Assessing Specific Grapho-Phonemic Skills in Elementary Students
Kelly P. Robbins, John L. Hosp, Michelle K. Hosp and Lindsay J. Flynn
Assessment for Effective Intervention 2010 36: 21 originally published online 5 October 2010
DOI: 10.1177/1534508410379845

The online version of this article can be found at:
http://aei.sagepub.com/content/36/1/21

Published by:
Hammill Institute on Disabilities

http://www.sagepublications.com

Additional services and information for Assessment for Effective Intervention can be found at:
Email Alerts: http://aei.sagepub.com/cgi/alerts
Subscriptions: http://aei.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav

Version of Record - Nov 11, 2010
OnlineFirst Version of Record - Oct 5, 2010
What is This?
Assessing Specific Grapho-Phonemic Skills in Elementary Students

Kelly P. Robbins, 1 John L. Hosp, 2 Michelle K. Hosp, 2 and Lindsay J. Flynn 3

Abstract

This study examines the relation between decoding and spelling performance on tasks that represent identical specific grapho-phonemic patterns. Elementary students (N = 206) were administered a 597 pseudoword decoding inventory representing 12 specific grapho-phonemic patterns and a 104 real-word spelling inventory representing identical grapho-phonemic patterns presented on the decoding inventory. Correlational and quantile regression analyses revealed a moderate to strong correlation between student performance in decoding and spelling tasks, with stronger relations found among more complex grapho-phonemic patterns and weaker relations found among less complex grapho-phonemic patterns. The results of this study support the utility of a spelling assessment with items representing a wide range of grapho-phonemic patterns as a screener for specific areas of grapho-phonemic knowledge weakness and provide educators with an efficient method of collecting the data needed to develop targeted interventions.

Keywords

assessment, decoding, spelling, grapho-phonemic skills

Accurately and efficiently assessing grapho-phonemic knowledge (i.e., the knowledge of letter–sound correspondences) is a vital component of designing effective instruction because it enables teachers to identify the specific grapho-phonemic skills that students have mastered and those that they are lacking. Research indicates that a moderate to strong relation (i.e., .32–.84) exists between whole word decoding and spelling when comparing both real and pseudowords (Guthrie, 1973; Juel, Griffith, & Gough, 1986; Nagy, Berninger, Abbott, Vaughan, & Vermeulen, 2003). Decoding and spelling skills have also been shown to follow a similar path of development (Ehri, 1998; Henderson, 1981). Furthermore, students who have a firm understanding of the predictability of the relation between phonemes in speech and graphemes in print are more likely to be proficient at decoding and spelling (Waters, Bruck, & Seidenberg, 1985). To accurately and efficiently assess grapho-phonemic knowledge, teachers must understand the developmental path that typifies decoding and spelling skill acquisition.

Development in Decoding and Spelling

Decoding and spelling development typically progress through four phases, beginning with mastery of simplistic grapho-phonemic patterns (e.g., consonant–vowel–consonant patterns), followed by mastery of more complex grapho-phonemic patterns (e.g., prefix and suffix patterns; Chall, 1996; Ehri, 1986; Henderson, 1981). In Phase 1 (i.e., the prealphabetic or preliterate phase) words are identified visually, and decoding and spelling errors are typified by errors based on visual cues, which do not demonstrate application of grapho-phonemic correspondence knowledge (Ehri, 1998; Henderson & Templeton, 1986). In Phase 2 (i.e., partial alphabetic phase or the letter name phase) unfamiliar words are rarely decoded or spelled accurately because of an incomplete understanding of the alphabetic system of grapho-phonemic correspondences (Ehri, 2000). In Phase 3 (i.e., full alphabetic phase or within-word pattern phase) students gain knowledge of the alphabetic principle and decode and spell systematically, one phoneme corresponding to one grapheme or by analogy (Ehri, 1998). In Phase 4 (i.e., consolidated alphabetic or syllabic juncture phase) students become more efficient at decoding and spelling as they are able to identify and spell common letter sequences, or groups that represent phonemes (i.e., vowel teams and affixes; Ehri, 1997; Henderson & Templeton, 1986).

1University of Utah, Salt Lake City, UT, USA
2University of Iowa, Iowa City, IA, USA
3University of California, Riverside, Riverside, CA, USA

Corresponding Author:
Kelly P. Robbins, University of Utah, 1705 Campus Dr., Rm 221 Milton Bennion Hall, Salt Lake City, UT 84112
Email: krobbins.utah@gmail.com
Understanding the role of grapho-phonemic correspondence in decoding and spelling is a necessary prerequisite for teachers who are interested in accurately identifying students’ phase of grapho-phonemic knowledge development.

The Role of Grapho-phonemic Correspondence in Decoding and Spelling

Research indicates that word spellings are predictable 50% of the time by grapho-phonemic correspondence alone (Fry, 2004). Furthermore, the level of complexity of specific grapho-phonemic correspondences has been shown to affect student decoding and spelling performance (Apel, Wolter, & Masterson, 2006; Willson, Rupley, Rodriguez, & Mergen, 1999). This complexity is typically found in the representation of vowels. For example, words where single vowel phonemes are represented by a single vowel grapheme (e.g., /ū/ in bug is represented by U) are less complex than words where single vowel phonemes are represented by grapheme sequences (e.g., the /ū/ in bright is represented by IGH). Moreover, research shows that students in first and second grade decode words with fewer phonemes more accurately than words with many phonemes (Willson et al., 1999). The frequency of the phoneme and grapheme sequences within a word has also been shown to influence the ability of students to decode and spell pseudowords (Apel et al., 2006). For example, students decoded and spelled pseudowords with high frequency phoneme and grapheme sequences (e.g., hess) more accurately than words with low frequency phoneme and grapheme sequences (e.g., gouz). The impact of grapho-phonemic complexity on decoding and spelling must be considered to identify the most efficient and accurate assessment tools.

