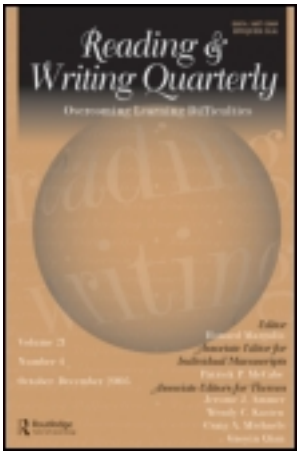

Reading & Writing Quarterly: Overcoming Learning Difficulties
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/urwl20

Examining the Impact of Teacher Quality on Fourth-Grade Students' Comprehension and Content-Area Achievement

Angela Hairrell [a], William H. Rupley [a], Meaghan Edmonds [b], Ross Larsen [a], Deborah Simmons [a], Victor Willson [a], Glenda Byrns [a] & Sharon Vaughn [b]

[a] Texas A&M University, College Station, Texas, USA

[b] The University of Texas at Austin, Austin, Texas, USA

Published online: 25 Apr 2011.

To cite this article: Angela Hairrell, William H. Rupley, Meaghan Edmonds, Ross Larsen, Deborah Simmons, Victor Willson, Glenda Byrns & Sharon Vaughn (2011) Examining the Impact of Teacher Quality on Fourth-Grade Students' Comprehension and Content-Area Achievement, Reading & Writing Quarterly: Overcoming Learning Difficulties, 27:3, 239-260, DOI: 10.1080/10573569.2011.560486

To link to this article: http://dx.doi.org/10.1080/10573569.2011.560486

Reading & Writing Quarterly, 27:239–260, 2011

ISSN: 1057-3569 print/1521-0693 online
DOI: 10.1080/10573569.2011.560486

Routledge
Taylor & Francis Group

Examining the Impact of Teacher Quality on Fourth-Grade Students' Comprehension and Content-Area Achievement

ANGELA HAIRRELL
WILLIAM H. RUPLEY
Texas A&M University, College Station, Texas, USA

MEAGHAN EDMONDS
The University of Texas at Austin, Austin, Texas, USA

ROSS LARSEN
DEBORAH SIMMONS
VICTOR WILLSON
GLENDA BYRNS
Texas A&M University, College Station, Texas, USA

SHARON VAUGHN
The University of Texas at Austin, Austin, Texas, USA

In this study we examined the effects of dimensions of teacher quality on students' comprehension and vocabulary performance. Participants were 36 teachers and their respective 679 students in 2 medium-size school districts in central Texas, both of which served high proportions of children from low–socioeconomic status households. We matched schools in Districts 1 and 2 on the previous year's reading achievement performance and then assigned them to 1 of 2 experimental conditions—comprehension or content vocabulary—through stratified random assignment, with each condition represented in multiple schools in each district. Teachers in each condition participated in a distributed 15 hr of content- and

This study was made possible through U.S. Department of Education, Institute of Education Sciences, Grant R305M050121A (*Enhancing the Quality of Expository Text Instruction and Comprehension Through Content and Case-Situated PD*) to Deborah Simmons, William H. Rupley, and Sharon Vaughn.

Address correspondence to William H. Rupley, Department of Teaching, Learning, & Culture, College of Education and Human Development, 4232 Texas A&M University, College Station, TX 77845, USA. E-mail: w-rupley@tamu.edu

curriculum-based professional development over an 18-week period. The intervention was for 30 min, 3 times a week. We examined the following dimensions of teacher quality: teacher qualifications, instructional practices, quality of strategy use, treatment fidelity, and instructional effectiveness. We used student measures of reading comprehension, content vocabulary, and social studies knowledge to explore the effects of teacher quality. Teachers' education, fidelity, and indicators of teacher quality were significantly related to student outcomes on a standardized measure of reading comprehension.

Effective content-area teachers draw upon their knowledge of reading practices, students' learning capabilities, and pedagogical knowledge to plan and enact quality instruction. In addition to understanding the constructs of reading, effective teachers understand learners and use ongoing assessment to monitor and adjust students' learning experiences through reading (Blair, Rupley, & Nichols, 2007). Effective content-area instruction also requires additional teacher capabilities, such as knowledge of the content (e.g., social studies, science, math) and processing and procedural strategies to assist students with reading and learning from informational text.

It is well established in the literature that the last part of third grade and the beginning of fourth grade are pivotal times for student learning (Blair et al., 2007; Chall, 1996; Gersten, Fuchs, Williams, & Baker, 2001). During this period, informational text assumes a prominent role in students' learning as they are transitioning from learning to read to using reading as a tool for learning (Chall, 1996). As noted by Ippolito, Steele, and Samson (2008), this transition makes adolescent literacy instruction distinctive as well as distinctly challenging. Although this period is referred to as a transition, for many students it becomes a period of stagnation, as they cannot keep pace with the curricular demands. As students move into the later elementary years and continue into the middle grades and high school, content-area textbooks play a larger role in their education and become a primary means of knowledge acquisition in the educational system.

Few content-area teachers have had adequate preparation in teaching students how to negotiate the complex concepts and organization found in content-area textbooks (Baxter & Reddy, 2007; Biancarosa & Snow, 2006). Informational text comprehension is a complex process that requires students to successfully manipulate and integrate information from multiple sources while engaged in reading (Kolligian & Sternberg, 1987). Informational texts pose particular challenges, as students are reading longer passages (Bereiter & Scardamalia, 1987), processing abstract logical–causal arguments (Stein & Trabasso, 1981), and making sense of the complicated and varied text structures (Kucan & Beck, 1997). Furthermore, the density

of information and unfamiliar technical vocabulary demand that students implement strategies to extract, summarize, and synthesize the content (Lapp, Flood, & Ranck-Buhr, 1995).

