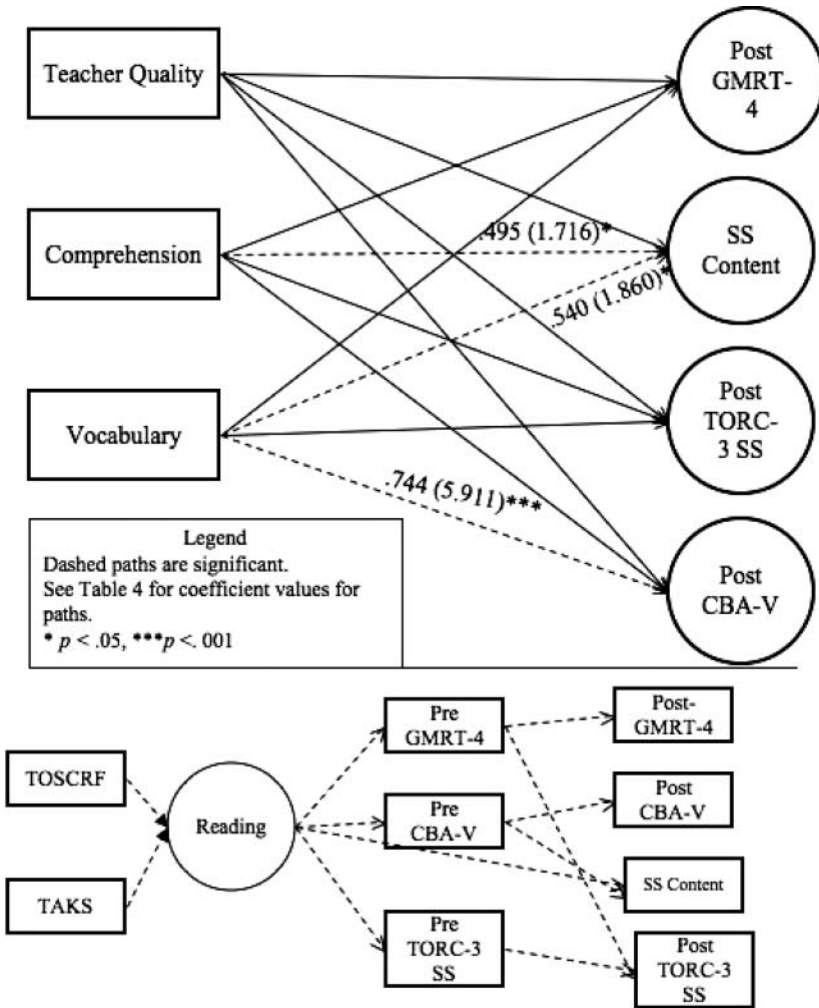
For model 1, the fit indices were $\chi^2/df = 2.21$, comparative fit index = .99, Tucker–Lewis Index = .96, root mean square error of approximation = .04, standardized root mean square residual (SRMR; within) = .02, and SRMR (between) = .11. Results of the within-in model (student level) showed that all posttest measures were predicted by pretest measures. The within-level covariates explained between 11% and 52% of the outcome variable. Specifically, GMRT-4 explained 52%; TORC-3 SS, 11%; and Social Studies Content, 28%. The CBA-V did not converge in the model and cannot be reported.

Table 3. Zero-correlations among outcome variables

	1.	2.	3.	4.
1. GMRT-4 Posttest	—	.44*	.61*	.40*
2. CBA-V Posttest	.44*	—	.34*	.42*
3. TORC-3 SS Posttest	.61*	.34*	—	.36*
4. Social Studies Content	.40*	.42*	.36*	—

Note. GMRT-4 = Gates-MacGinitie Reading Test–4th Edition–Passage Comprehension Subtest; CBA-V = Curriculum-Based Vocabulary Assessment; TORC-3 SS = Test of Reading Comprehension, Social Studies Vocabulary subtest.

* $p < .01$.



Note. Analysis controls for classroom structure (self-contained or departmentalized) and for bilingual classrooms. All errors were correlated. Standardized path coefficients are reported first, with unstandardized path coefficients in parentheses.

Figure 2. Model 1: Comparison of comprehension and vocabulary interventions to typical practice. Note. GMRT-4 = Gates-MacGinitie Reading Test–4th Edition–Passage Comprehension Subtest; SS = social studies; TORC-3 SS = Test of Reading Comprehension, Social Studies Vocabulary subtest; CBA-V = Curriculum-Based Vocabulary Assessment; TOSCRF = Test of Silent Contextual Reading Fluency; TAKS = Texas Assessment of Knowledge and Skills.

