Journal of Educational Psychology
2002, Vol. 94, No. 4, 669–686

0022-0663/02/$5.00 DOI: 10.1037//0022-0663.94.4.669

Contribution of Spelling Instruction to the Spelling, Writing, and Reading of Poor Spellers

Steve Graham and Karen R. Harris
University of Maryland

Barbara Fink Chorzempa
State University of New York College at Geneseo

The contribution of supplemental spelling instruction to spelling, writing, and reading was examined. Second-grade children experiencing difficulties learning to spell participated in 48 20-min sessions designed to improve their spelling skills. In comparison with peers in a contact control condition receiving mathematics instruction, students in the spelling condition made greater improvements on norm-referenced spelling measures, a writing-fluency test, and a reading word-attack measure following instruction. Six months later, students in the spelling treatment maintained their advantage in spelling but not on the writing-fluency and reading word-attack measures. However, spelling instruction had a positive effect at maintenance on the reading word-recognition skills of children who scored lowest on this measure at pretest.

Spelling and Writing

Difficulties with spelling can potentially effect writing in several ways. One misspelled word can blur the message that an author is trying to convey. This point was illustrated in the following sentence, in which a child committed a spelling miscue that clouds his initial intention: "Louis XVI was gelatined during the French Revolution" (Abbington, 1932, p. 18). Such spelling miscues are not just limited to children, however, as this advertisement illustrates: "Local club needs experienced kook to run kitchen."

Poor spelling may also influence perceptions about a child's competence as a writer. When teachers or other adults are asked to evaluate two or more versions of a paper differing only in number of spelling miscues, papers with fewer spelling errors are assigned higher marks for writing quality than papers with more spelling errors (Marshall & Powers, 1969). Perhaps even more important, difficulties with spelling may interfere with the execution of composing processes during the act of writing (Graham, 1990; Scardamalia, Bereiter, & Goleman, 1982). Having to consciously attend to the process of spelling while composing may tax the writer's processing memory (see Berninger, 1999), interfering with other writing processes, such as generating content and planning. For instance, having to switch attention during composing to mechanical demands, such as having to think about how to spell a particular word, may lead the writer to forget plans or ideas already held in working memory, influencing writing output. Simultaneously allocating attention to spelling words while planning the next unit of text when writing may also affect the coherence and complexity of content integration, influencing the overall quality of writing. It is further possible that if attention is occupied with spelling concerns, there may be less opportunity for the writer to make his or her expressions more precisely fit intentions at the point of translation, affecting the process of translating ideas or words into sentences. Finally, spelling difficulties may constrain a child's development as a writer. Difficulties mastering spelling skills may lead young children to avoid writing and develop a mind set that they cannot write, resulting in arrested writing development (Graham, 1999).

One purpose of the present study was to examine the relationship between spelling and the process of writing. McCutchen (1995) and others (Berninger, 1999; Graham, 1999) have argued that the resource demands imposed by spelling, especially for beginning writers, can affect the use of other writing processes, such as planning, text generation, and sentence construction. Studies examining the correlation between spelling and writing performance of young children provide some support for this contention, as correlations typically range between .4 to .5 (Graham, 2000). For instance, Juel (1988) found that spelling performance accounted for 29% of the variance in the writing scores of first-grade children. Such correlational data must be interpreted cautiously, however, as spelling may also be considered an outcome of writing practice or instruction (Graham, 2000).

To more directly investigate the link between spelling and writing, Berninger et al. (1998) examined the effects of spelling instruction on children's writing performance. Second graders who were poor spellers were randomly assigned to seven spelling treatment groups and a contact control condition that received

Editor's Note. Carole Beale served as the action editor for this article.—MP

Steve Graham and Karen R. Harris, Department of Special Education, University of Maryland; Barbara Fink Chorzempa, Department of Curriculum and Instruction, State University of New York College at Geneseo.

Preparation of this article was supported by the Center to Accelerate Student Learning, funded by the U.S. Department of Education's Office of Special Education Programs Grant H324V980001. We wish to acknowledge the contributions of Rebecca Jones, Sonja McKenna, Helen Fives, Susan Moran, Bruce Saddler, and Linda Mason who either served as an instructor or assisted in scoring spelling, writing, or reading samples.

Correspondence concerning this article should be addressed to Steve Graham, Department of Special Education, University of Maryland, College Park, Maryland 20742. E-mail: sg23@umail.umd.edu

instruction in phonological and orthographic awareness skills. Specially trained tutors provided approximately 8 hr of instruction to pairs of students. This special instruction supplemented children's regular spelling program. All of the spelling treatments resulted in improved spelling performance, and one of the treatments had a positive impact on how much students wrote. Students in this group were taught common phoneme–spelling associations and practiced new spellings by pointing to each letter in a left-to-right order while simultaneously saying the sound in words they were studying.

Although the Berninger et al. (1998) study provides some preliminary support for the contention that there is a direct link between spelling and writing, at least in terms of content generation, additional research is needed. This includes replicating the effects for content generation or writing output and further examining the impact of spelling instruction on other aspects of composing, such as writing quality and sentence construction. As noted previously, each of these aspects of writing may be influenced by the resource demands of spelling. Additional replication is also needed because a text-production skill like spelling may not pose the same kind of resource demand problems as handwriting (see Berninger et al., 1997; Graham, Harris, & Fink, 2000; Jones & Christensen, 1999). Young writers cannot avoid the demands of handwriting when composing (unless they dictate or type their compositions), but they may routinely lessen spelling demands by not worrying about correct spelling.

Spelling and Reading

A second purpose of the present study was to examine the link between spelling and reading. As Treiman (1993) noted, some investigators believe that the relationship between the two skills are very close and that learning about spelling enhances reading proficiency. This viewpoint was expressed by Adams (1990) in her review of the early reading literature. She indicated that learning about spelling contributes to reading development, including children's ability to pronounce words correctly and decode unknown words.

Ehri (1987, 1989) proposed that spelling contributes to reading development by shaping children's knowledge of phonemic awareness, strengthening their grasp of the alphabetic principle, and making sight words easier to remember. When children encounter a new spelling, for example, their examination of how the letters and phonemes in the word line up provides additional cues about the phonemic structure of the word. As Ehri (1989) indicated, "orthography penetrates and commingles with readers' phonological knowledge to determine how phonemes in words are conceptualized" (p. 360). Spelling not only provides information about the sounds in words but may further enhance children's understanding of the concept that visible symbols stand for separate sounds in words. Knowledge of the alphabetic principle and decoding are presumably strengthened when children analyze the connection between the letters and sounds in words, attempt to spell unknown words phonetically, and are taught that specific letters stand for particular sounds. Finally, knowledge of the spelling system provides schemata that should help children make sense of the words they read, making them easier to remember. As children's knowledge of the spelling system grows, their sight word reading should becomes more accurate, as "they are able to make more complete connections between the spellings and pronunciations of words in memory" (Ehri, 1989, p. 359).

One source of evidence supporting the view that spelling and reading are closely linked is the finding that children who are good readers are usually good spellers. The correlation between spelling and reading range from .5 to .9 (Ehri, 1987; Horn, 1960). As Foorman and Francis (1994) indicated, knowledge of a word's spelling almost always aids the reading of that word. This correlational data are buttressed by experimental studies showing that spelling instruction improves reading performance (e.g., Berninger et al., 1998; Ehri & Wilce, 1987; Uhry & Shepherd, 1993). For instance, Berninger et al. (1998) found that spelling instruction improved the word-recognition performance of second graders who were poor spellers.

Although these data support the claim that spelling and reading are related, additional research is needed to verify this contention. As Treiman (1993) observed, the correlation between reading and spelling are not perfect, and some children who experience difficulty with spelling read quite well. Bryant and Bradley (1980) further indicated that spelling and reading may be particularly disconnected for young children, as they found that normal 6- and 7-year-olds could spell a word but not read it.

Prevention

In addition to examining the theoretical links between spelling and the skills of writing and reading, the present study addressed an important practical issue as well. How can we prevent writing difficulties? Although the work by Englert et al. (1995) demonstrated that a well-designed literacy program can have a positive impact on the writing performance of primary-grade children who experience learning difficulties, there is very little data on which skills or what aspects of instruction need to be emphasized to prevent writing difficulties (Graham & Harris, 2002). Several previous studies have shown that early, supplementary instruction in another text-transcription skill, handwriting, can facilitate the writing development of children with poor handwriting skills (Berninger et al., 1997; Graham et al., 2000; Jones & Christensen, 1999). The finding that there were transfer effects from spelling to writing in the Berninger et al. (1998) study suggests that supplementary spelling instruction during the primary grades may also be an important element in prevention efforts. Thus, a third purpose of the present study was to examine the impact of supplementary spelling instruction on the writing performance of young children who were poor spellers. We also examined the effects of such instruction on reading performance, as previous research (e.g., Adams, 1990; Ehri, 1989) suggested that spelling may be one of the keys to preventing reading disabilities.

Present Study

Second-grade students who were experiencing difficulty learning to spell were provided supplemental spelling instruction. This instruction reinforced and extended children's regular classroom spelling program by providing additional instruction and practice with spelling skills emphasized in both the participating schools' second-grade curriculum guides and the basal program used by the teachers. The supplementary program was developed so that it addressed two basic operations identified in several influential

models of spelling performance (Ehri, 1986; Jorm, 1983; Simon, 1976). According to these models, writers first access their lexical memory to see whether the spelling of the target word is stored there. If it is not, they generate a plausible spelling, using their current knowledge of the spelling system to make an informed guess. They might further verify the correctness of their spelling by decoding it or examining it to see whether it looks right.

Students' lexical knowledge was strengthened by teaching them how to spell correctly words that frequently occur in the writing of primary-grade children. Their knowledge of the spelling system was enhanced by teaching them common sound−letter combinations (e.g., consonants, blends, etc.), spelling patterns or rules involving long and short vowels (e.g., the doubling rule for adding a suffix), and frequently occurring phonograms or rimes (e.g., *ig*, *at*, etc.). With the word-sorting, word-building, and peer-practice activities used to teach these skills, both accuracy and fluency were stressed. Verifying the correctness of their spelling was emphasized as students learned how to spell frequently occurring words (i.e., misspellings made during study practice were corrected) and when building words using onsets and rimes (i.e., they examined each built word to determine whether it was a real or made-up word).

In contrast, students in the contact control condition were taught math skills. This treatment was chosen for the control condition because we did not expect that it would facilitate the process of learning to write or read. In addition, the math program that was modified for use in this study has been effective with a wide range of second-grade students (Fuchs, Fuchs, Hamlett, Phillips, & Bentz, l994; Fuchs, Fuchs, Phillips, Hamlett, & Karns, l995).

In terms of examining the link between spelling and writing, we extended previous research (i.e., the Berninger et al., 1998, study) in four important ways. One, we used multiple measures of writing performance to assess the impact of supplemental spelling instruction on learning to write. This included a writing-fluency measure from a norm-referenced test as well as a story-writing probe that yielded scores for writing output and compositional quality. Two, we examined whether the effects of supplemental spelling instruction were maintained over time. Six months following the completion of instruction, spelling, writing, and reading measures were readministered. Three, students were provided with 16 hr of supplemental spelling instruction. This doubled the amount of instructional time provided by Berninger et al. (1998). Four, the study builds on the previous Berninger et al. (1998) study by including students with a greater range of academic difficulties. Ten percent of the children in the Berninger et al. (1998) study received speech and language services, whereas 43% of the students in this study were identified as having a disability. This included children with speech and language difficulties, learning disabilities, attention-deficit/hyperactivity disorder (ADHD), developmental disabilities, and behavioral difficulties.

With regards to reading, the present study enhanced previous research in two important ways. One, it examined whether the effects of supplemental spelling instruction were maintained over time. There is a paucity of data examining this issue. Two, it examined the effects of spelling instruction on both the word-recognition and word-attack skills of the participating students. Although Berninger et al. (1998) reported that supplemental spelling instruction improved children's word-identification skills, they did not examine transfer to word-attack skills.