Many students are unable to use expository texts as a learning tool because they lack the strategies to make expository text comprehensible (Weaver & Kintsch, 1991). By fourth grade, students are expected not only to acquire content knowledge through reading but also to use cognitive strategies to learn from reading informational textbooks. Examples of such strategies are activating background knowledge, using text structures and organizations, analyzing and synthesizing main ideas, learning new vocabulary through context, drawing and supporting conclusions with evidence from text, and monitoring comprehension.

Although there is considerable agreement about the skills and strategies adolescent readers need to comprehend and learn from content-area text, at issue is how to effectively prepare teachers to incorporate these strategies into their content-area instruction. What methods of professional development (PD) prepare content-area teachers to effectively teach students to incorporate these strategies into their reading? Experimental studies evaluating the application of vocabulary and comprehension strategies in content areas have been limited (Baumann, Edwards, Boland, Olejnik, & Kame'enui, 2003; Gersten & Okolo, 2007; Okolo, Englert, Bouck, & Heutsche, 2007). In addition, few studies have experimentally examined the qualities of teachers who effectively integrate reading strategies into their teaching of content-area classes.

In the present study, we examined the effects on students' learning of a PD model designed to promote teachers' teaching of reading comprehension strategies and vocabulary during social studies. We also examined the role of teacher qualifications and teacher practice.

DIMENSIONS OF TEACHER QUALITY

Teacher quality is a multidimensional construct that includes a range of variables, including teacher qualifications, teacher characteristics, teacher practices, and teacher effectiveness (Goe & Stickler, 2008). This section reviews the literature on teacher qualifications and teacher practices, with particular emphasis on those dimensions of teacher quality that are malleable through PD.

Teacher Qualifications

Teacher qualifications include many variables, such as subject-level knowledge, advanced degrees, certification exam scores, degree-granting institutions, certification, and content-based pedagogical knowledge. For content-area teachers, the qualifications of subject area expertise and pedagogical knowledge would seem theoretically to be of primary importance; however, little research has

investigated the relation of these qualifications to student outcomes. Monk (1994) found that both teachers' mathematical knowledge and their content-based pedagogical knowledge were positively associated with student outcomes. Other studies have reported similar findings that teachers content knowledge (Aaronson, Barrow, & Sanders, 2003; Wenglinsky, 2000, 2002) and pedagogical knowledge (Hill, Rowan, & Ball, 2005) are factors associated with students' learning in mathematics and science. Content and pedagogical knowledge were noted by Schacter and Thum (2004) as one of 12 teacher variables that positively influenced student outcomes in reading, mathematics, and language.

In addition to teacher knowledge, several teacher practices have been shown to be associated with student learning outcomes. These practices are instruction aligned with assessment, objectives that are clearly stated, modeling and think-alouds, progress monitoring, and active learning in reading and math (Goe, 2007; Kimball, White, Milanowski, & Borman, 2004; Smith, Lee, & Newmann, 2001). Learning objectives clearly stated and understood by the students were linked to increased student achievement in reading and mathematics at both the middle and elementary school levels (Matsumura et al., 2006; Schacter & Thum, 2004). In addition, the use of progress monitoring (Schacter & Thum, 2004; Wenglinsky, 2000, 2002) and active learning (Frome, Lasater, & Cooney, 2005; Smith et al., 2001; Wenglinsky, 2000, 2002) were linked with higher student achievement. Alignment of the curriculum and assessments and the use of active teaching methods were found to have effect sizes ranging from 0.09 to 0.18 for reading in Grades 1–6 (Rowan, Correnti, & Miller, 2002).

CHARACTERISTICS OF EFFECTIVE PD

Most educators agree that a fundamental outcome of PD should be improvement in student achievement (Banilower, Heck, & Weiss, 2007). Although student achievement outcomes are the gold standard (Blanton & Kaput, 2005) of PD, it is generally assumed that changes in teacher knowledge and instructional practices precede and strongly influence student learning (Harris & Sass, 2007).

PD can be a primary method of supporting teachers' knowledge and practice and increasing their quality of instruction. There is emerging evidence on what qualities and goals of effective PD are most beneficial to increasing students' achievement (Ball & Cohen, 1996; Darling-Hammond, 2000; Garet, Porter, Desimone, Birman, & Yoon, 2001; Lawless & Pellegrino, 2007; Penuel, Fishman, Yamaguchi, & Gallagher, 2007; Yell, 2006). Evidence suggests that effective PD is intensive, distributed over time, content focused, and anchored to student learning (Borasi & Fonzi, 2002; Garet et al., 2001; Penuel et al., 2007). Recent studies have indicated that effective PD programs

are content driven and explicitly focused on student achievement. Research notes the significance of providing ongoing support for teachers through mentoring, peer group planning and discussions, and lesson studies (Belzer, Drennon, & Smith, 2001; National Reading Panel, 2000). In a large-scale empirical study of more than 1,000 teachers (Garet et al., 2001), the elements of PD on teachers' learning were evaluated. Three core features of PD were identified that resulted in significant and positive effects on teachers' self-reported content knowledge: a focus on content knowledge, opportunities for active learning, and coherence with other learning activities.

In a recent study of 454 science teachers, Penuel and colleagues (2007) found that certain elements of PD were associated with implementation. For instance, they found that PD that provided "opportunity to localize" or opportunity to "tailor the curriculum to teachers" was significantly related to program implementation. PD activities that engaged teachers in aligning their subject content with national, state, and regional standards and that were consistent with curricular goals were also highly associated with program implementation. One of their key conclusions was that there must be a "fit" (p. 952) between the PD and the local context. Penuel and colleagues referred to this as *coherence* and found that the perceived coherence of the PD and district goals was a strong predictor of program implementation. Furthermore, when curriculum concerns were addressed and time was provided for planning, implementation of new practices increased (Spillane, 1999).

PD that addresses content knowledge in union with pedagogical knowledge is well supported (Hill et al., 2005; Knapp, 2003; Supovitz & Turner, 2000) in both teacher program implementation and students' achievement. Cohen and Hill (1998) reported that the consequence of such PD is that it deepens teachers' subject matter knowledge while simultaneously increasing their pedagogical proficiency. Similar results have been reported in PD studies with math and science teachers (Hill et al., 2005; Schacter & Thum, 2004; Supovitz & Turner, 2000).