Table 4. Between-level estimates for Model 1: Comprehension and vocabulary interventions compared to typical practice

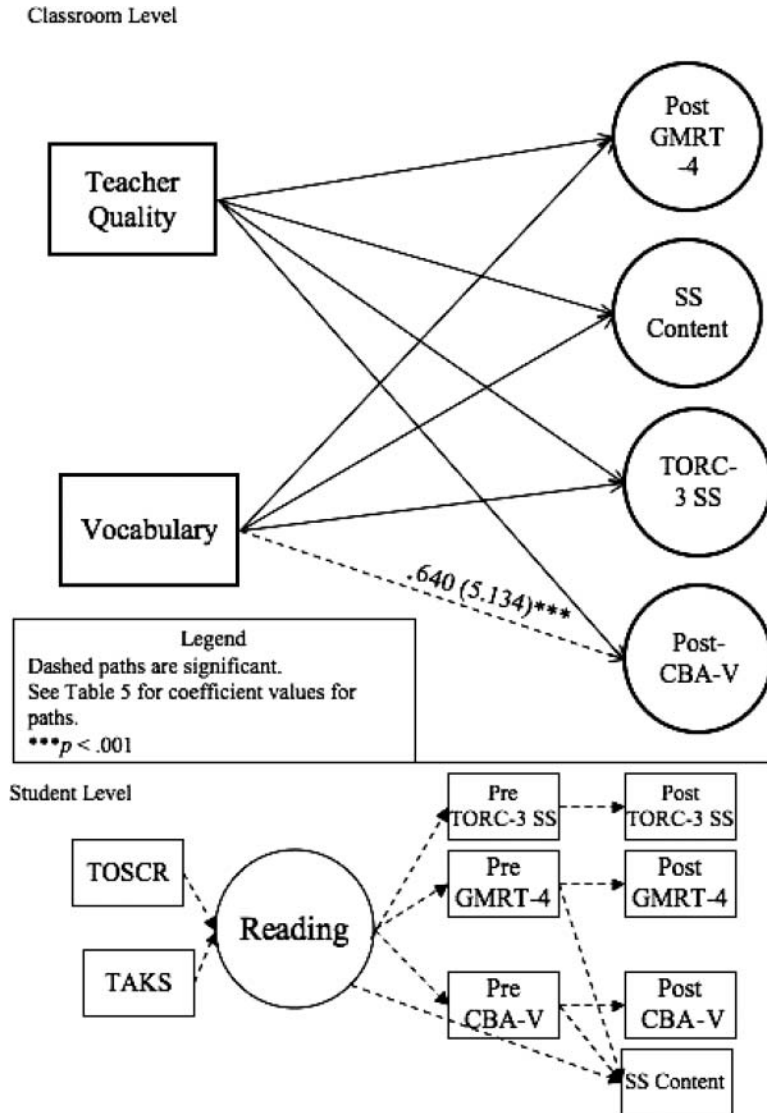
Path	B	SE	β	<i>p</i>
Teaching quality				
Post-GMRT-4	.235	.271	.258	.382
Post-TORC-3SS	.042	.136	.112	.760
Social Studies Content	.104	.224	.091	.636
Post-CBA-V	.036	.417	.013	.931
Comprehension				
Post-GMRT-4	-.064	.668	-.026	.941
Post-TORC-3SS	.350	.485	.313	.475
Social Studies Content	1.716	.788	.495	.020
Post-CBA-V	.266	1.532	.034	.860
Vocabulary				
Post-GMRT-4	.181	1.054	.071	.860
Post-TORC-3SS	.396	.581	.360	.487
Social Studies Content	1.860	.931	.540	.031
Post-CBA-V	5.911	1.634	.744	.001

Note. B and SE represent unstandardized coefficient values. β represents standardized coefficient values. Post-GMRT-4 = Gates-MacGinitie Reading Test—4th Edition—Passage Comprehension Subtest; Post-TORC-3 SS = Test of Reading Comprehension, Social Studies Vocabulary Subtest; Post-CBA-V = curriculum-based vocabulary measure.

In the between-level model (teacher level), results indicated that both the comprehension and vocabulary groups significantly outperformed typical practice on the social studies content test (see Table 4). For the comprehension condition, the standardized path coefficient or ES (*B*) was .50 (*B* = 1.72, *p* = .02). Similarly, the standardized path coefficient from vocabulary to the content assessment was .54 (*B* = 1.86, *p* = .03). The standardized path coefficient from vocabulary to the CBA-V was .74 (*B* = 5.91, *p* < .001). At posttest, students in the vocabulary condition knew almost six more words on the CBA-V measure when compared to typical practice students. No reliable differences between typical practice and either experimental intervention were observed on the general standardized measures of reading comprehension (GMRT-4) or social studies vocabulary (TORC-3 SS). However, both conditions yielded substantively important effects (comprehension ES = .31; vocabulary ES = .36) on the TORC-3 SS.