We anticipated that supplemental spelling instruction would have a positive impact on the participating children's spelling performance, leading to improvement on the specific skills taught as well as more generalized improvement to norm-referenced measures of spelling achievement. It was also expected that spelling instruction would lead to improvements in writing, as spelling would require fewer attentional resources as it improved, allowing children to devote more resources to other writing processes. Until transcription processes, such as spelling, become relatively automatic, they presumably impose a significant drain on available resources and limit other writing processes, such as content generation (McCutchen, 1995). Although we predicted that supplemental spelling instruction would boost children's writing output and fluency, no prediction was made concerning the quality of their writing. Graham, Berninger, Abbott, Abbott, and Whitaker (1997) found that the spelling performance of primary-grade children predicted how much they wrote under timed conditions but not the quality of their compositions. On the other hand, several studies reported a statistically significant relationship between spelling and writing quality (see Deno, Marston, & Mirkin, 1982; Hogan & Mishler, 1980). It was further expected that spelling instruction would have a salutary effect on children's word-identification and word-attack skills. Both theory and previous research indicate that spelling can enhance children's ability to identify and decode words correctly (e.g., Adams, 1990; Berninger et al., 1998; Ehri, 1989; Ehri & Wilce, 1987). Finally, we anticipated that the hypothesized effects of supplemental spelling instruction would be maintained over time. Students who spell better and are more knowledgeable about the relationship between sound and letters are more likely to write and read and will experience fewer negative experiences when learning either of these skills.

Method

Participants

Screening. At the beginning of October, 291 second-grade children in 12 classrooms attending four schools (in a single school district) located in urban middle-class neighborhoods outside of Washington, D.C., were administered the Spelling subtest from the Wechsler Individual Achievement Test (WIAT; Wechsler, 1992). This measure assesses children's ability to spell correctly in writing words dictated by an examiner. The test was used to identify children as at risk in spelling. To qualify as at risk, a child's standard score on the test had to fall two thirds of a standard deviation or more below the mean for the normative sample. All tests were scored by Steve Graham (after all identifying information had been removed by the teachers), and half of them were rescored by a graduate student unfamiliar with the design and purpose of the study. Interrater reliability was .99.

Sixty-seven children were identified as at risk for spelling using the criteria specified previously. Interviews with each child's teacher indicated that all but one of these children was experiencing difficulty with spelling. Of the 66 children identified as at risk by both the spelling test and the teacher, parents of 60 children granted informed consent for participation in the study.

Groups. The 60 second-grade children who granted informed consent were randomly assigned to two groups: spelling and mathematics instruction. In the 6-month period during which instruction took place (November through April), 5 of these students moved to another school district, and 1 child was removed from the study by his parents. The parents decided to remove the child from the study because he was experiencing difficulty in a number of subjects, and they wanted him to be in his regular classroom

full time. One of the children who moved was in the spelling group, whereas the other 5 children were in the mathematics group. We examined all available information for the 6 children to determine whether they differed in some significant way from the remaining 54 participants. Their personal characteristics (chronological age, gender, race, and socioeconomic status) as well as scores on spelling, writing, and reading pretest measures were similar to those of the group as a whole. When 1 of these 6 students was no longer available to participate in the study, he was replaced by another child who was identified by the teacher as experiencing spelling difficulties. Data on the replacement children's performance was not included as part of the study, as these children did not receive the full instructional program. It was necessary to replace the 6 missing children, as children were paired for instruction in both the spelling and math conditions, and most of the lessons in each condition included cooperative learning activities.

The mean age of the 54 participants was 7 years 3 months. English was the primary language for all children. Thirty-seven of the students were boys, whereas 17 were girls. Sixty-five percent of the children were Black, 26% were White, 6% were Hispanic, and 3% were of mixed race. The racial composition of the sample was consistent with that of the student body of the participating schools. Sixty-five percent of the participants received a free or reduced lunch. Again this distribution was equivalent with that of the students' schools, which were located in relatively poor urban settings. Mother's mean educational level was 13.1 years. All but 2 of the students' mothers had completed high school. Twenty-three students were identified as having a disability.

Of the 23 students with a disability, 12 experienced speech and language difficulties, 5 were classified as having a learning disability, 3 were identified by a clinical psychologist as having ADHD, 2 were labeled as emotionally disturbed, and 1 received services for developmental disabilities. Examination of students' files indicated that a variety of measures were used to assess and identify the children with special needs. Because the overlap in the measures administered from 1 child to the next was minimal, it was not possible to aggregate test scores for these students as a group or by disability category. However, all of the students with speech and language difficulties scored one standard deviation or more below the mean on one or more standardized tests of language or speech functioning (e.g., Battelle Developmental Inventory, Developmental Indicators for the Assessment of Learning—Revised; Newborg, Stock, & Wnek, 1984), Comprehensive Receptive and Expressive Vocabulary Test (Wallace & Hammill, 1994), Expressive One-Word Picture Vocabulary Test—Revised (Gardner, 1990), Peabody Picture Vocabulary Test—III (Dunn & Dunn, 1997), Comprehensive Receptive and Expressive Vocabulary Tests (Wallace & Hammill, 1994), or Test for Auditory Comprehension of Language (Carrow-Wolfolk, 1985). All 5 children with learning disabilities scored within the normal range on an individually administered intelligence test (e.g., Wechsler Preschool and Primary Scale of Intelligence—Revised [Wechsler, 1967] or the Wechsler Intelligence Scale for Children—III [Wechsler, 1991]) and one or more standard deviation below the mean in either reading or math on the Woodcock–Johnson Psycho-Educational Battery—Revised (WJ–R; Woodcock & Johnson, 1990). The 3 children with ADHD all took Ritalin and scored one or more standard deviations below the mean on the same academic measure administered to the children with learning disabilities. The 2 students classified as having an emotional disturbance primarily experienced difficulty with acting-out behaviors and aggression. The child with a developmental delay scored two standard deviations below the mean on both the Battelle Developmental Inventory (Newborg et al., 1984) and the Test for Auditory Comprehension of Language (Carrow-Wolfolk, 1985).

Additional evidence that the participating students were having difficulty with spelling was obtained using a second norm-referenced measure, the Test of Written Spelling—3 (TWS–3; Larsen & Hammill, 1994), also administered before the start of instruction (see the *Measures* section). Students' mean standard score on this test was 75.2 (range = 64 to 96), almost two standard deviations below the mean for the normative sample ($M = 100$; $SD = 15$). Two students (both in the math condition) had a standard score above 90 on the TWS–3. Their standard score on the WIAT spelling test was 89. Although there was a significant correlation between students' standard scores on these two tests ($r = .41$; $p < .01$), students' mean standard score on the TWS–3 was lower than their mean standard score on the WIAT Spelling subtest. The relatively modest correlation between scores for the two tests suggests that they do not measure the exact same content.

In addition to spelling problems, teachers further indicated that each child had difficulties with writing. The results from a standardized writing test, the Writing Fluency subtest of the WJ–R, administered before instruction began (see the *Measures* section), provided some confirmation for the teachers' observations. The mean standard score of the 54 students was 90.9 ($SD = 9.2$; range = 73 to 125). The mean for this test is 100 ($SD = 15$). Thus, as a group, the mean writing performance of the participating students was almost two thirds of a standard deviation below the normative sample. It should be noted that 60% of the students obtained a standard score of 90 or lower, but 10% had a standard score of 100 or better.

Many of the students also experienced difficulties with basic reading skills. On the Word Identification and Word Attack subtests of the Woodcock Reading Mastery Test—Revised (WRMT–R; Woodcock, 1987), administered before the start of instruction (see the *Measures* section), students' mean standard scores were 89.7 (range = 50 to 129) and 79.1 (range = 56 to 104), respectively. The mean for these tests is 100 ($SD = 15$). On the Word Recognition subtest, almost 70% of the students had a standard score below 90, whereas 13% scored above 100. On the word-attack measure, 87% of the students had standard scores below 90; one student had a standard score above 100.

Information on the characteristics of students by condition is presented in Table 1. Using *t* tests for independent means, we found that there were no statistically significant differences between students assigned to the two conditions at the start of the study in terms of chronological age or

Table 1
Student Characteristics by Instructional Condition at the Start of the Study

Variable	Condition	
	Spelling	Mathematics
Age (months)		
M	89.4	87.2
SD	5.5	5.0
Gender		
Female	8	9
Male	17	20
Race		
Black	15	20
White	8	6
Hispanic	2	1
Mixed	0	2
Free or reduced lunch		
Yes	15	20
No	10	9
Mother's educational level		
M	13.0	13.0
SD	1.8	1.6
Students with disabilities		
Yes	11	12
No	14	17

Note. $N = 25$ for the spelling condition and $N = 29$ for the math condition.

mother's educational level (both *p*s > .13). Chi-square analysis also showed that there was no statistically significant differences between conditions in terms of gender as well as the number of students receiving a free or reduced lunch or identified as having a disability (all *p*s > .55). Using *t* tests for independent means, we found that there were also no statistically significant differences between students in the spelling and math conditions on the standardized spelling, writing, and reading tests administered before instruction: WIAT Spelling ($p = .97$), TWS–3 Predictable Words ($p = .23$), TWS–3 Unpredictable Words ($p = .41$), WJ–R Writing Fluency ($p = .18$), WRMT–R Word Attack ($p = .81$), and WRMT–R Word Identification ($p = .90$).

Instructional environment. Although there was considerable variability in the amount of time students' teachers spent teaching spelling, ranging from 25 to 150 min a week ($M = 80.9$ min), a new list of spelling words was introduced each week. These words were taken from the spelling component of a basal reading program used by the participating school district. The list of weekly spelling words ranged from 11 to 16 words, depending on the teacher. Two of the 12 teachers indicated that weaker spellers worked on only 6 to 8 words a week.

At the start of each week, two thirds of the teachers administered a pretest on the words and asked students to correct misspelled words. All of the teachers gave a test at the end of the week to assess students' mastery of the words. During the week, 7 of the 12 teachers provided some class time for children to practice their spelling words. This was typically no more than 15 min a week. Most of the teachers indicated that students learned spelling skills, such as spelling patterns and phoneme−grapheme combinations, using the phonics activities included in the basal reading program.

Instructional adaptations for weaker spellers were relatively rare, with three teachers indicating that they made no modifications for weaker spellers. When teachers did make adjustments for weaker spellers, they typically involved extra review, small group instruction, or instruction on how to sound out and spell unknown words.

There was also considerable variability in the amount of time teachers devoted to writing instruction, ranging from 100 to 225 min a week ($M = 140$ min). One half of the teachers used a traditional process approach to writing instruction (see Calkins, 1985). Four teachers at one school used a writing curriculum developed by a local educator. This curriculum was based on the process approach to writing but included more sequenced and systematic instruction in basic writing skills than is typically offered in the process approach. The remaining two teachers used a modified version of the process approach, whereby students planned, drafted, and edited their papers, but instruction mostly centered around learning writing strategies developed for state-wide high-stakes tests administered during third grade.

General Instructional Procedures

Instruction was delivered to students in the spelling and math instructional conditions by six graduate students who were majoring in education. Before the start of the study, instructors were taught how to implement both instructional conditions. For each instructional condition, instructors were provided with a notebook that contained detailed directions for implementing each activity and lesson. The value of both of the instructional conditions was stressed, so that instructors would not be predisposed to one condition over the other. Over the course of a 2-week period, they practiced implementing both conditions until they were able to do each without error. To help ensure that both conditions were correctly implemented, instructors were provided with a checklist for each lesson that furnished step-by-step directions. As they taught a lesson, they were asked to check off each step as it was completed.

Instructors were assigned to one of the four participating schools. Three times a week, they worked with each of their assigned pairs of students for 20 min a session. Pairs of students were formed in the following way. First, we told the teachers in each school which students had been randomly assigned to each condition. Then, in conjunction with the teachers, students in the same condition were paired together on the basis of teachers' opinions of their compatibility and issues of scheduling.

A single lesson was conducted in each session, and all children participated in 48 instructional sessions. Each instructor worked with pairs of students from the two instructional conditions. For each instructional condition, there were 15 pairs of students, and students remained in the same instructional pair throughout the course of the experiment. In two of the schools, instruction was delivered in a quiet location in the classroom. In the other two schools, the instructors worked in a quiet location outside of the classroom (i.e., an empty room or at a table set up in the hallway). To ensure that none of the students missed any of their regular classroom spelling instruction, instructors were not scheduled to work with students when this instruction occurred. Furthermore, the children's teachers were periodically contacted to verify that the participating students were not missing classroom spelling instruction. We were unable to identify any instances in which children missed classroom spelling instruction as a result of participating in the study.