PURPOSE

The purposes of this research were (a) to explore relations among dimensions of teacher quality and student achievement among groups of fourth-grade social studies teachers who participated in vocabulary and comprehension PD and (b) to evaluate the feasibility of integrating vocabulary and comprehension strategies within the context of fourth-grade social studies.

Research questions guiding this inquiry were as follows:

1. What influences do teacher qualifications, characteristics, quality, and practices have on student reading comprehension, vocabulary development, and content knowledge?

2. What are teachers' perceptions of the feasibility and practicality of integrating a vocabulary or comprehension intervention into their social studies instruction?

METHODS

Students

This study was conducted in two medium-size school districts in central Texas. All 5 elementary schools in District 1 and 6 of 15 elementary schools in District 2 participated. School personnel identified the six schools in District 2 as candidates for the study based on third-grade reading performance on state achievement tests. In District 1, 77% of the students qualified for the free or reduced-price lunch program; 65% qualified in District 2. School demographics were obtained. Schools were matched both within and across the districts based on student ethnic composition and randomly assigned to one of two treatment conditions of social studies instruction: comprehension or content vocabulary. Primary matches were made based on either predominant African American, Latino, or Anglo composition with reasonable correspondence and on the previous year's reading achievement performance on the Texas Assessment of Knowledge and Skills (TAKS; Texas Education Agency, 2005–2006), with each condition represented in multiple schools in each district.

Informed consent was obtained for 679 fourth-grade students from 36 participating social studies classes, with 17 classes in the comprehension (C) condition and 19 in the vocabulary (V) condition. Demographic information on the students is provided in Table 1. Students were ethnically and racially diverse, with average percentages across all groups of 16.9% African American (C = 19.7%, V = 14.1%), 67.0% Hispanic (C = 66.2%,

TABLE 1 Student Demographic Information by Condition Prior to Professional Development

	Comprehension		Vocabulary	
Variable	*n*	%	*n*	%
Number	309		306	
Gender				
Male	153	49.5	152	49.7
Female	156	50.5	154	0.3
Ethnicity				
African American	60	19.4	43	14.1
Hispanic	202	65.4	207	67.6
Euro-American	39	12.6	49	16.0
Other	4	2.6	5	2.3
Qualify for special education	3	0.01	5	0.02

V = 67.9%), and 14.4% Euro-American (C = 12.8%, V = 16.1%). Note that 1% of the students reported other races/ethnicities; 1.4% received special education services (C = 1.1%, V = 1.7%). Group equivalence was evaluated using analyses of variance for the continuous variables (e.g., age) and chi-square tests for the categorical variables (e.g., gender, ethnicity, special education services). In addition, pretesting on the Gates-MacGinitie Reading Tests (GMRT) Fourth Edition, Comprehension Subtest; the Test of Reading Comprehension–Social Studies; and the Test of Silent Contextual Reading was all nonsignificant at an alpha of .05: $F(1, 678) = .08$, $p = .78$; $F(1, 678) = .59$, $p = .45$; and $F(1, 678) = 1.02$, $p = .31$, respectively; and is considered in the subsequent data analysis.

Teachers

A total of 36 fourth-grade social studies teachers participated in the study during the 2006–2007 school year. The number of years of teaching experience averaged 5.72 in the comprehension group and 9.18 in the vocabulary group. Table 2 provides teacher demographic information, including gender, race/ethnicity, teaching experience, degrees earned, and certification. Chi-square analyses conducted on the categorical variables and one-way analyses of variance conducted on the continuous variables to assess teacher equivalence indicated no significant differences among groups on gender, years of teaching experience, years teaching social studies, or highest degree

TABLE 2 Teacher Demographics by Condition Prior to Professional Development

	Comprehension		Vocabulary	
Variable	*n*	%	*n*	%
Number	17		19	
Gender				
Male	2	11.8	0	0.0
Female	15	88.2	19	100.0
Ethnicity				
African American	5	29.4	0	0.0
Hispanic	4	23.5	4	21.1
Euro-American	7	41.2	15	78.9
Education				
Bachelor's	16	94.1	16	84.2
Master's	1	5.9	3	15.8
Certifications				
Bilingual	4	23.5	4	21.1
Special education	0	0.0	2	10.5
Organization				
Self-contained	12	70.6	9	47.4
Departmentalized	5	29.4	10	52.6
Experience				
Mean years	5.72		9.18	
Mean years in fourth grade	3.22		4.63	

earned. There was a significant relation among groups on ethnicity, $\chi^2(1, N = 36) = 7.709$, $p = .02$, $df = 2$, with a higher percentage of African American teachers in the comprehension condition.

MEASURES

We examined three sets of variables associated with teacher quality that emerged from our literature review: (a) characteristics and qualifications, (b) instructional practices and fidelity, and (c) overall teacher quality. Several measures were used to collect information about these variables.

Measures of Teacher Characteristics and Qualifications

Prior to PD, teachers completed a demographic survey with questions related to ethnicity, professional training, and degrees. Gender and ethnicity were used as indicators describing invariable teacher characteristics. Teacher qualification variables included degrees held and a self-report measure of perceived reading strategy knowledge. For the knowledge perceptions measure, teachers were asked to rate their knowledge of specific comprehension and vocabulary strategies at the conclusion of the PD. Teachers rated their knowledge on a scale of 1 (*not at all familiar*) to 3 (*very familiar*). Responses were summed to create a perceived knowledge score (10 items, for a maximum score of 30). See Table A-1 in the Appendix for sample items.