Model 2: Comparing Vocabulary and Comprehension Conditions

Model 2 (see Figure 3) compares the vocabulary and comprehension conditions to each other in the presence of teaching quality, classroom structure, and



Note. Analysis controls for classroom structure (self-contained or departmentalized) and for bilingual classrooms. All errors were correlated. Standardized path coefficients are reported first, with unstandardized path coefficients in parentheses.

Figure 3. Model 2: Comparison of comprehension and vocabulary interventions, including teaching quality. *Note.* GMRT-4 = Gates-MacGinitie Reading Test–4th Edition–Passage Comprehension Subtest; SS = social studies; TORC-3 SS = Test of Reading Comprehension, Social Studies Vocabulary subtest; CBA-V = Curriculum-Based Vocabulary Assessment; TOSCRF = Test of Silent Contextual Reading Fluency; TAKS = Texas Assessment of Knowledge and Skills.

Table 5. Between-level estimates for Model 2: Comparison of comprehension and vocabulary interventions

Path	B	SE	β	<i>p</i>
Teaching quality				
Post-GMRT-4	.349	.289	.317	.220
Post-TORC-3 SS	.030	.160	.072	.842
Social Studies Content	.118	.239	.091	.618
Post-CBA-V	.248	.470	.074	.598
Vocabulary				
Post-GMRT-4	-.114	.794	-.045	.879
Post-TORC-3 SS	-.010	.450	-.013	.976
Social Studies Content	.205	.683	.062	.775
Post-CBA-V	5.134	1.291	.640	.001

Note. B and SE represent unstandardized coefficient values. β represents standardized coefficient values. Post-GMRT-4 = Gates-MacGinitie Reading Test–4th Edition–Passage Comprehension Subtest; Post-TORC-3 SS = Test of Reading Comprehension, Social Studies Vocabulary Subtest; Post-CBA-V = curriculum-based vocabulary measure.

bilingual classroom covariates. Student level covariates from Model 1 were maintained, and results were comparable to the findings reported previously. For Model 2, fit indices were $\chi^2/df = 2.53$, comparative fit index = .98, Tucker–Lewis Index = .96, root mean square error of approximation = .047, SRMR (within) = .02, and SRMR (between) = .12.

Findings indicated that the effect of the two experimental interventions differed reliably on only one measure, CBA-V, with a standardized coefficient (*B*) of .64 ($B = 5.13$, $p < .001$) (See Table 5). At posttest, students in the vocabulary condition knew on average five words more on the CBA-V than students in the comprehension condition.

Dosage, Fidelity to Treatment, and Quality of Instruction

Each week, classroom teachers reported through online logs the amount of time they taught social studies. Teachers in the comprehension and content vocabulary conditions reported an average of 1.46 and 1.47 hr per week, respectively (range = .96–1.96 for comprehension; .88–1.86 hr for content vocabulary). Teachers in the typical practice condition reported 1.68 average hr per week, with a range of 1.26 to 1.96 hr. There were no statistically significant differences between conditions, $F(2, 43) = 2.37$, $p = .11$, regarding dosage or amount of social studies instruction provided per week. Follow-up analyses indicated that the amount of time experimental teachers devoted to social studies instruction did not differ significantly over the three 6-week units.

Fidelity to treatment was assessed only for teachers in the content vocabulary and comprehension treatments. Fidelity was initially included as a teacher-level variable in the multilevel analyses for each outcome measure. In addition, Fidelity \times Treatment Condition interactions were included as a predictor to consider differential effects of treatment for higher or lower levels of implementation. Overall mean fidelity scores were 4.75 ($SD = 1.65$) for the comprehension condition and 5.52 ($SD = 0.79$) for the vocabulary condition. These means were not statistically different from one another. Mean fidelity ratings for the comprehension condition were 4.75, 4.80, and 4.12 for the three observation points and 5.67, 5.45, and 4.61 for the vocabulary condition, respectively. Simple correlations between fidelity and student achievement outcomes never exceeded 0.13, with only one being significant at $p < .01$, the CBA-V. Although there were differences between vocabulary (higher) and comprehension on fidelity, the correlation was reduced from 0.047 to 0.019 when teaching quality was partialled out. This suggests that fidelity would not have differentially contributed to any achievement outcomes.

Regarding teaching quality, across observations, mean quality on a 1-to-7 scale averaged 4.75 ($SD = 1.44$) for the comprehension condition, 5.26 ($SD = .93$) for content vocabulary, and 2.73 ($SD = .65$) for typical practice. Measures of reliability or stability of teaching quality as indexed by correlations between observations were .70 for comprehension, .60 for content vocabulary, and .67 for typical practice. As indicated in Tables 4 and 5, the effect of teaching quality did not reach statistical significance on any of the student outcome measures at the $p < .05$ level. ESs were $d = .26$ on the GMRT-4 reading comprehension measure on the model comparing experimental interventions to typical practice and $d = .32$ on the comprehension to vocabulary model indicating a substantively important impact.