Spelling Instruction

Spelling instruction consisted of 48 lessons that were divided into 8 units (6 lessons per unit). Each unit focused on two or more spelling patterns involving long and/or short vowel sounds. The spelling patterns for each unit were as follows: Unit 1 (short vowel sound for /a/, /e/, and /i/ in CVC-, CCVC-, and CVCC-type words), Unit 2 (short vowel sound for /o/ and /u/ in CVC-, CCVC-, and CVCC-type words), Unit 3 (short vowel sound for /a/ in CVC-, CCVC-, and CVCC-type words and long vowel sound for /a/ in CVCe-type words), Unit 4 (short vowel sound for /o/ in CVC-, CCVC-, and CVCC-type words and long vowel sound for /o/ in CVCe- and CCVCe-type words), Unit 5 (short vowel sound for /i/ in CVC-, CCVC-, and CVCC-type words and long vowel sound for /i/ in CVCe- and CCVCe-type words), Unit 6 (the short vowel sound and /ck/ at the end of monosyllabic words and the long vowel sound and /ke/ at the end of monosyllabic words), Unit 7 (adding the suffix *ed* to monosyllabic words with a short vowel or a long vowel sound; i.e., the doubling rule), and Unit 8 (adding the suffix *ing* to monosyllabic words with a short vowel or a long vowel sound; i.e., the doubling rule).

The lessons for each unit followed a predictable pattern that included seven different types of activities (see Appendix). During the first lessons of each unit, each student pair participated in a word-sorting activity (Bear, Invernizzi, Templeton, & Johnston, 2000; Graham, Harris, & Loynachan, 1996; Zutell, 1993), in which words were sorted or categorized by the particular spelling patterns or features emphasized in that unit (Activity 1). For example, in Unit 1, words were sorted into three categories: words containing a short vowel sound for /a/, /e/, or /i/. First, the instructor directed the word-sorting activity, involving students in the thinking and decision-making process. The instructor placed two or more master word cards next to each other (i.e., *mad*, *sled*, and *pig* for Unit 1). The instructor pronounced each word and then said the word again emphasizing the target feature (e.g., short vowel sound for /a/ in *mad*). Students were asked how the words were similar and different, emphasizing both sounds and letters. Next, the instructor drew a word card from a pack of 11 to 12 word cards, pronounced the word, said it again (emphasizing the target feature), and placed it under the appropriate master word card (e.g., last was placed *mad*). Before placing the word under the appropriate category, the instructor modeled aloud the thinking processes that led to that decision. The instructor continued to do this until the students understood the process, and then the instructor encouraged them to categorize and place the remaining word cards while thinking aloud. If an error was made in placing a word, the instructor corrected and modeled aloud his or her thoughts on where the word should be placed. Once all of the words were placed, the instructor helped the students state a rule for the patterns emphasized in that word sort (e.g., in Unit 3, students generated the following rule for the

long /a/ vowel sound in CVCe words: when you hear a long /a/ in a small word, the *a* is often followed by a consonant and silent *e*.). Students then generated words of their own, writing them on blank word cards and placed them under the appropriate category. If an exception word was generated (e.g., *have* for the short vowel sound of /a/), a new category was started for exception words (a card with *a?* was used to designate this category). This procedure helped students learn that there are exceptions (or inconsistencies) to the rules or patterns they were discovering. For more than half of the units (i.e., 3, 4, 5, 7, and 8), exception words were included in the word pack, and an exception master card was placed beside the other master cards at the start of the sort.

Once the instructor-directed word sort was completed, the word cards were reshuffled, and the student pair completed the sort by themselves. The instructor timed how long it took the students to complete the word sort, and if there was enough time remaining, the students did the sort again. The goal at this point for the students was to do the sort correctly and beat their previous time. If students made an error during either of these two sorts, the teacher gave feedback and corrected as necessary.

At the end of Lesson 1, students were encouraged to "hunt" for words that fit the patterns that they were working on in that unit (Activity 2). During each of the subsequent lessons, students shared the words that they had found with the instructor. For each appropriate word, students received a sticker that they placed in a sticker book.

During Lesson 2, each student was given 8 spelling words to study (Activity 3). Over the course of the experiment, they studied 64 words. All of these words frequently occur in the writing of young children. They were selected from lists of the most frequently occurring words in young children's writing compiled in studies by Farr, Beverstock, and Robbins (1988); Graham, Harris, and Loynachan (1993); and Henderson (1978). The student had missed each of these words on a pretest given before the start of the unit. The pretest list contained 25 words that fit the spelling patterns emphasized in that unit. Words on the list were randomly assigned by pattern so that a word for one pattern was followed by a word for another pattern and so forth (e.g., *camp*, *best*, *fish*, *last*, *west*, *sick*). Once the student missed 8 words, the pretest was ended.

In Lessons 2−5, students used two basic study procedures to learn the 8 spelling words (13 min were allotted to studying words in each lesson). One procedure was called "graph busters" and involved the student recording on a graph the number of times she or he correctly studied the words using a traditional study strategy and then trying to beat this tally the next time the procedure was used. The strategy involved the following steps: (a) say the word and study the letters, (b) close your eyes and say the letters, (c) study the letters again, (d) write the word three times without looking at it, and (e) check the spellings and correct any misspellings. This strategy has been used to teach spelling words to poor spellers in several previous studies (Harris, 1986; Harris, Graham, Reid, McElroy, & Hamby, 1994). Students always used graph busters in Lessons 2 and 3, and it could be used in Lesson 5 at the discretion of either the students or the instructor. The goal in Lessons 3 and 5 was to beat the number of correct practices achieved in the previous lesson. During Unit 1, the instructor modeled how to use the strategy and provided needed assistance in its use, until students could do graph busters independently.

The second study procedure involved studying words while playing a game. Working in pairs, students played a game that required the child to produce the correct written spelling of a word to complete a move. If the word was spelled incorrectly, the misspelled word was corrected by the student's partner. All of the games were taken from the Spell-it Write spelling series (Harris, Graham, Zutell, & Gentry, 1998).[1] An example of a game was Spelling Road Race. This included a laminated board with a racing track marked off into 30 segments. When a child correctly spelled a word, he or she moved a space for each letter in the word. The spelling game was not introduced until Lesson 4, after students had gained some facility in spelling their words. The spelling game was typically used in Lesson 5 as well but could be replaced by graph busters at the discretion of either the students or the instructor.

At the start of Lessons 2−5 (and the end of Lesson 6 if time permitted), students participated in a short activity (2 min) entitled phonics warm-up (Activity 4). The purpose of this activity was to improve their skills in correctly identifying the letter(s) corresponding to sounds for short vowels, consonants, blends, and digraphs. Across the course of the instructional program, students worked on 46 different sound−letter combinations. Each sound−letter combination was represented on a card with a picture on one side (e.g., a picture of a dog) and a corresponding letter on the other side (e.g., *d*). Cards were divided into sets, ranging from 9 to 16 cards. For the first minute of phonics warm-up, 1 student would hold up a card and identify the word pictured on the card as well as the location and sound of the target letter (e.g., "Dog starts with the sound /d/"); the student was then asked to identify the letter that corresponded to that sound. If an incorrect letter was named, then the child with the card said the correct response. The children continued cycling through the cards in this manner until 1 min had elapsed, and then they changed roles. Once students had mastered one set of cards, they would start to work on the next set.

During Lessons 2–5, the student pair worked together (for 4 min) to build words that corresponded to the spelling patterns emphasized in that unit (Activity 5). With the exception of Lesson 5, a single rime was introduced in each lesson. The instructor began by placing a card containing the rime (e.g., *ig*) on the table and said the sound that the rime made. The instructor then modeled building a word by placing a card containing either a consonant or blend at the front of the rime. Next, the students worked together to develop as many real words as they could, using 18 letter cards containing either a consonant, blend, or digraph. For each rime, at least 10 of these letter cards made a real word. If the students made a nonsense word, the instructor told them that it was not a real word. The instructor timed how long it took the pair to complete the word-building activity. If there was enough time remaining, the students repeated the word-building activity, with the goal of beating their previous time. In Lesson 5, students practiced building words with the three rimes introduced in Lessons 2−4.

At the start of Lesson 6, students were given a unit test to determine whether they had mastered the 8 words they had studied during Lessons 2−5 (Activity 6). After the test was completed, they corrected any misspelled words and recorded their performance on a graph that included data points for all 8 units. They then set a goal for how many words they would spell correctly on the next unit test and recorded this on the graph. Whenever a student met his or her goal for a unit test, the instructor used a magic marker to place a star over that point on the graph. Also, during this lesson, students completed a test assessing their spelling of nine words that contained the rimes emphasized during word building (see word-building spelling test in the *Measures* section). This test was not corrected, and no feedback was given to students on their performance.

Starting with the second unit, a review of the patterns introduced in the prior unit were undertaken immediately following testing (Activity 7). This included building words with the rimes introduced in the previous unit as well as reviewing the generalizations that students generated during the previous unit.

Mathematics Instruction

Like spelling instruction, the math control condition included 48 instructional sessions that were divided into 8 units (6 lessons per unit). The units focused on the following skills: Unit 1 (addition facts), Unit 2 (addition without regrouping), Unit 3 (addition with regrouping), Unit 4 (subtraction

[1] We are indebted to Zaner-Bloser (Columbus, OH) for providing us with games and word-sorting cards. Copies of the instructional manuals are available from Steve Graham.

facts), Unit 5 (subtraction without borrowing), Unit 6 (subtraction with borrowing), Unit 7 (word problems), and Unit 8 (word problems). The math program was a modified version of the Peer-Assisted Learning math program developed by Fuchs and Fuchs (see Fuchs et al., l994, l995) for students in second grade.

With the exception of Lesson 6, all lessons contained four activities. The first activity in these lessons was math warm-up (4 min). The purpose of this activity was to improve students' mastery of basic addition and subtraction math facts. Across the course of the instructional program, students worked on 66 addition facts and 66 subtraction facts. This included all addition and subtraction facts involving the numbers 0 to 10. Each fact was printed on a card with the problem on one side (e.g., 4 + 5 =) and the answer on the other side (i.e., 9). Cards were divided into sets, ranging from 19 to 26 cards. For the first 2 min of math warm-up, one student would hold up a fact card and say, "What number does this make?" If an incorrect answer was given, then the child with the card named the correct number. The children continued cycling through the cards in this manner until 2 min had elapsed, and then they switched roles. Once students had mastered one set of cards, they would start to work on the next set. Each day the students recorded the number of math facts that each answered correctly. If a student beat the previous day's performance, then the instructor used a magic marker to place a star over that point on the graph.

The second activity in Lessons 1−5 was peer coaching (10 min). With this activity, the instructor modeled aloud how to do two sample problems (e.g., 43 + 6 =) for the skill emphasized in that unit (i.e., adding without regrouping). To guide the process, the instructor used a "coaching card," which included verbal directions for completing the problem. Using chips to illustrate each number in the problem (a red chip = 1; a green chip = 10), the instructor showed the students how to do the problem concretely while following the steps on the coaching card.

Next, students took turns coaching each other as they worked on problems similar to the ones modeled by the instructor. As 1 student (the player) worked on a problem, saying each step on the coaching sheet aloud, the other student (the coach) monitored the process, telling the first student when something was done incorrectly. To facilitate this process, the coach had a booklet with the correct answers. If an answer was incorrect, the player was asked to fix it and could obtain help from either the coach or the instructor. If one or both of the students did not appear to understand how to do the problems, then the instructor demonstrated how to do the problems using manipulatives, such as base-10 blocks, or connected the problem to something in real life (e.g., "If you had eight donuts and your brother ate three, how many would be left?"). After two problems, the coach and the player switched roles. This continued until it was time to move on to the next activity.