Measures of Instructional Practices

Members of the research team used a fidelity protocol to code each recorded lesson for both the presence and quality of the comprehension and vocabulary intervention components (e.g., generating and answering questions, introducing concept-laden words, providing multiple exposures to vocabulary words). The protocol was adapted from similar instruments used in other research studies (Vaughn, 2002; Vaughn, Klingner, & Bryant, 2001) and included six items, each of which was rated on a 4-point scale ranging from 0 (*not at all*) to 3 (*exemplary quality of implementation*). For this analysis, the rating on the target strategy for the coded lesson was used. In other words, because strategy instruction and practice were distributed so that a new target strategy was introduced each 6 weeks, we used the fidelity rating for the targeted strategy and not previously taught strategies when creating this variable. Teachers also completed weekly online logs to document the amount of time that they taught social studies. Comprehension teachers reported an average of 106.1 min weekly, whereas vocabulary teachers reported 119.3 min total (ranges = 57.5–151.7 min for comprehension, 61.3–194.6 min for vocabulary).

Measures of Teacher Effectiveness

We conceptualized teacher effectiveness in two ways: the quality of instruction and student achievement. Quality of instructional lessons was determined through coding audio-recorded intervention implementation lessons. Researchers coded one lesson from each 6-week social studies unit, for a total of three recorded lessons. Researchers coded two randomly selected recordings of instruction for each teacher (approximately 34 recordings). Interrater reliability, calculated as the number of agreements divided by the number of agreements plus disagreements across all items, was .88 prior to coding independently. Moreover, 20% of lesson recordings were double-coded, with reliabilities ranging from .75 to 1.00. When there was a discrepancy between coders, the questionable item was rated by an expert coder.

Overall quality of instruction was rated using a 7-point scale ranging from 1 (*not at all effective*) to 7 (*highly effective*; Foorman & Schatschneider, 2003; Taylor, Pearson, Peterson, & Rodriguez, 2003). When rating teacher quality, coders considered generally accepted features of effective teaching supported by research and theory (Blair et al., 2007; Heilman, Blair, & Rupley, 2002; Margolis & Brannigan, 2009), such as lesson pacing, corrective feedback, and explicitness of instruction. Overall teacher quality ratings did not consider intervention fidelity. Teachers could have been rated high on teacher quality even if they did not implement the intervention with 100% fidelity.

Student achievement was measured using a standardized comprehension measure and curriculum-based measures of vocabulary and content knowledge. In addition, the state reading assessment was used as a covariate measure of prior reading achievement.

The GMRT–4 Passage Comprehension Subtest (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) was administered to classes of students pre- and postintervention and served as a measure of general comprehension achievement. The GMRT–4 is a group-administered 35-min 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-forms reliability ranges from .80 to .87 in national samples.

A curriculum-based vocabulary measure (CBM–V) was constructed by the researchers to assess students' knowledge of social studies vocabulary taught as part of the lessons featured in the PD. The 5-min timed CBM–V was administered pre- and postintervention and required students to match vocabulary words presented in the first column with the correct definition of the word in the second column. Words and definitions were randomly ordered. The measure was modeled after commonly used curriculum-based vocabulary measures reported in previous research (Espin, Busch, Shin, & Kruschwitz, 2001). Reliability averages in this study were approximately .70.

A curriculum-based measure of social studies content was developed by a task force of social studies teachers from School District 2 and administered by teachers in all classrooms at the end of each 6-week period (three administrations) to assess mastery of taught social studies content. These measures were untimed and comprised 10 multiple-choice items (four options) that focused on social studies content. Each item had a value of 1 point. For the purpose of this analysis, the results of these three individual assessments were summed to create a content knowledge score. Reliability within treatment was .78.

The TAKS Grade 3 reading score was obtained for all students and used as a covariate in statistical analyses. The TAKS is administered to all Grade 3 students and is a measure of general reading achievement. Internal consistency reliability coefficients range from the high .80 s to the low .90 s (Texas Education Agency, 2005–2006).

Measures of PD Efficacy Perceptions and Instructional Feasibility

All PD teachers provided feedback on the effectiveness of the strategies, the impact of strategies on students' learning and engagement, and the influence of PD on their social studies teaching at the conclusion of each case (the end of a 6-week unit of instruction). Surveys were a mix of open-ended and Likert-type scale items. Likert items addressed the effectiveness of the strategies, the impact of strategies on students' learning and engagement, and the influence of PD on teachers' social studies teaching. Open-ended items addressed feedback on intervention components and anecdotes from the classroom. Following completion of their 18-week PD, teachers completed a fourth online survey on their perceptions of their effectiveness in teaching social studies, the effectiveness of the overall PD, and the effectiveness of the intervention.

PD Model

Teachers participated in either reading comprehension or vocabulary activities in the content- and curriculum-based PD. Although the content knowledge shared during these sessions differed (one focused on reading comprehension strategies and one on vocabulary strategies), the instructors and delivery processes were the same. Three tenets guided the PD model for both strategy areas. The first tenet referred to the idea that the PD should be grounded in the actual work of teachers. For this reason, intervention materials were guided by state standards and also met district standards. The second tenet emphasized that instructional strategies presented in the PD sessions should be designed using district-adopted materials and integrated into the social studies content. The third tenet required that the presentation and content of the PD be based on evidence-based best practices

(e.g., the comprehension and vocabulary strategies were based on empirical evidence). Because presentations were hierarchical, they required the integration of both procedural knowledge and conceptual knowledge, meaning that as each case progressed, activities built on previous activities and supported the presentation and assimilation of more advanced work. Other supports were provided through online resources and direct access to research staff to answer questions.

Principal instructional components of the PD model incorporated the following:

- Vocabulary evidence-based strategies that included methods for (a) identifying and prioritizing critical vocabulary, (b) previewing and connecting vocabulary to the bigger concepts and content in a social studies unit and chapter by using graphic organizers, (c) using a vocabulary map to illustrate relationships among words, (d) activating and building background knowledge, (e) using context to learn meanings of unfamiliar vocabulary, and (f) using strategies to assess and maintain vocabulary knowledge.
- Comprehension evidence-based strategies for (a) teaching students to ask more advanced levels of questions and (b) enabling students to create summary statements (gist statements) of informational/expository text.
- Teacher implementation of PD instructional strategies for (a) modeling learning activities and (b) using graphic organizers (unit overviews, chapter overviews, vocabulary maps, question cards, and game cards) that utilized explicit and systematic teaching of concepts.