DISCUSSION

This study was conducted to advance our knowledge of instructional methods to promote fourth-grade students' understanding of content-area text, an area of critical importance with limited experimental research. Our first set of findings involves the relative benefit of two parsimonious multistrategy interventions compared to typical social studies instruction. In general, the impact of both interventions was stronger on measures more proximal to the content of instruction than standardized vocabulary and comprehension measures. On the district measure of social studies learning, findings indicated that students who participated in either form of multiple-strategy instruction (comprehension or vocabulary) learned more content than in typical social studies instruction. This finding is encouraging, as prior studies focused on social studies text (e.g., Baumann et al., 2003; Klingner et al., 1998) found comparable but not

differential performance of vocabulary and comprehension interventions over typical practice on content-area measures.

It is important to note that the social studies content measure in the current study was district-developed, and is more like researcher-designed measures than more global or general measures of content learning. ESs of both interventions (comprehension ES = .50; vocabulary ES = .54) were lower than effects (ES = .89) documented by Edmonds et al. (2009) on researcher-developed measures. This difference is attributable to two possible factors. First, in the present study, we included the full range of learners in the classroom, whereas students in the Edmonds et al. study were older, struggling readers, where the impact of intervention is often larger. Second, in the Edmonds' et al. meta-analyses, narrative text was predominant compared to expository text.

Although the pattern of findings on social studies learning was similar for both experimental conditions compared to typical practice, a different, but not unexpected effect was documented favoring the vocabulary intervention over typical practice on the curriculum-based assessment of vocabulary (ES = .74). Consistent with Baumann et al.'s (2003) results, students in the current study who were explicitly taught vocabulary and independent vocabulary learning strategies outperformed students who received typical social studies instruction. In the current study, the vocabulary condition effect translated into a 6-word difference over the typical practice conditions. It is important to note that the vocabulary words on the 30-item curriculum-based assessment were a stratified random sample from the 62 words taught during the 18-week intervention (i.e., 10 words were randomly selected from each 6-week unit's taught vocabulary). Therefore, one might expect a more extensive impact in terms of the number of vocabulary learned by the vocabulary group had the measure included all taught words. The vocabulary difference was not replicated in the comprehension to typical practice comparison and provides support for some level of vocabulary instruction in social studies instruction if vocabulary is a target outcome.

In contrast to the effects of interventions on proximal measures of social studies, results indicated that comprehension and content vocabulary strategies did not produce statistically significant effects on standardized measures of general comprehension or vocabulary. This finding is consistent with prior research where intervention effects on standardized measures were typically more modest in nature than for more proximal measures (Elleman et al., 2009; Pearson et al., 2007). Moreover, this finding is consistent with findings of the large-scale fifth-grade study of supplemental reading comprehension interventions (James-Burdumy et al., 2009). Both experimental conditions did have a substantively important impact on the standardized measure of social studies vocabulary (ES = .31 comprehension and .36 vocabulary; TORC-3 SS). This finding, though not statistically significant provides important, preliminary evidence of the potential impact of both vocabulary and comprehension strategies on a standardized measure of content vocabulary knowledge. The small magnitude of this effect may be attributable, in part, to the fact that the standardized

measure included a broader spectrum of social studies vocabulary and content than addressed in the current interventions.

Our second set of findings provides preliminary information regarding the comparable benefit of multiple-strategy reading comprehension and vocabulary interventions on content-specific and general measures of vocabulary and comprehension. This is one of a limited number of studies that has experimentally compared the effects of different multiple-strategy interventions, particularly comprehension versus vocabulary strategy comparisons. Our overall conclusion is that both conditions were equally effective on the range of measures with the exception of the curriculum-based assessment of vocabulary. Though we speculated that the comprehension condition might show stronger effects on comprehension measures than the content vocabulary condition, their comparable impact was understandable.

When designing both the content vocabulary and comprehension conditions, we drew from prior research to identify instructional elements associated with positive effects (Baumann et al., 2003; Beck et al., 2002; Bryant et al., 2003; Edmonds et al., 2009; Gersten et al., 2001; Kim et al., 2004; Kintsch, 1998; National Reading Panel, 2000; Sencibaugh, 2007; Stahl & Fairbanks, 1986; Taylor et al., 2003). Next, we developed the instructional routines in both conditions to engage students in analyzing and synthesizing content from text. The benefit of content engagement in informational text has been recently documented (McKeown, Beck, & Blake, 2009) and offers a potential explanation for gains in both experimental conditions in this study. In both experimental conditions, students were actively engaged in text, and findings suggest this was sufficient to result in comparable effects in social studies learning across the vocabulary and comprehension conditions. In the comprehension condition, students engaged in a range of evidence-supported strategies (e.g., generating questions, developing main-idea statements, using graphic organizers to summarize text). Likewise, in the vocabulary condition, students not only focused on specific word meanings that were essential to understanding content, they were involved in evaluating and activating prior knowledge through activities that integrated key vocabulary and in summarizing chapter content using targeted vocabulary. For example, target vocabulary were selected for their importance to the chapter content (e.g., *The Exploration of Texas*) and incorporated in semantic organizers that required students to analyze and synthesize information (e.g., "What were Cortez, Coronado, and La Salle's reasons for exploration?"). At the end of chapters, classroom discussions focused on topics that incorporated key vocabulary such as, "I think the explorers in Texas changed the lives of the Native Americans by. . . ." Vocabulary was strategically integrated into questions and summaries and in many instances focused explicitly on integrating vocabulary into important concepts and content. Relatedly, the effect of vocabulary intervention on the social studies content measure is attributable, in part, to the impact of vocabulary knowledge and content comprehension. In the following example item the relevance of vocabulary is clear.