The third activity, independent practice, provided students an opportunity to practice independently the problems emphasized in that unit (5 min). In Lessons 1−3, each student was directed to answer correctly as many problems as he or she could in a period of 3 min. The instructor then scored their responses, clearing up any misunderstandings that occurred. The students graphed the number of problems answered correctly. The goal was to beat the number of correct answers obtained in the previous lesson. In Lessons 4 and 5, students played a game with each other as they completed their independent practice problems. Using the same games that were applied with the spelling program, students took turns working on separate problems. When a child correctly answered a problem, he or she moved a space for each digit in the answer. If a problem was not answered correctly, the student made no move and received help from either the other student or the instructor.

The fourth activity, sticker reward, occurred at the end of each lesson (1 min). Students were rewarded with one or more stickers if they stayed on task and worked hard during the 20-min period. Students placed their stickers in a sticker book.

Before each unit, children in the math condition took the same spelling pretests that students in the spelling condition were administered. The first 8 words spelled incorrectly by each student in the control group on the pretest were tested at the start of Lesson 6 (as was done with children in the spelling condition). Students in the math condition, however, did not correct this test, graph the number of words spelled correctly, or set a goal for number of words to be spelled correctly on the next unit test. Like students in the spelling condition, students in the control condition completed the word-building spelling tests as well (see *Measures* section).

During Lesson 6, students in the math condition also received stickers for on-task behavior and hard work and did the math warm-up activity if time permitted. Starting with Unit 2, students reviewed the skills taught in the previous lesson. This involved the instructor modeling how to do a sample problem, with students then demonstrating how to do several problems while saying each step aloud. The instructor provided assistance as needed.

Fidelity of Treatment Implementation

To ensure that instructional procedures were delivered as planned, the following safeguards were implemented. First, instructors received intensive practice in applying the instructional procedures for each condition. Second, instructors met with Steve Graham weekly to discuss any glitches that occurred in implementing procedures. Reported glitches or deviations from instructional plans occurred rarely and usually involved an inadvertent mistake by an instructor. Third, instructors were provided with a checklist that contained step-by-step directions for each lesson. As they completed a step, instructors were asked to check it off. Examination of these checklists once instruction ended showed that instructors completed 99.1% of the steps during spelling instruction across lesson plans and 98.3% during math instruction. Fourth, one fourth of all lessons were tape-recorded and checked to determine whether each step of a lesson was executed as intended by the instructor. The percentage of steps completed correctly across lesson plans by the six instructors was 98.0% for spelling instruction and 97.6% for math instruction.

Measures

Each student was individually assessed before the start of instruction at the beginning of October (pretest), after instruction ended at the beginning of May (posttest), and again 6 months later at the beginning of November (maintenance). The order of tests was counterbalanced and administered across a series of six to eight sessions to minimize fatigue. The only exceptions involved the spelling unit, maintenance, and word-building spelling tests, which were administered during instruction; two math tests (math facts and computation tests) and the sound−letter test, which were administered at pretest and posttest only; and another math test (word problems), which was administered just at posttest. The math tests were included to determine whether the modified Peer-Assisted Learning program in math was effective. Because the effects of spelling instruction, not math, were the focus of this study, the math tests were not administered at maintenance. The word-problem test was only administered at posttest, because students' teachers expressed reservations about administering this test at the start of the study, indicating that these students would not be able to do these types of problems at this point.

Spelling. During Lesson 6 of each unit, students were administered two spelling tests: the spelling unit test and the word-building spelling test. For each student, the words on the spelling-unit test were the first 8 words missed on a pretest administered before the start of each unit. This pretest contained 25 words that fit the spelling patterns emphasized in that unit. Students in the spelling condition studied these words during the course of the unit; thus, this test measured their mastery of the studied words. Because the words on each spelling unit test were not the same for all children, we examined whether the words that children in the spelling and

control groups were asked to spell on the unit tests differed in terms of their difficulty or frequency. Using data from the New Iowa Spelling Scales (Greene, 1954), each word tested was assigned a difficulty score. For each word tested, this scale provided a percentage of students at second grade in the normative sample who were able to spell a target word correctly. Furthermore, we used data compiled in the *American Heritage Word Frequency Book* (Carroll, Davies, & Richman, 1971) to obtain a measure of how frequently each word occurred in the reading material at the elementary-school level. This measure provided an index of frequency of occurrence per one million words. Across the eight unit spelling tests, there was no statistically significant difference between the spelling and control condition in terms of the spelling difficulty of the words tested ($p = .48$) or the frequency with which they occurred in children's reading material ($p = .87$).

For the even-numbered units, half of the words from the unit spelling test were assessed one session later providing a short-term maintenance measure (i.e., immediate-maintenance spelling test). For the odd-numbered units, half of the words from the unit spelling test were tested 2 weeks later, providing a maintenance test of longer duration (delayed-maintenance spelling probe). The words selected for testing on both of these maintenance probes were randomly selected from each student's list of 8 words. For each of these measures (i.e., the unit spelling test and the two maintenance tests), a student's score was the number of words spelled correctly.

The second measure administered during Lesson 6 was the word-building spelling test. For each unit, this test included 9 words, and each word contained one of the 3 rimes taught during that unit (none of these words were practiced or constructed during instruction, however). For each rime, an equal number of words was included on the test. This test directly assessed skills learned during the word-building activities and the phonics warm-up activities. These tests differed from the unit spelling tests in that all students were tested on the same words.

For the word-building spelling tests and the unit spelling tests (including maintenance), the student was asked to write the correct spelling of a word, after the examiner said the word, used it in a sentence, and said the word again. The score was the number of words spelled correctly. All tests were scored by the examiner as well as a graduate student unfamiliar with the design and purpose of the study. Interrater reliability ranged from .99 to 1.00.

To assess students' knowledge of sound−letter combinations, students were asked to name the letter or letters made by 30 different sounds during a 1-min period. The 30 sounds included consonants, short vowels, consonant blends, and consonant digraphs. These were the same types of sound−letter combinations targeted for instruction during the phonics warm-up activity in the spelling condition. After the examiner said a sound, the student was directed to say the corresponding letter(s). The student's score was the number of letters produced correctly in 1 min. This task was scored by the examiner. One half of the recording sheets were rescored by a graduate student who was unfamiliar with the design and purpose of the study. Interrater reliability was 1.00.

Two norm-referenced measures of spelling were also administered to each student: the Spelling subtest from the WIAT (Wechsler, 1992) and the TWS–3 (Larsen & Hammill, 1994). With both of these measures, the students were asked to write the correct spelling of a word, after the examiner said the word, used it in a sentence, and said the word again. The Spelling subtest for the WIAT contains 45 items, and internal consistency and test−retest reliability for this subtest, as reported in the WIAT manual, are excellent for the age-group tested (.89 or above). The TWS–3 contains two spelling scales with 50 words each: one that assesses students' spelling of words that are consistently governed by the rules of standard American English (Predictable Words subtest) and another that assesses words that do not follow these rules (Unpredictable Words subtest). Internal consistency reliability for both subtests, as reported in the TWS–3 manual, is also very strong for the age group tested (greater than .90 in most studies). All of the norm-referenced spelling tests were scored by the examiner as well as a graduate student unfamiliar with the design and purpose of the study. Interrater reliability ranged from .98 to .99. The score for these tests was a standard score, computed following the procedures in the test manual.

Writing. To assess effects of spelling instruction on writing, two tasks were administered to each student. With one of the tasks, the Writing Fluency subtest from the WJ−R (Woodcock & Johnson, 1990), the child composed a sentence from three written words that accompanied a picture. The student's response to each item was scored as correct if all three words were used without modification in a grammatically complete sentence. As specified in the test manual, testing was stopped if three correct responses were not produced in 2 min, whereas an additional 5 min of testing were provided if this criteria was met or exceeded. The score for this test was a standard score, computed following the procedures in the test manual. Although this test involves little idea generation or planning as words are provided, it does require text generation to link the provided words into a sentence (translation) under timed conditions (Berninger et al., 1997). The test−retest reliability for this subtest, as reported in the WJ−R manual, is .87 for the age group tested, and interrater reliability between two scorers (Steve Graham and a teacher unfamiliar with the design and purpose of the study) was .99.

For the second writing task, story writing, students were asked to write a story to go with a black-and-white picture. Pictures were selected so that they were interesting and open ended enough that students would be able to devise their own story line when responding to them. Two elementary school teachers and a second-grade child selected 4 pictures that met these criteria from a set of 75 pictures. At pretest and posttest, students were given 2 pictures and asked to select one to write a story about. This further increased the possibility that students would be interested in the picture. The assignment of the two sets of pictures was counterbalanced so that (a) each child responded to a different set of pictures at pretest, posttest, and maintenance, and (b) at each testing probe, an equivalent number of students in each condition responded to each set of pictures. This ensured that students always wrote a story in response to a new picture, while controlling for possible variations in students' writing performance because of differences in the content or the subject of the pictures. Before writing, students were told that the instructor could not provide any help with writing the story (this included help with spelling words).

At maintenance, story-writing performance was assessed using the Story Construction subtest of the Test of Written Language—3 (Hammill & Larsen, 1996). With this test, students are allotted 15 min to write a story in response to a black-and-white picture. As with the story-writing probes administered at pretest and posttest, students were provided no help or feedback when writing their story.

Two scores were computed for all stories: total number of words written and compositional quality. Total number of words included all written words, regardless of spelling, that represented a spoken word. Compositional quality was assessed using a traditional holistic rating scale. Examiners were asked to read each story attentively to obtain a general impression of overall writing quality. Compositions were then scored on an 8-point scale, with 1 representing the lowest quality of writing and 8 representing the highest quality. In determining the score for each story, examiners were told that imagination, organization, grammar, sentence structure, and aptness of word choice should all be taken into account in forming a judgment about overall quality and that no one factor should receive undue weight. To guide the examiners in the scoring process, we provided them with a representative story for a low, middle, and high score. These compositions were collected in November and obtained from students in two second-grade classes in one of the participating schools. These classrooms did not participate in the study. All of the students in these two classes were asked to write a story in response to a picture. Two former elementary school teachers then selected the best, average, and poorest stories on the basis of the scoring criteria described previously.

All stories were typed before scoring. Identifying information was removed, and the spelling, punctuation, and capitalization in each story was

corrected. This was done to minimize any bias that might occur when examiners scored stories. Previous research has shown that the appearance of text or surface-level features, such as handwriting legibility, can influence judgments about writing quality (Graham, 1999). All stories were scored for total number of words by Steve Graham, and one half of these stories were rescored by a graduate student who was unfamiliar with the design and purpose of the study. Interrater reliability was 1.00. For compositional quality, all compositions were independently scored by two graduate students who were unfamiliar with the design and the purpose of the study. Interrater reliability was .89. The quality score for a student's paper was the average score for the two raters.

Reading. To assess effects of spelling instruction on reading, two tasks were administered to each student: the Word Identification and Word Attack subtests from the WRMT–R (Woodcock, 1987). With the Word Identification subtest, students read aloud single printed words. In contrast, the Word Attack subtest assesses students skills in using phonic and structural analysis to read aloud nonsense words. The internal consistency reliabilities for these two subtests, as reported in the WRMT–R manual, are excellent for the age group tested (.90 or greater). Both subtests were scored by the examiner as well as a graduate student unfamiliar with the design and purpose of the study. Interrater reliability was .99 for both subtests. The score for these tests was a standard score, computed following the procedures in the test manual.

Math. Three tests were administered to assess the skills taught to students in the mathematics control condition. With the math facts test, students were allotted 2 min to complete 25 addition and subtraction problems involving basic facts (e.g., 7 − 3 =). These problems included addition and subtraction facts taught in Units 1 and 4. With the computation test, students were allotted 10 min to complete 16 problems involving addition with and without regrouping (e.g., 46 + 9 =) and subtraction with and without borrowing (e.g., 36 − 19 =). These problems included the skills taught in Units 2, 3, 5, and 6. With the word-problem test, students were allotted 10 min to complete word problems (e.g., "There are 32 students in Joe's class. If there are 15 boys in the class, how many girls are there?"), similar to the ones taught in Units 7 and 8. These tests were based on the measures used by Fuchs and Fuchs (Fuchs et al., l994, 1995) to examine the effectiveness of Peer-Assisted Learning in mathematics with children at the second-grade level. The score for each of these tests was the number of problems answered correctly during the specified time frame. All tests were scored by the instructors and a graduate student unfamiliar with the design and purpose of the study. Interrater reliability ranged from .99 to 1.00.