This model of PD is an integrated and supportive one in which each component builds on and reinforces the other components. Table A-2 in the Appendix provides an overview of the instructional content for each case.

The PD consisted of three face-to-face combined sessions. These three sessions were implemented as cross-district sessions to ensure consistency of information. The initial session was approximately 5 hr in length and occurred prior to teachers teaching the first instructional unit. The second and third sessions were after school and lasted approximately 2 hr. In addition to the combined PD sessions, teachers met in Teacher Study Teams (TSTs). TSTs were small groups of teachers from the same instructional condition (vocabulary or comprehension) and included all teachers from two schools in the same district. The TSTs were 90-min after-school sessions and occurred at the midpoint of each 6 weeks of PD implementation. The purposes of TSTs were to identify what was working, examine student performance, review relevant research, learn new instructional methods, and practice previously learned methods.

DATA ANALYSIS

A two-level structural equation model was developed to account properly for both student-level pretest and posttest data and teacher-level variables associated with the PD interventions conducted in the study. We developed a model based on the characteristics of teacher quality identified by Goe and Stickler (2008). Measures of teacher qualifications included teachers' highest degree earned, years of teaching experience, and perceived knowledge score. Teacher practice was represented by teachers' use of PD strategies and a fidelity measure. Teacher quality was based on an overall measure of teacher quality. Prior analysis suggested that there was no difference between the vocabulary or comprehension PD conditions in regard to the outcome measures and therefore, these two PD conditions could be combined for the analysis and are referred to as the PD condition.

At the student level, the pretest score for each outcome measure was used as the most proximate covariate for the outcome to account for initial differences in the outcome along with each student's TAKS reading score from third grade. The two-level model was analyzed using MPLUS 5.2 (Muthén, & Muthén, 1998–2007).

To complement the analysis, we performed a qualitative analysis of data from TSTs, instructional proficiency forms, and open-ended survey items. All data sources provided information on teachers' perceptions of the intervention's efficacy and implementation. Using an approach consistent with grounded theory (Strauss & Corbin, 1998), we first identified categories that represented the major themes of comments across the sources. We then grouped teacher responses into these categories and summarized those that were most prevalent. We did not attempt to quantify or tally responses because not all teachers answered every question on the forms. Rather, our findings represent common themes from obtained responses.

RESULTS

Correlations for all independent variables and outcome measures can be found in Tables 3 and 4. The majority of the significant correlations were low, ranging from.34 to.40. Of the correlations among the dependent variables, only the relationship between fidelity and teacher quality was statistically significant at the .01 level, with $r = .73$. Table 5 reports both standardized and unstandardized regression coefficients. Figure 1 depicts the structural equation model. All fit statistics fell within the acceptable ranges: $\chi^2/df = 4.56$, comparative fit index = .94, Tucker–Lewis index = .76, root mean square error of approximation = .10, standardized root-mean-square residual (within) = .004, standardized root-mean-square residual (between) = .16. Within-level (student-level) results showed that all posttest

TABLE 3 Pearson Correlations of All Dependent Variables

Variable	1	2	3	4	5	6	7
1. Education	—	.30	.38	.09	.24	.12	−.29
2. Experience	.30	—	.01	.20	.08	.17	.20
3. Teacher knowledge	.38	.01	—	−.24	−.08	−.09	−.48*
4. Quality of strategy use	.09	.20	−.24	—	.49*	.17	.46*
5. Fidelity	.24	.08	−.08	.49*	—	.73**	−.15
6. Teacher quality	.12	.17	−.09	.17	.73**	—	−.19
7. Professional development content	−.29	.20	−.48*	.46*	−.15	−.19	—

*$p < .05$ (2-tailed). **$p < .01$ (2-tailed).

measures were predicted by pretest measures. The within-level covariates explained 12% to 52% of the outcome variance.

Model Results

We modeled the characteristics identified by Goe and Stickler (2008) to our sample of teachers and students to explore the relationship of these characteristics to student achievement in reading comprehension, vocabulary knowledge, and content knowledge. Three variables representing teacher qualifications included degree, experience, and perceived knowledge. Teacher knowledge and years of experience were not significant predictors of any of the outcome measures. However, degree was a significant predictor, with students whose teachers held a master's degree outperforming those whose teachers held a bachelor's degree on the GMRT–4 ($\beta = -.29$, $B = -3.92$, $p < .05$). Our model included two measures of teacher practice (i.e., quality of strategy use and fidelity), one of which was significant. The path from fidelity to GMRT–4 resulted in a standardized path coefficient of −.55 ($B = -2.30$, $p < .04$). Overall teacher quality was positively related to increased achievement on the GMRT–4 ($\beta = .64$, $B = 2.28$, $p < .005$). Given that the PD conditions did not differ in terms of teacher quality, fidelity, or dosage, we can conclude that the higher a teacher was in teacher quality and fidelity in either PD condition (vocabulary or comprehension), the higher score the students scored on the GMRT–4. None of the variables were significant predictors of either

TABLE 4 Pearson Correlations of Independent Variables

Variable	1	2	3
1. Post-GMRT–4	—	.30**	.30**
2. Post-CBM–V	.30**	—	.43**
3. Social studies content	.30**	.43**	—

Note. GMRT–4 = Gates-MacGinitie Reading Tests (GMRT) Fourth Edition, Comprehension Subtest; CBM–V = curriculum-based vocabulary measure.
**$p < .01$ (2-tailed).