Why was Coronado's expedition not a total failure?

- a. *He did not find any gold or treasure.*
- b. *He discovered Palo Duro Canyon*
- c. *His exploration of the Southwest paved the way for future exploration*
- d. *He would not give up*

Both *expedition* and *exploration* were target vocabulary that were studied and integrated throughout the vocabulary condition.

The only significant difference between the two experimental conditions was on the curriculum-based assessment of vocabulary ($ES = .64$). As expected, students in the content vocabulary condition were more successful in learning and demonstrating knowledge of vocabulary assessed on the proximal measure of social studies vocabulary than those in the comprehension condition, findings that are consistent with prior research (Baumann et al., 2003; Baxter & Reddy, 2007; Lubliner & Smetana, 2005).

Finally, we anticipated that higher levels of teaching quality would be associated with higher levels of comprehension and vocabulary, irrespective of condition. This hypothesis was not supported on the majority of measures. Teaching quality did have a practical effect on both models on the standardized reading comprehension measure. Recent studies document the impact of certain teacher attributes on comprehension, including teacher-managed explicit instruction (e.g., Connor, Morrison, & Petrella, 2004). However, limited prior research has examined quality of instruction and its impact on comprehension or vocabulary outcomes. One notable study by Foorman et al. (2006) found significant but weak effects (i.e., effect accounted for 4% of between-level variance) of ratings of teacher effectiveness on third-grade students' passage comprehension.

In the present study, it is plausible that although ratings of teaching quality were designed to assess effective teaching behaviors, in general, the experimental interventions embedded many quality teaching components in the interventions themselves such as modeling new information, soliciting student responses, and so on. Therefore, it is likely that quality and intervention effects may have not been entirely separable, an outcome that would explain why the typical practice teachers' quality scores were lower than the experimental teachers' scores. Likewise, this may also explain why fidelity scores did not converge in the SEM model in the presence of teaching quality. That is, the overlap between intervention fidelity and quality teaching behaviors was considerable ($r = .76$). Unpacking the impact and interaction of teaching quality, fidelity, and interventions remains a complex matter that warrants further investigation.

Regarding, fidelity of treatment implementation, among teachers who participated in experimental conditions, there was variability within the conditions that need to be examined. In the vocabulary condition, fidelity scores ranged

from 2.00 to 7.00; in the comprehension condition there was a slightly wider range, 1.00 to 7.00. Admittedly, the strategies teachers were asked to implement were complicated instructional procedures; thus, more than the allocated time may be required to attain levels of proficiency and the quality necessary to influence comprehension and vocabulary acquisition. Although a few teachers in the experimental conditions attained high levels of teaching quality (scores of 6 or higher on a 7-point scale), the majority scored in the mid-range (4–5), indicating moderate levels of implementation. The range of quality scores compels us to revisit the professional development schedule and process. Specifically, it may be advantageous to differentiate professional development based on fidelity observations rather than assume that all teachers need comparable amounts and types of instructional support.

LIMITATIONS

The findings of this study must be placed in the context of several limitations focusing on the generalizability of findings, technical adequacy and ecological validity of measures, and adequacy of professional development.

First, the study was conducted in two school districts with a single social studies curriculum; thus, further research is needed to determine whether findings generalize beyond the curriculum and districts. Because strategies were selected from a body of evidence-based practices and then applied to existing curricula, we have no reason to expect that the effects of strategies would differ by district or program. Nonetheless, this limitation warrants further study. Second, the measure of social studies learning was district constructed. Although ecologically valid, indicators of technical adequacy were not established and require further examination. When designing this study, we examined multiple sources to identify measures of social studies learning and found none that measured the content of focus other than the district-constructed measure. This problem pervades comprehension research, particularly with older students (Lang et al., 2009), and places limits on our ability to compare the effects of interventions. Third, we did not include a direct measure assessing whether students actually learned the strategies introduced in interventions. Although Cromley and Azevedo (2007) did not find a significant path linking strategy performance to comprehension, an interim measure that assesses students' acquisition and application of strategies could be potentially important and informative in refining interventions. Finally, the ratings of teaching quality and fidelity were based on summative indices of a combination of observable pedagogical behaviors (e.g., modeling, student engagement, positive feedback). Although similar measures have been used in prior research with documented validity and reliability (Foorman & Schatschneider, 2003), further investigations of content-area instruction should consider how to evaluate the roles of quality and fidelity independent of and in interaction with intervention.