Results

Analysis

We first examined the effects of supplemental spelling instruction on the measures that were administered during treatment (i.e., spelling-unit, maintenance, and word-building spelling tests). Effects on the measures administered immediately following treatment were then examined, followed by an examination of the effects of treatment on the maintenance measures administered 6 months later. Because students worked together in small groups of 2, the unit of analysis was the mean performance of the pair. This arrangement did not allow us to examine separately the effects of treatment on the performance of students with special needs, as students with and without special needs were not assigned to separate pairs. We did conduct preliminary tests, however, using students' individual scores as the unit of analysis, to determine whether the outcome scores of students with and without disabilities differed significantly or whether there was an interaction between student type and treatment condition. A significant main effect for student type was only obtained for two outcome variables (i.e., the quality of students' stories at posttest and the number of computation problems completed correctly at posttest). For both of these measures, students without a disability outperformed students with a disability. Most important, a significant interaction between student type and treatment condition was not obtained for any of the outcome variables administered in this study. Thus, it appears that the treatment conditions had a similar effect on students with and without disabilities.

With the exception of the assessments administered during instruction and the word-problem measure administered just at posttest, all analyses involved the use of a one-way analysis of covariance (ANCOVA). The independent variable was treatment condition and the covariate was performance on the pretest. Entry performance was taken into account when analyzing the effects of treatment, because pretreatment differences have been found to influence instructional outcomes (Pressley, Goodchild, Fleet, Zajchowski, & Evans, 1989). Prior to conducting each ANCOVA, we used hierarchical regression analyses to test the homogeneity of regression assumption. If this assumption was met, then ANCOVA was used. If the assumption was not met, then the Johnson-Neyman procedure was used. Failure to meet the assumption of homogeneity of regression indicated the treatments had differential effects on students scoring low and high on the pretest. The Johnson-Neyman procedure was used to establish the regions of significance for the interaction between pretest and treatment conditions, thereby specifying the upper and lower values in the range of pretest scores associated with superior outcomes for one of the treatment conditions (Pedhazur, 1982).

To estimate the practical significance of effects, we computed effect sizes for the assessments administered during treatment as well as the word-problem measure by subtracting the difference between means and dividing by the standard deviation for the control condition (i.e., the math condition). For effects involving ANCOVA (see Glass, McGaw, & Smith, 1981), the difference between the two adjusted means were subtracted and divided by the square root of the term: [(*MS* error) × (*df* for error – 1) / (1 – the correlation between pretest and outcome scores squared) × (*df* for error – 2)].

It should be noted that some of the outcome scores were not normally distributed (either the measure for kurtosis or skewness exceeded 1.0). These included number of words written on the posttest story, WJ–R Writing Fluency at posttest, the math facts test at posttest, and TWS–3 Unpredictable Words at maintenance. Consequently, we also analyzed these measures using the Mann-Whitney Test, a nonparametric procedure (Siegel, 1956). The results of the nonparametric and parametric tests were identical.

Progress During Instruction

During instruction, students were periodically administered spelling-unit, maintenance, and word-building spelling tests. These measures were directly tied to three components of spelling instruction. The spelling-unit and maintenance tests directly assessed the word-study component, whereas the word-building spelling test was directly tied to the word-building activity and phonics warm-up. A *t* test for independent means was used to analyze each of the measures administered during instruction. The independent variable for each analysis was treatment (spelling vs.

math). Each dependent measure (e.g., spelling-unit tests) was the average score for each student pair on all tests administered during instruction. For example, for the spelling-unit tests, the dependent measure was the average score on the eight spelling-unit tests. For all four analysis, the homogeneity of variance assumption was met. Means and standard deviations for the unit, maintenance (immediate and delayed), and word-building spelling tests are presented in Table 2.

Students in the spelling condition learned most of the spelling words they studied. On the spelling-unit tests administered at the end of each unit, they spelled 89% of the words correctly; students in the math condition spelled only 20% correct. As expected, there was a statistically significant difference between the mean performance of students in the two conditions (see Table 2), favoring the spelling treatment, $t(28) = 25.7$, $p = .00$ (effect size = 11.81). In addition, students in the spelling condition retained the correct spelling for most of the words they studied. On the maintenance tests administered one session after the spelling-unit tests, they spelled 92% of the words correctly; students in the math control condition spelled only 24% correctly. There was a statistically significant difference between the mean performance of students in the two conditions (see Table 2), once again favoring the spelling treatment, $t(28) = 15.07$, $p = .00$ (effect size = 3.72). On the 2-week delayed-maintenance tests, students in the spelling condition spelled 76% of the words correctly, whereas children in the math condition spelled only 21% correct (see Table 2). The difference between the mean performance of the two groups was also statistically significant (see Table 2), favoring the spelling treatment, $t(28) = 14.16$, $p = .00$ (effect size = 2.86). On the word-building spelling tests, students who received the spelling treatment spelled twice as many words correct as their peers in the math control condition (see Table 2). The difference between these two conditions was statistically significant, $t(28) = 6.03$, $p = .00$ (effect size = 2.86).

Posttest

Means and standard deviations for each measure administered at pretest and posttest are presented in Table 3. Also included in Table 3 are posttest scores that have been adjusted for initial pretest differences.

Spelling. The sound–letter test directly assessed the skills taught during phonics warm-up. After adjusting for initial pretest differences, there was a statistically significant difference between conditions, favoring the spelling treatment, $F(1, 27) = 5.88$, $MSE = 12.96$, $p = .02$ (effect size = .70). Thus, as expected, students in the spelling condition were more fluent in correctly identifying sound–letter combinations following instruction than their counterparts in the math condition.

The norm-referenced measures of spelling performance (i.e., WIAT Spelling, TWS–3 Predictable Words, TWS–3 Unpredictable Words) that were administered to students assessed the impact of the spelling treatment beyond the specific words and skills that were taught. After adjusting for initial pretest differences, there was a statistically significant difference between conditions, favoring the spelling treatment on all three of these measures: WIAT, $F(1, 27) = 5.35$, $MSE = 41.06$, $p = .03$ (effect size = .77); TWS–3 Predictable Words, $F(1, 27) = 16.25$, $MSE = 20.91$, $p = .00$ (effect size = 1.05); and TWS–3 Unpredictable Words, $F(1, 27) = 4.26$, $MSE = 35.24$, $p = .05$ (effect size = 0.64). It should be noted that by the end of instruction, students in the spelling condition improved their normative standing on each of these measures. In comparison with the test's normative sample, scores for students in the spelling condition rose by two fifths, one half, and one third of a standard deviation on the WIAT, TWS–3 Predictable Words, and TWS–3 Unpredictable Words, respectively (see Table 3). In contrast, the normative standing of children in the math control condition on each of these measures evidenced little change from pretest to posttest.

To determine how many students in the spelling condition were responsive to treatment, we examined how many students' standard scores improved from pretest to posttest on the WIAT and TWS–3 Predictable Words. All but 1 student in the spelling condition evidenced improved performance from pretest and posttest on at least one of these two tests. Eighty-eight percent of students' scores improved by 5 or more points (one third of a standard deviation or more for the tests' normative sample) on one of these two tests. It should be noted that the 1 student who evidenced no growth on either of the two tests was similar to his peers in the spelling condition in terms of his performance during instruction. He averaged 7.4 words correct on the spelling-unit tests, 3.3 words correct on the immediate-maintenance tests, 2.8 words correct on the delayed-maintenance tests, and 7.0 words correct on the word-building spelling tests.

Writing. Consistent with our predictions, the writing-fluency skills of students in the spelling condition improved more than those of the children in the math control condition. After adjusting for initial pretest differences, there was a statistically significant difference between conditions on the Writing Fluency subtest of the WJ–R, favoring the spelling treatment, $F(1, 27) = 5.31$, $MSE = 68.66$, $p = .03$ (effect size = 0.78). In comparison with the test's normative sample, scores for children in the spelling condition rose by almost two fifths of a standard deviation, whereas scores for students in the math condition dropped slightly (see Table 3).

Contrary to expectations, however, supplemental spelling instruction did not result in a statistically significant effect for either story length or story quality. Although students in both conditions

Table 2
Means and Standard Deviations for Spelling Measures Administered During Instruction

	Condition			
	Spelling		Mathematics	
Measure: spelling tests	*M*	*SD*	*M*	*SD*
Unit	7.15	0.69	1.60	0.47
Immediate maintenance	3.62	0.38	0.97	0.72
Delayed maintenance	3.03	0.57	0.95	0.73
Word building	5.51	1.41	2.76	1.07

Note. Mean performance on each measure averaged the tests administered (eight tests were administered for the unit and word-building spelling measures, whereas four tests were administered for each of the maintenance spelling measures); the spelling-unit test contained eight items; the word-building test contained nine items, and both maintenance tests contained four items each. $N = 25$ for the spelling condition and $N = 29$ for the math condition.

Table 3
Means and Standard Deviations for Pretest, Posttest, and Adjusted Posttest Scores by Condition

	Spelling instruction						Math instruction					
	Pretest		Posttest		Adjusted		Pretest		Posttest		Adjusted	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Sound–letter	12.4	3.0	18.6	4.8	18.9		13.0	3.2	16.0	4.8	15.7	
Spelling: WIAT	87.4	4.7	93.5	9.2	93.6		87.5	2.8	88.2	3.0	88.1	
Predictable Words: TWS–3	78.1	6.3	86.1	8.8	87.3		80.4	5.8	81.6	6.4	80.4	
Unpredictable Words: TWS–3	81.5	5.1	86.5	8.8	87.1		82.9	4.0	83.3	5.1	82.6	
Writing Fluency: WJ–R	88.8	6.3	94.5	10.2	94.7		89.4	4.1	87.9	7.3	87.7	
Story length	25.6	16.6	35.3	19.7	33.7		22.4	11.3	45.4	39.0	46.9	
Compositional quality	1.8	0.8	3.0	1.6	2.8		1.5	0.5	2.9	1.8	3.0	
Word Attack: WRMT–R	79.4	8.4	84.5	11.1	84.2		78.7	10.6	74.7	12.2	75.0	
Word ID: WRMT–R	90.2	11.4	89.8	9.0	89.5		89.4	7.6	87.1	6.8	87.3	
Basic facts: Math	5.7	3.3	9.3	4.8	9.6		6.2	2.4	14.9	3.6	14.7	
Computation: Math	2.2	1.5	5.4	2.5	5.4		2.3	1.5	8.2	2.4	8.1	
Word problem: Math			1.4	1.4					2.4	1.3		

Note. Adjusted score is the posttest score adjusted for initial pretest differences. Standard scores were reported for the WIAT, TWS–3, WJ–R, and WRMT–R ($M = 100$; $SD = 15$). $N = 25$ for the spelling condition and $N = 29$ for the math condition. WIAT = Wechsler Individual Achievement Test; TWS–3 = Test of Written Spelling—3; WJ–R = Woodcock–Johnson Psycho-Educational Battery—Revised; WRMT–R = Woodcock Reading Mastery Test—Revised; Word ID = Word Identification.

increased the length and quality of their stories by posttest (see Table 3), differences between the two treatment conditions were not greater than chance for either of these variables. It is interesting to note that there was considerable variability at posttest in the length of students' stories, especially for children in the math control condition. Variability in young children's writing is not unusual (see, e.g., Graham et al., 1997), and the especially high variability in the writing scores of children in the math control condition appeared to be primarily due to three outliers. The posttest stories of these 3 children were 246-, 160-, and 134-words long, respectively. When the scores for these 3 students are dropped from the analysis, the standard deviations of the math control group drops from 39.0 to 24.9 words. This more closely resembles the standard deviation of 20.7 words for students in the spelling treatment condition.

Reading. As expected, the word-attack skills of students in the spelling condition improved more than those of the children in the math control condition. After adjusting for initial pretest differences, there was a statistically significant difference between conditions on the Word Attack subtest of the WJ–R, favoring the spelling treatment, $F(1, 27) = 9.99$, $MSE = 63.96$, $p < .01$ (effect size = .82). In comparison with the test's normative sample, scores for students in the spelling condition rose by almost one third of a standard deviation, whereas scores for students in the math control condition dropped slightly (see Table 3).