TABLE 5 Between-Level Model Estimates for the Structural Equation Model

Path (path number from Figure 1)	*B*	*SE*	*β*	*p*
Education				
Post-GMRT–4 (1)	–3.92	.20	–.29	.05
Post-CBM–V (2)	–0.18	.19	–.08	.66
Social studies content (3)	1.88	.22	.26	.24
Experience				
Post-GMRT–4 (4)	–0.10	.20	–.13	.51
Post-CBM–V (5)	0.14	.20	.17	.38
Social studies content (6)	0.10	.29	.25	.39
Teacher knowledge				
Post-GMRT–4 (7)	–0.48	.21	–.28	.17
Post-CBM–V (8)	0.40	.25	.23	.36
Social studies content (9)	0.42	.45	.46	.31
Quality of strategy use				
Post-GMRT–4 (10)	0.06	.19	.02	.91
Post-CBM–V (11)	–0.71	.27	–.25	.34
Social studies content (12)	0.27	.31	.18	.56
Fidelity				
Post-GMRT–4 (13)	–2.30	.15	–.55	.04
Post-CBM–V (14)	–0.20	.36	–.05	.90
Social studies content (15)	1.03	.44	.46	.30
Teacher quality				
Post-GMRT–4 (16)	2.28	.12	.64	.005
Post-CBM–V (17)	0.72	.34	.20	.56
Social studies content (18)	0.12	.56	.06	.91
Professional development				
Post-GMRT–4 (19)	0.43	.23	.05	.85
Post-CBM–V (20)	–0.39	.20	–.39	.21
Social studies content (21)	0.83	.47	.16	.73

Note. GMRT–4 = Gates-MacGinitie Reading Tests (GMRT) Fourth Edition, Comprehension Subtest; CBM–V = curriculum-based vocabulary measure.

the CBM–V or the social studies content test. All path weights can be found in Table 5. *R*-squares for the outcomes measures are as follows: GMRT–4 = .83, CBM–V = .49, social studies content = .43.

Teachers' Perceptions of PD Feasibility and Student Engagement

Implementing the PD practices

The most prevalent and consistent implementation challenge reported by teachers was an inadequate amount of time to fully implement the PD practices in their teaching. Reported time constraints included the ability to complete all PD lesson components within the suggested lesson length and the limited amount of time allocated to social studies teaching. Time constraints for social studies teaching resulted from external pressure to spend time in those subject areas on which students would be assessed on the upcoming state assessments.

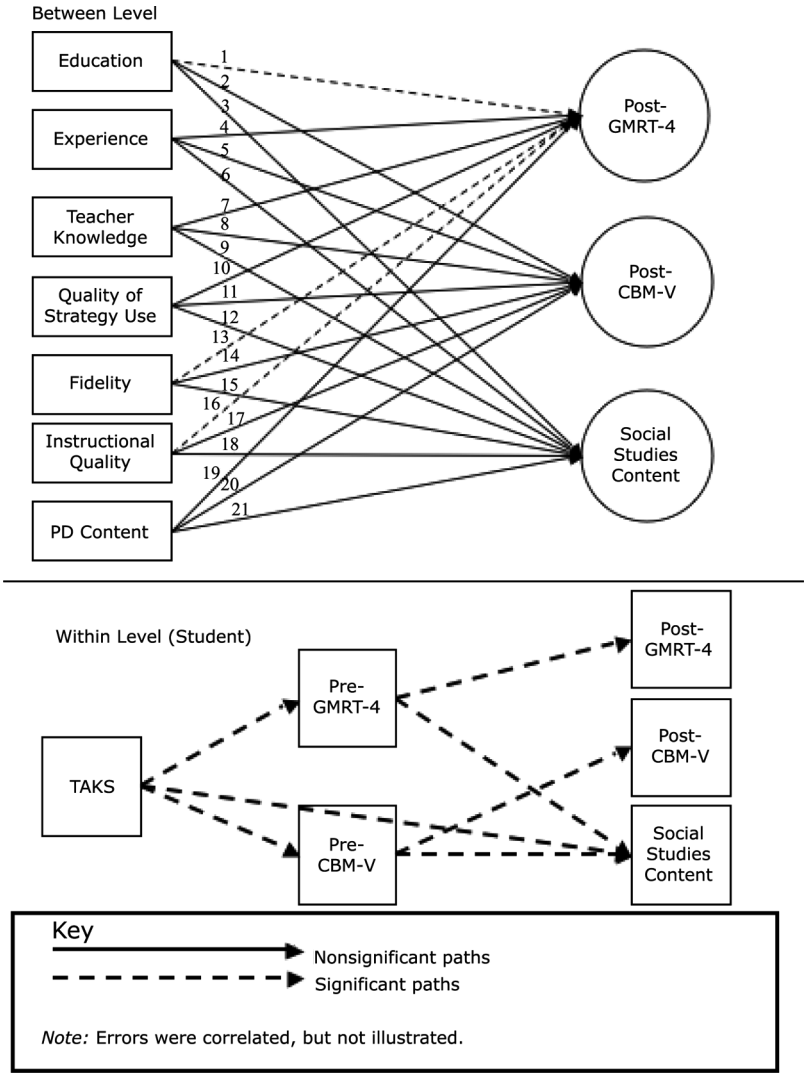

FIGURE 1 Structural equation model of characteristics of teacher quality. GMRT–4 = Gates-MacGinitie Reading Tests (GMRT) Fourth Edition, Comprehension Subtest; CBM–V = curriculum-based vocabulary measure; PD = professional development; TAKS = Texas Assessment of Knowledge and Skills.

DISCUSSION AND CONCLUSIONS

This study examined the impact of dimensions of teacher quality on student achievement among a group of teachers who participated in an 18-week PD/curriculum implementation experience. Results of structural equation modeling indicated that three teacher variables were related to student performance on a standardized measure of reading comprehension.

Three variables in our model showed a positive relationship to the reading comprehension measure. As identified in the review by Goe and Sticker (2008), although education has been shown to impact student outcomes in math and science, these findings have not been replicated in reading or social studies. Our findings show that education may be a factor in teacher quality for social studies. This is plausible considering the vast amount and breadth of information found in the social studies curriculum.