IMPLICATIONS FOR RESEARCH AND PRACTICE

Content-area text poses multiple challenges for students as they progress beyond the primary grades. Research syntheses and guidance documents promote the use of multiple comprehension and vocabulary strategies to address the complexities of content-area text, yet much of the current evidence is based on strategy application in narrative text and limited to studies comparing experimental treatments to typical practice or no-treatment comparison groups (Edmonds et al., 2009). Moreover, the sheer number and range of available strategies is broad, and the tension of how to teach content and strategies within allocated times is real. Teachers of content-area subjects that rely extensively on text-based learning require effective and efficient methods that address the competing yet complementary priorities of comprehension and content.

From this study, we are able to derive implications that contribute to our understanding of content-area instruction and design of future research. Specifically current findings suggest that allocating a portion of instructional time to comprehension or vocabulary strategies that are embedded in social studies text may enhance students' access to and understanding of content-specific text and content vocabulary.

A second and related implication involves how to design interventions that address the complexity of content-area comprehension. Central to this study was the question of how to improve learning from content-area text with the constraints of instructional time. Toward that end, we constructed multiple-strategy combinations that were empirically grounded and conceptually organized, yet perhaps instructionally incomplete. Content-area instruction by design assumes that students become more knowledgeable of the subject matter they are studying. Although current findings indicate that students in the experimental conditions acquired statistically higher levels of content knowledge, few students were able to correctly answer all questions on the district content test. Moreover, findings failed to generalize to standardized measures of comprehension and vocabulary, which may implicate the need interventions of greater duration and comprehensiveness. Teachers reported that they frequently did not have sufficient time to implement the full set of strategies; therefore, further research is needed to evaluate the full potential of experimental interventions and the time necessary to do so.

At issue is the right balance of content and strategy instructions, or perhaps more effective methods of integrating the two. Comprehension strategies that exclude vocabulary instruction may not yield optimal results on the full range of measures. Likewise, vocabulary interventions that do not incorporate comprehension strategies may have diminished impact. Therefore, further research needs to examine the most effective combination of strategies to address the multiple dimensions of social studies text.

Another critical implication from this study is the amount and type of professional development necessary to result in optimal levels of implementation.

Although findings indicated that teaching quality had no significant effect on student outcomes independent of instruction condition, our analyses suggested that few teachers in any condition received the highest possible rating of instructional quality. This finding implicates the amount and type of professional development as a probable factor to examine in future research. Although prior research has documented the importance of distributed professional development (Darling-Hammond, 2000; Garet, Porter, Desimone, Birman, & Yoon, 2001; Joyce & Showers, 1996), limited research has investigated the amount and structure of professional development necessary to implement high-quality instruction in complex content areas. Therefore, future research to improve content-area comprehension will require a multidimensional process that examines the multiple contents (social studies, reading), the instructional contexts, and the professional development process.

ACKNOWLEDGMENTS

The research reported in this article was supported in part by Grant Number R305M050121A, Teacher Quality Reading/Writing, Institute of Education Sciences, U.S. Department of Education. The contents do not necessarily represent the positions or policies of the funding agency.

We gratefully acknowledge the districts, schools, teachers, and students who participated in this study and research staff who served critical roles in data collection and manuscript preparation.

REFERENCES

- Armbruster, B. B., & Anderson, T. H. (1984). Structures of explanations in history textbooks or so what if Governor Stanford missed the spike and hit the rail? *Journal of Curriculum Studies, 16*, 139–161.
- Ausubel, D. P. (1990). *The acquisition and retention of knowledge: A cognitive view*. Dordrecht, The Netherlands: Kluwer Academic.
- Baumann, J. (2009). Vocabulary and reading comprehension: The nexus of meaning. In S. Israel & G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 323–345). 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. (1991). Research on vocabulary instruction: Ode to Voltaire. In J. Flood, J. J. Lapp, & J. R. Squire (Eds.), *Handbook of research on teaching the English language arts* (pp. 604–632). New York: Macmillan.
- Baxter, S., & Reddy, L. (2007). *What content-area teachers should know about adolescent literacy*. Jessup, MD: National Institute for Literacy.

- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York: Guilford.
- Biancarosa, C., & Snow, C. E. (2004). *Reading next: A vision for action and research in middle and high school literacy: A report to the Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.
- Biemiller, A. (2003). Vocabulary: Needed if more children are to read well. *Reading Psychology, 24*, 323–335.
- Boardman, A. G., Roberts, G., Vaughn, S., Wexler, J., Murray, C. S., & Kosanovich, M. (2008). *Effective instruction for adolescent struggling readers: A practice brief*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Boyd, C. D., Gay, G., Geiger, R., Kracht, J. B., Pang, V. O., Risinger, C. F., & Sanchez, S. M. (2003). *Scott Foresman Social Studies Texas*. Glenview, IL: Pearson Education, Inc.
- Brown, V. L., Hammill, D. D., & Wiederholt, J. L. (1995). *The Test of Reading Comprehension* (3rd ed.). Austin, TX: Pro-Ed.
- Bryant, D. P., Goodwin, M., Bryant, B. R., & Higgins, K. (2003). Vocabulary instruction for students with learning disabilities: A review of the research. *Learning Disability Quarterly, 26*, 117–128.
- Carnine, D., Crawford, D. B., Harniss, M. K., Hollenbeck, K. L., & Miller, S. K. (2002). Effective strategies for teaching social studies. In E. J. Kame'enui, D. W. Carnine, R. C. Dixon, D. C. Simmons, & M. D. Coyne (Eds.), *Effective teaching strategies that accommodate diverse learners* (Vol. 2, pp. 177–202). Upper Saddle River, NJ: Pearson.
- Chall, J. S. (1996). *Stages of reading development* (2nd ed.). Fort Worth, TX: Harcourt Brace.
- Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Chen, H. C., & Graves, M. F. (1995). Effects of previewing and providing background knowledge on Taiwanese college students' comprehension of American short stories. *TESOL Quarterly, 29*, 663–686.
- Connor, C. M., Morrison, F. J., & Petrella, J. N. (2004). Effective reading comprehension instruction: Examining Child × Instruction interactions. *Journal of Educational Psychology, 96*, 682–698.
- Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 99*, 311–325.
- Darling-Hammond, L. (2000). Teaching quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives, 8*(1), 1–42.
- Dole, J. A., Valencia, S. W., Greer, E. A., & Wardrop, J. L. (1991). Effects of two types of prereading instruction on the comprehension of narrative and expository text. *Reading Research Quarterly, 26*, 142–159.
- Duke, N. K. (2000). 3.6 minutes per day: The scarcity of expository texts in first grade. *Reading Research Quarterly, 35*, 202–224.
- Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., et al. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79*, 262–300.

- Elleman, A., Lindo, E. J., Murphy, P., & Compton, D. (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), 1–44.
- Espin, C. A., Busch, T. W., Shin, J., & Kruschwitz, R. (2001). Curriculum-based measurement in the content areas: validity of vocabulary-matching as an indicator of performance in social studies. *Learning Disabilities Research & Practice*, 16, 142–151.
- Foorman, B. R., & Schatschneider, C. (2003). Measurement of teaching practices during reading/language arts instruction and its relationship to student achievement. In S. Vaughn & K. Briggs (Eds.), *Reading in the classroom: Systems for observation of teaching and learning* (pp. 1–30). Baltimore: Brookes.
- Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in Grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology*, 31, 1–29.
- 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*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Fukkink, R. G., & de Glopper, K. (1998). Effects of instruction in deriving word meaning from context: A meta-analysis. *Review of Educational Research*, 68, 450–469.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38, 915–945.
- Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, 71, 279–320.
- Graves, M. F. (2006). *The vocabulary book: Learning and instruction*. New York: Teachers College Press.
- James-Burdumy, S., Mansfield, W., Deke, J., Carey, N., Lugo-Gil, J., Hershey, A., et al. (2009). *Effectiveness of selected supplemental reading comprehension interventions: Impacts on a first cohort of fifth-grade students* (NCEE 2009-4032). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Jitendra, A. K., Edwards, L. L., & Sacks, G. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children*, 70, 299–322.
- Johnson, G., Gersten, R., & Carnine, D. W. (1986). Effects of instructional design variables on vocabulary acquisition of LD students: A study of computer-assisted instruction. *Journal of Learning Disabilities*, 20, 206–213.
- Joyce, B., & Showers, B. (1996). Staff development as a comprehensive service organization. *Journal of Staff Development*, 17(1), 2–6.
- Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Slinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide* (NCEE 2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