Contrary to expectations, however, supplemental spelling instruction did not result in a statistically significant effect for word recognition. Performance on the WRMT–R Word Identification subtest declined slightly for students in both conditions over the course of the treatment (see Table 3).

Math. The modified Peer-Assisted Learning math program used in this study was effective in teaching most of the math skills targeted for instruction (see Table 3). After adjusting for initial pretest differences, there was a statistically significant difference between conditions, favoring the math treatment on the basic-facts test, $F(1, 27) = 14.61$, $MSE = 13.20$, $p < .01$ (effect size = 1.19), and the computation test, $F(1, 27) = 13.08$, $MSE = 4.19$, $p < .01$ (effect size = 1.11).

Students' performance on the word-problem test administered at posttest only was analyzed with a *t* test for independent means. The mean difference between the two conditions was not statistically significant. Students in the math and spelling condition only averaged 2.4 and 1.4 problems correct at posttest, respectively (see Table 3).

Maintenance

Means and standard deviations for each measure administered at pretest and maintenance are presented in Table 4. Also included in Table 4 are maintenance scores that have been adjusted for initial pretest differences. Maintenance measures were administered 6 months after instruction ended. During this time period (May to November), 15 students moved out of the school district and were unavailable for testing. This was 27.8% of the 54 students who received instruction (the mobility rate in the four participating schools averaged 40.0%). Eight of these 15 students (53.0%) were in the spelling condition. One additional student moved during maintenance testing, providing an incomplete data set for this child. In the math condition, at least 1 student in all 15 of the original instructional pairs was tested at maintenance. In the spelling condition, this was the case for 12 of the 15 pairs.

It is important to note that the remaining students appeared to be representative of the sample as a whole, as there was no statistically significant difference between them and the students who moved in terms of their performance on the spelling measures administered during instruction (all ps $> .37$); on the spelling (all ps $> .23$), writing (all ps $> .39$), reading (all ps $> .65$), and math (all ps $> .57$) measures administered at pretest; or on the spelling (all ps $> .70$), writing (all ps $> .34$), reading (all ps $> .70$), and math (all ps $> .22$) measures administered at posttest. The two

Table 4
Means and Standard Deviations for Pretest, Maintenance (Mainten.), and Adjusted Maintenance (Adjusted) Scores by Condition

	Spelling instruction						Math instruction					
	Pretest		Mainten.		Adjusted		Pretest		Mainten.		Adjusted	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Spelling: WIAT	87.4	4.7	95.9	8.0	96.7		87.5	2.8	89.3	9.4	88.8	
Predictable Words: TWS–3	78.1	6.3	92.5	11.0	93.8		80.4	5.8	85.8	7.4	84.8	
Unpredictable Words: TWS–3	81.5	5.1	94.3	10.7	95.4		82.9	4.0	90.1	7.2	89.2	
Writing Fluency: WJ–R	88.8	6.3	96.3	8.2	96.4		89.4	4.1	90.9	9.8	90.9	
Story length	25.6	16.6	69.7	30.2	69.5		22.4	11.3	67.6	32.8	67.7	
Compositional quality	1.8	0.8	4.0	1.4	3.9		1.5	0.5	3.6	1.8	3.7	
Word Attack: WRMT–R	79.4	8.4	83.7	11.7	83.6		78.7	10.6	77.1	14.5	77.2	
Word ID: WRMT–R	90.2	11.4	93.0	7.2			89.4	7.6	87.4	11.8		

Note. Adjusted score is the maintenance score adjusted for initial pretest differences. Standard scores were reported for the WIAT, TWS–3, WJ–R, and WRMT–R ($M = 100$; $SD = 15$). $N = 17$ for the spelling condition and $N = 22$ for the math condition, except for story length and compositional quality, where there were 23 subjects in the math condition. WIAT = Wechsler Individual Achievement Test; TWS–3 = Test of Written Spelling—3; WJ–R = Woodcock–Johnson Psycho-Educational Battery—Revised; WRMT–R = Woodcock Reading Mastery Test—Revised; Word ID = Word Identification.

groups were also similar in chronological age ($p = .99$) as well as mother's educational level ($p = .30$). Furthermore, chi-square analyses indicated that the gender, race, or disability status of the remaining students did not differ from the departing ones (all ps $> .22$).

Spelling. As expected, the gains in spelling performance obtained at posttest were maintained 6 months later. After adjusting for initial pretest differences, there was a statistically significant difference between conditions, favoring the spelling treatment, on the three norm-referenced measures of spelling performance: WIAT, $F(1, 24) = 6.41$, $MSE = 63.53$, $p = .02$ (effect size = 0.92); TWS–3 Predictable Words, $F(1, 24) = 10.43$, $MSE = 50.30$, $p < .01$ (effect size = 1.07); and TWS–3 Unpredictable Words, $F(1, 24) = 4.38$, $MSE = 55.45$, $p < .05$ (effect size = .70). It is interesting to note that by the time of the maintenance probes (1 year from the start of treatment), students in both conditions had improved their normative standing on each of these measures (see Table 4). In comparison with the test's normative sample, standard scores for children in the spelling condition rose by more than one half a standard deviation on the WIAT, almost a full standard deviation on the TWS–3 Predictable Words, and four fifths of a standard deviation on the TWS–3 Unpredictable Words (see Table 4). In contrast, standard scores for children in the math control condition rose by just one tenth of a standard deviation on the WIAT, one third of a standard deviation on the TWS–3 Predictable Words, and slightly less than one half a standard deviation on the TWS–3 Unpredictable Words.

Writing. Contrary to our expectations, the statistically significant treatment effect for improved writing fluency obtained at posttest was not replicated at maintenance. The one-way ANCOVA for WJ–R Writing Fluency subtest was nonsignificant. Both treatment conditions evidenced small gains from posttest to maintenance (see Tables 3 and 4). It is also interesting to note that the effect size for adjusted maintenance scores was 0.57, favoring the spelling condition.

Similar to the findings at posttest, supplemental spelling instruction did not result in a statistically significant effect for story length or story quality 6 months later. Over the course of a year, students in both conditions almost tripled the length of their stories, whereas the quality of students' stories improved by slightly more than 2 points on an 8-point scale (see Table 4).

Reading. Contrary to our predictions, the statistically significant treatment effect for improved word-attack skills obtained at posttest was not replicated at maintenance. The one-way ANCOVA for the WRMT–R Word Attack subtest was nonsignificant. Although the normative standing of students in the spelling condition dropped slightly (by one half of a point) from posttest to maintenance, the relative standing of students in the math control condition improved by slightly more than 2 standard points during this same time period (see Tables 3 and 4). It is interesting to note, however, that the effect size for adjusted maintenance scores was 0.47, favoring the spelling condition.

For the WRMT–R Word Recognition subtest, the test for homogeneity of regression was significant at maintenance. Because the homogeneity of regression assumption was not met, the Johnson-Neyman procedure was used to establish regions of statistical significance for the difference between treatment conditions' regression lines relating to the pretest covariate (X) to the maintenance outcome (Y). Regression lines were determined separately for the spelling and math conditions as $Y' = 67.05 + 294.00X$ and $Y' = 16.81 + 1.17X$, respectively. These are plotted in Figure 1. The point of intersection occurs when $X = 96.17$ and $Y = 95.32$. As can be seen in Figure 1, to the left of that point, students in the spelling condition scored higher on maintenance, whereas to the right of that point, students in the math condition scored higher. Near the point of intersection, group differences between the lines are not statistically significant. The Johnson-Neyman technique establishes the regions to the left and the right of the intersection where statistically significant group differences exist. Following the strategy outlined in Pedhazur (1982), only one such region of statistical significance ($p < .05$) was observed to fall within the observed range of the pretest. This is the region below (to the left of) $X = 89.07$, in which students in the spelling condition scored higher on maintenance than students in the math condition. To the right of this value and continuing throughout the pretest range, there is no statistically significant difference between the two conditions (see Figure 1). Therefore, students whose initial pretest word-recognition performance was 89 or below

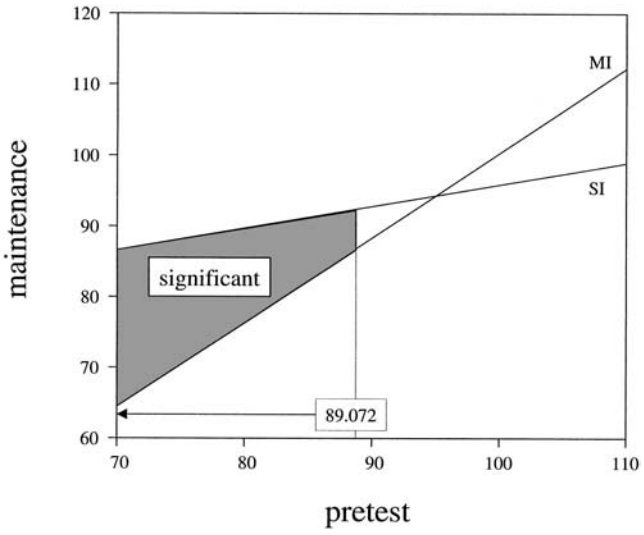

Figure 1. Interaction between pretest and treatment conditions at maintenance. MI = math instruction; SI = spelling instruction.

benefited significantly more from spelling instruction than math instruction at maintenance.

Discussion

In this study, we examined the links between spelling and the skills of learning to write and read. Previous research has shown that individual differences in spelling skills are related to individual differences in both writing and reading (see Ehri, 1987; Graham et al., 1997) and that spelling instruction can boost children's reading performance (e.g., Berninger et al., 1998; Ehri & Wilce, 1987; Uhry & Shepherd, 1993). With the exception of a single study by Berninger et al. (1998), however, there is little data on the effects of spelling instruction on children's writing. To assess the relationships between spelling and early writing and reading development, we provided supplemental spelling instruction to second-grade children, who were experiencing difficulty with spelling, and then assessed the immediate as well as long-term effects of such instruction on their spelling, writing, and reading performance.

Spelling

The spelling instruction provided to the participating students in this study was based on models of spelling (Ehri, 1986; Jorm, 1983; Simon, 1976) that emphasized two sources of information for spelling words: lexical knowledge (i.e., memory for the spelling of specific words) and knowledge of the spelling system (e.g., knowledge of sound–letter combinations, how to segment pronunciations into phonemes, spelling patterns and rules, etc.). Lexical knowledge was addressed by teaching students how to spell words that young children frequently use in their writing. Knowledge of the spelling system was approached by teaching students common sound–letter combinations, spelling patterns or rules, and frequently occurring phonograms or rimes. Data collected before, during, and after instruction showed that the supplemental spelling program was effective in boosting both kinds of knowledge. In terms of lexical knowledge, students correctly spelled almost 90% of the high-frequency words they studied at the end of each instructional unit and continued to spell 84% of the words they mastered 2 weeks later.

Similar results were obtained with two measures that assessed knowledge of the spelling system. Students who received supplementary spelling instruction were able to identify correctly more sound–letter combinations under timed conditions than their peers in the contact control condition (i.e., math instruction). On the word-building test administered at the end of each unit, they also correctly spelled twice as many words as their math control counterparts. This measure assessed students' skills in applying the onsets and rimes taught during spelling instruction.

Students in the spelling condition further outscored their math counterparts on three norm-referenced spelling tests administered immediately following instruction and 6 months later. These findings provide additional support that the spelling program enhanced students' knowledge of the spelling systems as well as their lexical knowledge, as it resulted in improvements in the ability to spell words that do and do not follow the rules of standard American English (i.e., predictable and unpredictable spelling words).

It is interesting to note that the relative standing of students in the contact control condition remained generally constant from the start to the end of instruction on all three norm-referenced spelling measures but that their relative standing improved on two of the three measures 6 months later at maintenance. One interpretation of this data is that, given enough time, weaker spellers eventually catch up with their better spelling counterparts. This contention is not justified for three reasons. One, students' relative standing did not improve on all three of the measures. Two, when relative standing did increase, it was on the two measures with the lowest initial mean pretest scores, suggesting that these changes may be due, at least in part, to regression to the mean. Three, in longitudinal studies of literacy development, poor spellers tend to remain poor spellers (e.g., Juel, 1988; Maeland & Karlsdottir, 1991).