Teacher quality and fidelity were both positively related to student achievement on the standardized measure of reading comprehension. Teacher quality was related to increased student achievement on a standardized measure of reading comprehension. This is no small feat, given that standardized measures result in much smaller effect sizes than researcher-developed measures (Elleman, Lindo, Morphy, & Compton, 2009). Our measure of teacher quality was based on the work of Foorman and Schatschneider (2003) and was consistent with the characteristics of teacher quality identified by others (Blair et al., 2007; Heilman et al., 2002; Margolis & Brannigan, 2009). Specifically, the teacher's use of corrective feedback, instructional pacing, and level of student engagement were among characteristics of teacher quality that were rated.

The teachers in this study who adhered more closely to the PD materials had a greater impact on student achievement than those who did not. The PD focused on evidence-based practices for vocabulary and comprehension instruction, as well as general effective instructional practices. Fidelity in our study was not exceptionally high, with an average of 4.7 on a scale of 1 to 10. The role of fidelity in the interpretation of findings and the importance of the teacher in maintaining fidelity to the treatment are critical in research (Hulleman & Cordray, 2009).

Together, these two findings reflect an important idea. Quality instructional materials combined with quality delivery can enhance student learning on standardized measures. It is important to address the lack of findings on researcher-developed measures. The social studies content measure was a district-developed measure. Although it was closely aligned to the instructional content, the reliability of this measure is questionable. The CBM–V measure did not result in significant results. This may be surprising, considering that one PD had a vocabulary focus. A possible explanation may be that the structural equation model, which combined many teacher and instructional variables, accounted for any possible variance. PD combining best practices for comprehension and vocabulary instruction and components of effective instruction may lead to increased student achievement. Although the present study is in no way definitive, it does pave the way for future studies to examine the relationship of PD content, classroom instruction, and student outcomes.

Although many teachers noted that they had used some of the strategies prior to participating in this study, most indicated that they had not taught

them explicitly or used them in combination in their teaching of social studies. Teachers often characterized the program featured in the PD as a break from the "traditional" approach to social studies instruction, including being more "in the text" than they were before their PD involvement.

Teachers in the vocabulary condition reported that the vocabulary maps were the most innovative and pedagogically influential component. Teachers in the comprehension condition found the gist strategy to be effective at helping students make sense of the social studies text. Teachers in the comprehension condition found some strategies difficult for students to master, especially the questioning strategy. In addition, many teachers in the vocabulary condition reported the decreased emphasis on comprehension to be the biggest change in their teaching. The intervention materials and structure appeared to aid implementation, and as a result of participating, some teachers indicated that they spent more time on social studies than in years past. In addition, several teachers indicated that they would continue to use the practices after the study ended. Yet despite the increased time, many teachers felt unable to cover as much content as they had when they used teacher-led discussions to teach social studies.

When asked how their students reacted to the lessons featured in the PD, teachers most often referred to the behavioral and social responses of their students, citing increased student engagement and on-task behavior. In addition, teachers reported that the materials and activities motivated students and piqued their interest in social studies. Teachers used words such as *fun, eager*, and *enjoyed* to describe their students' reactions to the instructional PD program components. However, some teachers reported that students tired of the instructional routines after several weeks. The most commonly reported academic changes as a result of the new practices included improved comprehension and recall of content and vocabulary. In addition, some teachers reported an awareness that some students were beginning to apply the strategies or use the vocabulary words in other curricular areas.

What is interesting is that many teachers reported that they felt that the practices were having a positive influence on students' comprehension and vocabulary knowledge in social studies. The analyses support their perceptions of students' vocabulary learning but not of students' comprehension abilities, suggesting that our measures may not have been sensitive to the changes observed by teachers. Perhaps a more proximal measure of students' social studies learning focusing on weekly content would have been more sensitive to students' comprehension growth.

The role of teacher quality in this study was most interesting. The PD significantly influenced overall teacher quality, and levels of teacher quality were also significantly related to vocabulary learning. Our analyses provide evidence that participation in high-quality PD is associated with higher ratings of teacher quality. The significant relationships between PD and overall teacher quality in the full model and between the vocabulary strand and

strategy instruction quality in the partial model indicates that a PD model that provides teachers with the knowledge and skills to incorporate effective strategy and vocabulary instruction into their social studies instruction can improve the quality with which instruction is delivered.

REFERENCES

Aaronson, D., Barrow, L., & Sanders, W. (2003). T*eachers and student achievement in the Chicago public schools*. (Working Paper No. 2002–28). Chicago, IL: Federal Reserve Bank in Chicago.

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher*, *25*(9), pp. 6–8, 14.

Banilower, E. R., Heck, D. J., & Weiss, I. R. (2007). Can professional development make the vision of the standards a reality? The impact of the National Science Foundation's local systemic change through teacher enhancement initiative. *Journal of Research in Science Teaching*, *44*, 375–395.

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.

Baxter, S., & Reddy, L. (2007). *What content-area teachers should know about adolescent literacy*. Jessup, MD: National Institute for Literacy.

Belzer, A., Drennon, C., & Smith, C. (2001). *Building professional development systems in adult basic education: Lessons from the field*. Retrieved from http://www.ncsall.net/?id=559

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.

Biancarosa, C., & Snow, C. E. (2006). *Reading next: A vision for action and research in middle and high school literacy: A report to the Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Blair, T., Rupley, W. H., & Nichols, W. D. (2007). The effect teacher of reading: Considering the "what" and "how" of instruction. *The Reading Teacher*, *60*, 432–438.

Blanton, M. L., & Kaput, J. J. (2005). Characterizing a classroom practice that promotes algebraic reasoning. *Journal for Research in Mathematics Education*, *36*, 412–446.

Borasi, R., & Fonzi, R. (2002). *Professional development that supports school mathematics reform. Foundations, Volume 3: A monograph for professional development in science, mathematics, and technology education*. Arlington, VA: National Science Foundation, Directorate for Education and Human Resources, Division of Elementary, Secondary, and Informal Education.

Chall, J. S. (1996). *Stages of reading development* (2nd ed.). New York, NY: Harcourt Brace College.