- Kempe, J., Corrin, W., Nelson, E., Salinger, T., Herrmann, S., & Drummond, K. (2008). *The enhanced reading opportunities study: Early impact and implementation findings* (NCEE 2008-4015). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education.
- 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.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Kline, R. (2005). *Principles and practice of structural equation modeling*. New York: Guilford.
- Klingner, J. K., Vaughn, S., Dimino, J., Schumm, J. S., & Bryant, D. P. (2001). *From clunk to click: Collaborative strategic reading*. Longmont, CO: Sopris West.
- Klingner, J. K., Vaughn, S., & Schumm, J. S. (1998). Collaborative strategic reading during social studies in heterogeneous fourth grade classrooms. *The Elementary School Journal, 99*, 3–22.
- Kuhn, M. R., & Stahl, S. A. (1998). Teaching children to learn word meanings from contexts: A synthesis and some questions. *Journal of Literacy Research, 30*, 119–138.
- Lang, L., Torgesen, J., Vogel, W., Chanter, C., Lefsky, E., & Petscher, Y. (2009). Exploring the relative effectiveness of reading interventions for high school students. *Journal of Educational Effectiveness, 2*, 149–175.
- Lubliner, S., & Smetana, L. (2005). The effects of comprehensive vocabulary instruction on Title I students' metacognitive word-learning skills and reading comprehension. *Journal of Literacy Research, 37*, 163–200.
- MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie Reading Tests* (4th ed.). Itasca, IL: Riverside.
- McKeown, M. G., Beck, I. L., & Blake, G. K. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly, 44*, 218–253.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.
- Okolo, C. M., Englert, C. S., Bouck, E. C., & Heutsche, A. M. (2007). Web-based history learning environments: Helping all students learn and like history. *Intervention in School and Clinic, 43*, 3–11.
- Pearson, P. D., Hiebert, E. H., & Kamil, M. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly, 42*, 282–296.
- Rand Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.
- Raphael, T. (1986). Teaching question answer relationships, revisited. *Reading Teacher, 39*(6), 516–522.
- Raphael, T., Highfield, K., & Au, K. (2006). *QAR now: A powerful and practical framework that develops comprehension and higher-level thinking in all students*. New York: Scholastic.
- Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., et al. (2007). *Reading interventions for adolescent struggling readers: A*

- meta-analysis with implications for practice*. Portsmouth, NH: RMC Research Corporation.
- Sencibaugh, J. M. (2007). Meta-analysis of reading comprehension interventions for students with learning disabilities: Strategies and implications. *Reading Improvement, 44*(1), 6–22.
- Shanahan, T., & Beck, I. L. (2006). Effective literacy teaching for English-language learners. In D. L. August & T. Shanahan (Eds.), *Developing literacy in a second language: Report of the National Literacy Panel* (pp. 415–488). Mahwah, NJ: Erlbaum.
- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*, 72–110.
- Swanson, H. L. (1999). Reading research for students with LD: A meta-analysis of intervention outcomes. *Journal of Learning Disabilities, 32*, 504–523.
- Swanson, H. L., Hoskyn, M., & Lee, C. (1999). *Interventions for students with learning disabilities: A meta-analysis of treatment outcomes*. New York: Guilford.
- Taylor, B. M., Pearson, P. D., Peterson, D. S., & Rodriguez, M. C. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *The Elementary School Journal, 104*, 3–28.
- Texas Education Agency. (2005–2006). *Technical digest* (Texas Education Agency Student Assessment Division). Austin, TX: Author. Retrieved February 18, 2008, from <http://www.tea.state.tx.us/studentassessment/>
- 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.
- Vaughn, S. (2002). Using response to treatment for identifying students with learning disabilities. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 549–554). Mahwah, NJ: Erlbaum.
- Vaughn, S., Klingner, J. K., & Bryant, D. P. (2001). Collaborative strategic reading as a means to enhance peer-mediated instruction for reading comprehension and content-area learning. *Remedial and Special Education, 22*, 66–74.
- What Works Clearinghouse. (2008, December). *Procedures and standards handbook, version 2.0*. Washington, DC: U.S. Department of Education's Institute of Education Sciences. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf
- Willson, V. L., & Rupley, W. H. (1997). A structural equation model for reading comprehension based on background, phonemic, and strategy knowledge. *Scientific Studies of Reading, 1*, 45–63.

APPENDIX

Table A1. Social studies vocabulary by week of introduction

Week 1	Week 2	Week 3	Week 4	Week 5
Unit 1				
Agriculture	Confederacy	Adobe	Conflict	Empire
Artifact	Migration	Pueblo	Treaty	City-state
Culture	Nomad	Irrigate	Battle	Tribute
Descendants	Reservation	Government	Frontier	Slave
Unit 2				
Conquistador	Mission	(Thanksgiving-	Revolution	Right
Colony	Religion	No Lessons)	Immigrant	Convention
Exploration	Villa		Empresario	Dictator
Origin	Presidio		Militia	Protestant
Region	Heritage		Cash crop	Language
			Vaquero	Decree
Unit 3				
Petition	Legend	Congress	Annexation	Boundary
Bill of Rights	Siege	Taxes	Resolution	Expand
Constitution	Surrender	Republic	legislature	Skirmish
Independence		Debt		
Revolution		Monument		
Declaration				
Consultation				

Copyright of Journal of Research on Educational Effectiveness is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.