Spelling and Writing

The findings from the present study indicate that there is a link between learning to spell and writing development. Immediately following instruction, students who received supplementary spelling instruction not only outperformed their counterparts in the math control condition on measures assessing spelling performance but evidenced greater progress in their ability to construct written sentences. Consequently, the effects of spelling instruction transferred to writing, resulting in improvements in children's text-production skills. These findings extend the previous Berninger et al. (1998) study by showing that spelling effects can transfer to more than just growth in writing output and that such effects can occur with a more academically challenging sample of students.

McCutchen (1995) and others (Berninger, 1999; Graham & Harris, 2000) have argued that spelling and other transcription skills impose a significant drain on the attentional resources of young writers and, consequently, limit their use of other writing processes. Although we were unable to assess the attentional resources needed for spelling, the findings from this study are consistent with this view, as children's sentence-writing skills improved as a result of spelling instruction. It is important to note,

however, that supplemental spelling instruction did not enhance the overall length or quality of students' stories in this investigation. Furthermore, supplemental spelling instruction had a positive effect on the writing output of only one of the seven spelling treatment groups in the study by Berninger et al. (1998). Thus, additional research is needed to replicate the present findings as well as to explore the connection between spelling and other writing skills.

Although spelling instruction enhanced children's sentence-writing skills immediately following instruction, there was no statistically significant difference between students in the spelling and math conditions on this or the other two writing measures (i.e., length and quality) 6 months after treatment ended. It is important to note, however, that the gains that students in the spelling condition made in sentence writing did not deteriorate over time, as their standard scores on this norm-referenced measure administered at maintenance were virtually identical to their scores at posttest.

Spelling and Reading

The findings from the present study also indicate that there is a link between learning to spell and reading development. Following instruction, students in the spelling condition outperformed their peers in the math control condition on a test measuring reading word-attack skills. Six months after instruction ended, transfer effects to word recognition were also obtained for students who had the lowest standard score (89 or below) on this measure at pretest. These findings are consistent with previous research showing that spelling instruction can enhance reading performance (e.g., Berninger et al., 1998; Ehri & Wilce, 1987; Uhry & Shepherd, 1993).

Ehri (1987, 1989) proposed that increased spelling knowledge helps children make sense of the words they read, making it easier to remember them, as it allows them to make more complete connections between spelling and pronunciations of words in memory. The findings from the present study and the Berninger et al. (1998) investigation are generally consistent with this view, as increased spelling knowledge was associated with improvements in word-recognition skills. However, in the Berninger et al. (1998) study, spelling effects on word recognition were obtained immediately following treatment, whereas in the present investigation, they were not evident until the maintenance test 6 months later, and then only for those students who initially had the poorest word-recognition skills. Although these differences may be related to student characteristics (the present sample was more academically challenging), it should be noted that students' performance on the word-recognition measure at the start of the two studies was generally equivalent, both in terms of mean performance and variability. Differences in outcomes were more likely due to differences in the spelling programs used in the two studies. In the present study, spelling instruction balanced the study of individual spelling words with exercises designed to foster the mastery of common sound–letter associations, rimes, and spelling rules. Students in the Berninger et al. (1998) study also learned common sound–letter associations in isolation, but most of their instruction focused on learning individual spelling words and making connections between the sounds and letters within those words. This elevated focus on individual words as the locus of spelling instruction in the Berninger et al. (1998) study may have facilitated more immediate transfer effects to word recognition, impacting not only the students with the poorest word recognition skills initially but the sample as a whole.

Unfortunately, it is not clear what impact spelling instruction had on children's word-attack skills in the Berninger et al. (1998) study, as this was not assessed. Ehri (1987, 1989) argued that increased spelling knowledge enhances children's phonological awareness skills and their grasp of the alphabetic principle, leading to improved decoding or word-attack skills. Although we did not assess the mechanisms (e.g., phonological awareness) that presumably facilitate improvements in word-attack skills, the findings from this study (at least in the short term) are consistent with Ehri's view. Children's word-attack skills improved as a result of spelling instruction that included activities designed, in part, to strengthen their phonological awareness skills, their grasp of the alphabetic principle, or both. With the phonics warm-up activity, for example, the tutor had to isolate and name the sound in a pictured word, whereas the tutee had to name the letter that represented that sound. Likewise, with the word-sorting activity, children were asked to analyze carefully the sounds in words and to examine the connections between sounds and letters.

Although spelling instruction enhanced children's word-attack skills immediately following instruction, there was no statistically significant difference between students in the spelling and math conditions 6 months later at maintenance. As with sentence-writing skills, however, the gains that students in the spelling condition made in word attack did not deteriorate over time, as their standard scores on this norm-referenced measure administered at maintenance dropped less than a single point during the intervening 6 months. Children in the math condition evidenced a slight rise during this period, but their average test scores still remained slightly below initial pretest levels.

Educational Implications

Periodically, there are calls to downplay or even eliminate spelling instruction as part of the school program (see Graham, 2000). In recent years, this has been evident in approaches such as whole-language and process writing, where greater emphasis is placed on informal and incidental methods for learning to spell and much less emphasis is placed on explicit and systematic instruction. The findings from the present study as well as the investigation by Berninger et al. (1998), however, indicate that explicit and systematic instruction is an important component of an effective spelling program for weaker spellers. Poor spellers in both of these studies became better spellers when they received extra spelling instruction. Gains in spelling performance were particularly pronounced in the present study, where the normative standing of spelling-instructed students evidenced considerable improvement on three norm-referenced measures of spelling. By the end of instruction, spelling-instructed students improved their normative standing on these tests by one third to one half of a standard deviation. In contrast, students in the contact control condition made less than one tenth of a standard deviation improvement on any of these measures. Six months later, spelling-instructed students' normative gains had increased to almost one full standard deviation on two of the spelling measures and more than one half a standard deviation on the other.

Although the present study supports the use of explicit and systematic instruction for poor spellers, it does not negate the importance of informal and incidental methods in learning to spell. As Graham (2000) indicated in a recent review, children learn the spelling of many words through reading and writing and acquire others through the use of informal teaching methods, such as teaching a spelling skill when the need arises. These approaches are less efficient for poor spellers, however (Graham, 1999). For example, poor spellers learn to correctly spell only a small portion of the words they are not directly taught during a given school year (e.g., Morris, Blanton, Blanton, & Perney, 1995). They are also less likely than good spellers to learn the correct spelling of a word by encountering it in text (Gilbert, 1934a, 1934b, 1935; Ormrod, 1986). Consequently, teachers may need to adjust how much emphasis they typically place on explicit instruction and the use of informal and incidental learning methods when working with weaker spellers.

With the advent of alternative modes of composing, such as word processing and speech synthesis, some teachers may assume that spelling can be ignored altogether. As one teacher told Steve Graham, "My kids don't need to worry about correct spelling, because they can get the right spelling by using the spell checker on the computer . . . soon they will be able to just say what they want to write." Although the use of spell checkers and speech synthesis has clearly increased in recent years, beginning writers still, and for the foreseeable future, do most of their writing by hand. Furthermore, spell checkers are only partly successful in eliminating spelling miscues. For instance, in a study by MacArthur, Graham, Haynes, and De La Paz (1996), poor spellers were only able to correct 37% of the spelling errors in their compositions when using a spell checker. Because poor spelling may impair communication, negatively influence teachers' evaluations (Marshall & Powers, 1969), and interfere with writing and its development (Berninger et al., 1998; Graham, 1990), these skills should not be avoided or downplayed until children do most of their writing with technological tools that provide a high rate of correct spelling.

The outcomes from the present study and other investigations (e.g., Berninger et al., 1998; Bradley & Bryant, 1983; Ehri & Wilce, 1987) indicate that extra spelling instruction may be an important ingredient in addressing the writing and reading problems experienced by some young children, especially those who have difficulty with spelling. Although extra spelling instruction did not influence long-term outcomes for most of the writing and reading skills assessed in this experiment, it did facilitate the initial acquisition of two critical writing and reading skills. For children who find writing and reading challenging, improvements in their sentence-writing and word-attack skills are significant, as both of these skills are essential building blocks in literacy development (see Adams, 1990; Berninger, Mizokawa, & Bragg, 1991; Hayes & Flower, 1986). Similarly, improvement in word-recognition scores, even when this takes longer to accrue and is limited to children who have the most difficulties with these skills, are beneficial.

Additional research is needed to replicate the present findings and to investigate the effectiveness of other techniques for preventing writing and reading difficulties. Such investigations are especially needed in writing, as educators and researchers have devoted little attention to the prevention of writing disabilities (Graham & Harris, 2002). Several recent studies indicate that early, supplementary handwriting instruction may be an important factor in the prevention of writing difficulties, as such instruction has led to immediate as well as long-term gains in the writing fluency of poor writers (Berninger et al., 1997; Graham et al., 2000; Jones & Christensen, 1999). Improvements in writing fluency may be especially important for struggling writers, as research by Berninger and her colleagues (Berninger et al., 1991) indicate that problems with writing fluency in the primary grade may be the genesis for writing difficulties in the upper grades.

We further anticipate that early and extra instruction in the self-regulatory aspects of writing, particularly planning and revising, will help prevent writing problems. Our own research has shown that poor writers experience difficulties regulating these processes when composing (De La Paz, Swanson, & Graham, 1998; Graham, 1997; Graham & Harris, 2000) and that teaching these processes to older elementary grade students who are poor writers improves how much and how well they write (Graham & Harris, 1996). It is also likely that efforts designed to improve the quality of the regular writing program (e.g., Englert et al., 1995) will be beneficial. Such instruction should reduce the number of cases of writing failure that are due to poor instruction and help to ameliorate the severity of writing difficulties experienced by other children whose primary problems are not instructional.

Finally, students in the contact control condition were taught math skills using a modified version of the Peer-Assisted Learning Strategy (PALS; Fuchs et al., 1994, 1995). Typically, this approach involves pairing a high-achieving student with a low-achieving student and having each child act as a tutor to the other. In the present study, we modified this format by (a) having two weaker students work together and (b) providing closer supervision of the tutoring process. We further altered PALS by adding an activity, math warm-up, to the start of each lesson. This activity provided extra practice in learning addition and subtraction facts. Although we did not compare our modified version of PALS with the original, the approach used in this study was successful in improving two of the math skills targeted for instruction: basic math facts and computation problems. There was, however, no statistically significant difference between spelling- and math-instructed students in the average number of word problems solved. Performance for both groups was quite low, averaging just one and two problems correct for the spelling and math conditions, respectively. Thus, additional modifications in the program may be needed to improve problem-solving skills when PALS is used with pairs of low-achieving students.

References

Abbington, A. (1932). *Pocket book of bonners: An omnibus of schoolboy howlers and unconscious humor.* New York: Pocket Books.

Adams, M. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Bear, D., Invernizzi, M., Templeton, S., & Johnston, F. (2000). *Words their way: Word study for phonics, vocabulary, and spelling instruction* (2nd ed.). Upper Saddle River, NJ: Merrill.

Berninger, V. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disability Quarterly, 22,* 99–112.

Berninger, V., Mizokawa, D., & Bragg, R. (1991). Theory-based diagnosis

and remediation of writing disabilities. *Journal of School Psychology, 29,* 57–97.

Berninger, V., Vaughn, K., Abbott, R., Abbott, S., Rogan, L., Brooks, A., et al. (1997). Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. *Journal of Educational Psychology, 89,* 652–666.

Berninger, V., Vaughn, K., Abbott, R., Brooks, A., Abbott, S., Rogan, L., et al. (1998). Early intervention for spelling problems: Teaching functional spelling units of varying size with a multiple-connections framework. *Journal of Educational Psychology, 90,* 587–605.

Bradley, L., & Bryant, P. (1983). *Rhyme and reason in reading and spelling.* Ann Arbor: University of Michigan Press.

Bryant, P., & Bradley, L. (1980). Why children sometimes write words which they do not read. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 353–370). London: Academic Press.

Calkins, L. (1985). "I am one who writes": New approaches to children's writing. *American Educator, 9,* 26–29, 42–44.