Cohen, D. K., & Hill, H. C. (1998). *Instructional policy and classroom performance: The mathematics reform in California*. (CPRE Research Report No. RR-39). Philadelphia, PA: Consortium for Policy Research in Education.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives*, *8*(1), 1–42.

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 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, MD: Brookes.

Frome, P., Lasater, B., & Cooney, S. (2005). *Well-qualified teachers and high-quality teaching: Are they the same?* (Research Brief). Atlanta, GA: Southern Regional Education Board.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes PD 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.

Gersten, R., & Okolo, C. M. (2007). Teaching history in all its splendid messiness to students with LD: Contemporary research. *Journal of Learning Disabilities*, *40*(2), 98–99.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.

Goe, L., & Stickler, L. M. (2008). *Teacher quality and student achievement: Making the most of recent research*. Washington, DC: National Comprehensive Center for Teacher Quality.

Harris, D., & Sass, T. (2007). *Teacher training, teacher quality, and student achievement* (Working Paper No. 3). Washington, DC: National Center for Analysis of Longitudinal Data in Education Research.

Heilman, A. J., Blair, T. R., & Rupley, W. H. (2002). *Principles and practices of teaching reading* (10th ed.). Columbus, OH: Merrill.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*, 371–406.

Hulleman, C., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, *2*(1), 88–110.

Ippolito, J., Steele, J. L., & Samson, J. F. (Eds.). (2008). Adolescent literacy [Special issue]. *Harvard Educational Review*, *78*(1).

Kimball, S. M., White, B., Milanowski, A. T., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education*, *79*(4), 54–78.

Kolligian, J., & Sternberg, R. J. (1987). Intelligence, information processing, and specific learning disabilities: A triarchic synthesis. *Journal of Learning Disabilities*, *20*(1), 8–17.

Knapp, M. S. (2003). PD as a policy pathway. *Review of Research in Education*, *27*, 109–157.

Kucan, L., & Beck, I. (1997). Thinking aloud and reading comprehension research: Inquiry, instruction, and social interaction. *Review of Educational Research*, *67*, 271–299. doi: 10.3102/00346543067003271

Lapp, D., Flood, J., & Ranck-Buhr, W. (1995). Using multiple text formats to explore scientific phenomena in middle school classrooms. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, *11*, 173–186.

Lawless, K. A., & Pellegrino, J. W. (2007). PD in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. *Review of Educational Research*, *77*, 575–611.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading tests* (4th ed.). Itasca, IL: Riverside.

Margolis, H., & Brannigan, G. G. (2009). *Reading disabilities: Beating the odds*. Voorhees, NJ: Reading2008 & Beyond.

Matsumura, L. C., Slater, S. C., Junker, B., Peterson, M., Boston, M., & Steele, M. (2006). *Measuring reading comprehension and mathematics instruction in urban middle schools: A pilot study of the Instructional Quality Assessment* (CSE Technical Report No. 681). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing. Retrieved from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/e0/f6.pdf

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement: Evidence from Cincinnati. *Peabody Journal of Education*, *79*(4), 33–53.

Muthén, L. K., & Muthén, B. O. (1998–2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.

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.

Penuel, W. R., Fishman, B. J., Yanaguchi, R., & Gallagher, L. P. (2007). What makes PD effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, *44*, 921–958.

Rowan, B., Correnti, F.-S., & Miller, R. J. (2002). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education*, *70*, 256–284.

Schacter, J., & Thum, Y. M. (2004). Paying for high- and low-quality teaching. *Economics of Education Review*, *23*, 411–430.

Smith, J. B., Lee, V. E., & Newmann, F. M. (2001). *Instruction and achievement in Chicago elementary schools*. Chicago, IL: Consortium on Chicago School Research. Retrieved from http://cesr.uchicago.edu/publication/p0f01.pdf

Spillane, J. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: The mediating role of teachers' zones of enactment. *Journal of Curriculum Studies*, *31*, 143–175.

Stein, N. L., & Trabasso, T. (1981). What's in a story: An approach to comprehension and instruction. In R. Glaser, (Ed.), *Advances in instructional psychology* (Vol. 2, pp. 213–267). Hillsdale, NJ: Erlbaum.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

Supovitz, J. A., & Turner, H. M. (2000). The effects of PD on science teaching practices and classroom culture. *Journal of Research in Science Teaching*, *37*, 963–980.

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 (2005–2006)*. Retrieved from http://www.tea.state.tx.us

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.

Weaver, C. A., III, & Kintsch, W. (1991). Expository text. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson, (Eds.), *Handbook of reading research* (Vol. 2, pp. 230–245). New York, NY: Longman.

Wenglinsky, H. (2000). *How teaching matters: Bringing the classroom back into discussion of teacher quality* (Policy Information Center Report). Princeton, NJ: Educational Testing Service.

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives*, *10*(12). Retrieved from http://epaa.asu.edu/epaa/v10n12/

Yell, M. L. (2006). *The law and special education* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

APPENDIX

TABLE A-1 Sample Items From the Perceived Teacher Knowledge Survey

Item	Not at all familiar	Somewhat familiar	Very familiar
Levels of questions	1	2	3
Main idea/gist statements	1	2	3
Vocabulary maps	1	2	3
Context clues	1	2	3

Note. Participants were instructed, "For each of the following vocabulary and comprehension strategies, rate how familiar you are with this strategy." Definitions of each strategy were provided to clarify the meaning.

TABLE A-2 Instructional Content of the Professional Development

Instructional content	Case 1	Case 2	Case 3
Comprehension			
Teach levels of questions	x	x	x
Teach the gist strategy	x	x	x
Teach gist for longer passages		x	x
Teach summarization			x
Vocabulary			
Identify and prioritize vocabulary	x	x	x
Teach words explicitly	x	x	x
Use graphic organizers (vocabulary maps, chapter and unit overviews)	x	x	x
Use practice activities	x	x	x
Build and activate background knowledge		x	x
Teach independent word learning strategies (context clues)			x