Carroll, J., Davies, P., & Richman, B. (1971). *The American heritage word frequency book.* Boston: Houghton Mifflin.

Carrow-Wolfolk, E. (1985). *Test for Auditory Comprehension of Language.* Allen, TX: Developmental Learning Materials.

De La Paz, S., Swanson, P., & Graham, S. (1998). Contribution of executive control to the revising problems of students with writing and learning difficulties. *Journal of Educational Psychology, 90,* 448–460.

Deno, S., Marston, D., & Mirkin, P. (1982). Valid measurement procedures for the continuous evaluation of written expression. *Exceptional Children, 48,* 368–371.

Dunn, L., & Dunn, M. (1997). *Peabody Picture Vocabulary Test—III.* Circle Pines, MN: American Guidance.

Ehri, L. (1986). Sources of difficulty in learning to spell and read. In M. Wolraich & D. Routh (Eds.), *Advances in developmental and behavioral pediatrics* (pp. 121–195). Greenwich, CT: JAI Press.

Ehri, L. (1987). Learning to read and spell words. *Journal of Reading Behavior, 19,* 5–31.

Ehri, L. (1989). The development of spelling knowledge and its role in reading acquisition and reading disabilities. *Journal of Learning Disabilities, 22,* 356–365.

Ehri, L., & Wilce, L. (1987). Does learning to spell help beginners learn to read real words? *Reading Research Quarterly, 18,* 47–65.

Englert, C., Garmon, A., Mariage, T., Rozendal, M., Tarrant, K., & Urba, J. (1995). The early literacy project: Connecting across the literacy curriculum. *Learning Disability Quarterly, 18,* 253–275.

Farr, R., Beverstock, C., & Robbins, B. (1988). *An examination of the writing vocabulary of children in grades two through eight.* Bloomington, IN: Indiana University, Center for Reading and Language Studies. (ERIC Document Reproduction Service No. ED 300820)

Foorman, B., & Francis, D. (1994). Exploring connections among reading, spelling, and phonemic segmentation during first grade. *Reading and Writing: An Interdisciplinary Journal, 6,* 65–91.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., & Bentz, J. (l994). Classwide curriculum-based measurement: Helping general educators meet the challenge of student diversity. *Exceptional Children, 60,* 518–537.

Fuchs, L. S., Fuchs, D., Phillips, N. B., Hamlett, C. L., & Karns, K. (l995). Acquisition and transfer effects of classwide peer-assisted learning strategies in mathematics for students with varying learning histories. *School Psychology Review, 24,* 604–620.

Gardner, M. (1990). *Expressive One-Word Picture Vocabulary Test—Revised.* Austin, TX: Pro-Ed.

Gilbert, L. (1934a). Effect of reading on spelling in the ninth grade. *School Review, 42,* 197–204.

Gilbert, L. (1934b). Effects of reading on spelling in the secondary schools. *California Quarterly of Secondary Education, 9,* 269–275.

Gilbert, L. (1935). A study of the effects of reading on spelling. *Journal of Educational Research, 28,* 570–576.

Glass, G., McGaw, B., & Smith, M. (1981). *Meta-analysis and social research.* Beverly Hills, CA: Sage.

Graham, S. (1990). The role of production factors in learning disabled students' compositions. *Journal of Educational Psychology, 82,* 781–791.

Graham, S. (1997). Executive control in the revising of students with learning and writing difficulties. *Journal of Educational Psychology, 89,* 223–234.

Graham, S. (1999). Handwriting and spelling instruction for students with learning disabilities: A review. *Learning Disability Quarterly, 22,* 78–98.

Graham, S. (2000). Should the natural learning approach replace spelling instruction? *Journal of Educational Psychology, 92,* 235–247.

Graham, S., Berninger, V., Abbott, R., Abbott, S., & Whitaker, D. (1997). The role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology, 89,* 170–182.

Graham, S., & Harris, K. R. (1996). Self-regulation and strategy instruction for students with writing and learning difficulties. In S. Ransdell & M. Levy (Eds.), *Science of writing: Theories, methods, individual differences, and applications* (pp. 347–360). Mahwah, NJ: Erlbaum.

Graham, S., & Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist, 35,* 3–12.

Graham, S., & Harris, K. R. (2002). Prevention and intervention for struggling writers. In M. Shinn, G. Stoner, & H. Walker (Eds.), *Interventions for academic and behavior problems: II. Preventive and remedial approaches* (pp. 589–610). Washington, DC: National Association of School Psychologists.

Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology, 92,* 620–633.

Graham, S., Harris, K. R., & Loynachan, C. (1993). The basic spelling vocabulary list. *Journal of Educational Research, 86,* 363–368.

Graham, S., Harris, K. R., & Loynachan, C. (1996). The directed spelling thinking activity: Applications with high frequency words. *Learning Disabilities Research and Practice, 11,* 34–40.

Greene, H. (1954). *The New Iowa Spelling Scales.* Iowa City: University of Iowa.

Hammill, D., & Larsen, S. (1996). *Test of Written Language—3.* Austin, TX: Pro-Ed.

Harris, K. R. (1986). Self-monitoring of attentional behavior vs. self-monitoring of productivity: Effects on task behavior and academic response rate among learning disabled children. *Journal of Applied Behavior Analysis, 19,* 417–423.

Harris, K. R., Graham, S., Reid, R., McElroy, K., & Hamby, R. (1994). Self-monitoring of attention versus self-monitoring of performance: Replication and cross-task comparison studies. *Learning Disability Quarterly, 17,* 121–139.

Harris, K. R., Graham, S., Zutell, J., Gentry, R. (1998). *Spell-it write.* Columbus, OH: Zaner-Bloser.

Hayes, J., & Flower, L. (1986). Writing research and the writer. *American Psychologist, 41,* 1106–1113.

Henderson, E. (1978). *Teaching spelling.* Boston, MA: Houghton Mifflin.

Hogan, T., & Mishler, C. (1980). Relationships between essay tests and objective measures of language skills for elementary students. *Journal of Educational Measurement, 17,* 219–227.

Horn, E. (1960). Spelling. In *Encyclopedia of educational research* (3rd ed., pp. 1337–1354). New York: MacMillan.

Jones, D., & Christensen, C. (1999). The relationship between automaticity in handwriting and students' ability to generate written text. *Journal of Educational Psychology, 91,* 44–49.

Jorm, A. (1983). *The psychology of reading and spelling difficulties.* London: Routledge & Kregan Paul.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grade. *Journal of Educational Psychology, 80,* 437–447.

Larsen, S., & Hammill, D. (1994). *Test of Written Spelling—3.* Austin, TX: Pro-Ed.

MacArthur, C., Graham, S., Haynes, J., & De La Paz, S. (1996). Spelling checkers and students with learning disabilities: Performance comparisons and impact on spelling. *Journal of Special Education, 30,* 35–57.

Maeland, A., & Karlsdottir, R. (1991). Development of reading, spelling, and writing skills from third to sixth grade in normal and dysgraphic school children. In J. Wann, A. Wings, & N. Sovik (Eds.), *Development of graphic skills: Research, perspectives and educational implications* (pp. 177–189). London: Academic Press.

Marshall, J., & Powers, J. (1969). Writing neatness, composition errors, and essay grades. *Journal of Educational Measurement, 6,* 97–101.

McCutchen, D. (1995). Cognitive processes in children's writing: Developmental and individual differences. *Issues in Education: Contribution From Educational Psychology, 1,* 123–160.

Morris, D., Blanton, L., Blanton, W., & Perney, J. (1995). Spelling instruction and achievement in six classrooms. *Elementary School Journal, 96,* 145–162.

Newborg, J., Stock, J., & Wnek, L. (1984). *Battelle Developmental Inventory Screening Tests.* Allen, TX: LINC.

Ormrod, J. (1986). Learning to spell: Three studies at the university level. *Research in the Teaching of English, 20,* 160–173.

Pedhazur, E. (1982). *Multiple regression in behavioral research.* Fort Worth, TX: Harcourt, Brace, Jovanovich.

Pressley, M., Goodchild, F., Fleet, J., Zajchowski, R., & Evans, E. (1989). The challenges of classroom strategy instruction. *Elementary School Journal, 89,* 301–342.

Scardamalia, M., Bereiter, C., & Goleman, H. (1982). The role of production factors in writing ability. In M. Nystrand (Ed.), *What writers know: The language, process, and structure of written discourse* (pp. 173–210). New York: Academic Press.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

Simon, D. (1976). Spelling: A task analysis. *Instructional Science, 5,* 277–302.

Treiman, R. (1993). *Beginning to spell.* New York: Oxford University Press.

Uhry, J., & Shepherd, M. (1993). Segmentation/spelling instruction as part of a first grade reading program: Effects on several measures of reading. *Reading Research Quarterly, 28,* 218–233.

Wallace, G., & Hammill, D. (1994). *Comprehensive Receptive and Expressive Vocabulary Tests.* Austin, TX: Pro-Ed.

Wechsler, D. (1967). *Wechsler Preschool and Primary Scale of Intelligence.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—III.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1992). *Wechsler Individual Achievement Test.* San Antonio, TX: Psychological Corporation.

Woodcock, R. (1987). *Woodcock Reading Mastery Test—Revised.* Circle Pines, MN: American Guidance Service.

Woodcock, R., & Johnson, M. (1990). *Woodcock−Johnson Psycho-Educational Battery—Revised.* Chicago: Riverside.

Zutell, J. (1993). Directed spelling thinking activity: A developmental, conceptual approach to advance spelling word knowledge. In *Conference papers: Literacy for the new millennium.* Melbourne, Australia: First International Conference and 19th Conference of the Australian Reading Association.

Appendix

Summary of the Sequence of Activities in Each Unit in the Spelling Treatment

Lesson 1

Word Sorting: Instructor and children sort words into the spelling patterns emphasized in the unit.

Word Hunt: Children are asked to hunt for words that fit the spelling patterns presented in the word sort.

Lesson 2

Word Hunt Check: Students identified words they found at school–home that fit the spelling-unit patterns.

Phonics Warm-Up: Students practice common sound–letter combinations.

Introduce Spelling Words: Students are asked to learn eight previously misspelled high-frequency words that correspond to the spelling-unit patterns.

Word Study: Students study new words using the five-step procedure and graph number of correct practices.

Word Building: Students build words using a rime that corresponds to one of the spelling-unit patterns.

Word Hunt: Children are asked to hunt for words that fit the spelling patterns presented in the word sort.

Lesson 3

Word Hunt Check: Students identified words they found at school–home that fit the spelling-unit patterns.

Phonics Warm-Up: Students practice common sound–letter combinations.

Word Study: Students continue to use the five-step procedure to study words assigned in Lesson 2 but try to increase their number of correct practices, assessing this goal by graphing the results.

Word Building: Students build words using a second rime that fits one of the spelling-unit's patterns.

Word Hunt: Children are asked to hunt for words that fit the spelling patterns presented in the word sort.

Lesson 4

Word Hunt Check: Students identified words they found at school–home that fit the spelling-unit's patterns.

Phonics Warm-Up: Students practice common sound–letter combinations.

(*Appendix continues*)

Word Study: Students studied the words assigned in Lesson 2 but study takes place within a game format.

Word Building: Students build words using a third rime that fits one of the spelling-unit's patterns.

Word Hunt: Children are asked to hunt for words that fit the spelling patterns presented in the word sort.

Lesson 5

Word Hunt Check: Students identified words they found at school–home that fit the spelling-unit's patterns.

Phonics Warm-Up: Students practice common sound–letter combinations.

Word Study: Students use either the five-step procedure or the spelling game format to study words.

Word Building: Students build words using the rimes introduced in the three previous lessons.

Lesson 6

Unit Spelling Test: Students tested on their mastery of the eight spelling words introduced in Lesson 2.

Correction of Test: Students correct any misspelled words, graph number correct, and set goal to spell all new words correctly on next unit test.

Word Building Test: Students tested on their ability to spell words that contain the three rimes used in the word-building activity.

Review: Starting with Unit 2, students review spelling patterns and skills introduced in earlier units.

Received May 2, 2001
Revision received December 17, 2001
Accepted January 11, 2002 ■