Assessment of Decoding and Spelling

Decoding and spelling are commonly assessed using norm-referenced (e.g., Word Attack subtest of the Woodcock–Johnson Tests of Achievement–Third Edition [WJIII; Woodcock, McGrew, & Mather, 2001] and the Spelling subtest of the Wide Range Achievement Test–Revised–Third Edition [Wilkinson, 1993]), criterion-referenced, qualitative tests (e.g., Nonsense Word Fluency subtest of the Dynamic Indicators of Basic Early Literacy Skills [NWF, DIBELS; Good & Kaminski, 2002] and Words Their Way spelling inventory [WTW; Bear, Invernizzi, Templeton, & Johnston, 2008]), and informal tests (e.g., decoding inventories designed by researchers for one-time use in a research study or by teachers for classroom use only). Many of these inventories are ordered for grapho-phonemic complexity, include 20 to 40 items, and often provide only a single opportunity for students to decode or spell a specific grapho-phonemic pattern. As a result, teachers and researchers are left with an incomplete assessment of student grapho-phonemic knowledge. Decoding inventories include real words or pseudowords (i.e., words that are formed from common grapho-phonemic patterns but are not real words; e.g., nan, zoo). Assessing decoding with isolated word lists, particularly using pseudowords, is the purest test of decoding because students are not able to rely on context (Wren, 2002) or memory (Greenberg, 1997) when decoding isolated words. These inventories typically require one-on-one testing. This is a time-consuming process and often reduces the amount of time available for instruction.

Spelling inventories are most commonly composed of isolated word lists; however, they typically include real words (most likely because of the vast number of graphemes that students could produce to plausibly represent individual phonemes). Assessing spelling with isolated word lists is a theoretically sound test of spelling skill because of the lack of composition interference affecting attention when spelling isolated words (Schlagel, 1989). Spelling inventories typically allow for group administration, conserving time and improving the efficiency of grapho-phonemic skill assessment. Despite the alignment between typical decoding and spelling assessments, current assessment practices do not fully take advantage of the relationship between decoding and spelling to create the most accurate and efficient grapho-phonemic knowledge assessment.

Rationale and Purpose for This Study

The purpose of this study was twofold. One aspect of this investigation compares student performance on identical, specific grapho-phonemic patterns in pseudoword decoding tasks and real-word spelling tasks. The second aspect investigates this relation across students with varying reading performance levels. Specifically, the research questions included the following: (a) What is the relation between elementary students’ decoding and spelling performance on tasks that represent identical specific grapho-phonemic patterns? and (b) What is the relation between decoding and spelling performance on tasks that represent identical specific grapho-phonemic patterns in elementary students of differing reading performance levels? Although past research has investigated the overall relation between decoding and spelling at the whole word level (Guthrie, 1973; Juel et al., 1986; Nagy et al., 2003), this study adds to the body of literature by investigating how closely these skills are related on the level of specific grapho-phonemic patterns across students of differing skill levels. This investigation offers insight into improving the efficiency of identifying student strengths and weaknesses in relation to specific grapho-phonemic patterns through the use of an integrated assessment approach.
This approach could be helpful to inform the development of targeted instructional interventions and improved student reading and spelling outcomes.

Method

Participants

This study took place within the context of a larger study. The original sample from the larger study consisted of 265 second-, third-, and fourth-grade English-only-speaking students, some with and some without disabilities, sampled from 23 classes in two urban schools and one suburban elementary school in two states in the intermountain west. Participants consisted of equal proportions of males and females (50% and 50%, respectively). A total of 75 participants (28.3%) qualified for free or reduced-price lunch, 29 (10.9%) did not qualify, and 161 (60.7%) did not report qualification status. The ethnic backgrounds of the students included 60.4% White, 18.4% Latino, 6.0% African American, 6.0% other backgrounds including Asian and Pacific Islander, and 9.1% not reported. In all, 19 participants (7.2%) had a special education classification. Through a multistage sampling procedure, three fourths of the original sample was randomly selected for participation in the current study. This resulted in a sample of 206 students. Participants consisted of equal proportions of males and females (48% and 52%, respectively). A total of 67 participants (32.5%) qualified for free or reduced-price lunch, 25 (12.1%) did not qualify, and 114 (55.3%) did not report qualification status. The ethnic backgrounds of the students included 61.6% White, 18.9% Latino, 5.8% African American, 6.9% other backgrounds including Asian and Pacific Islander, and 6.8% not reported. In all, 15 participants (7.3%) had a special education classification. According to diagnostic missing data analysis, fewer than 1% (0.2%) of data points were missing (308 of 139,365). Examination of frequency distributions and demographics did not reveal any discernable pattern of missingness (McKnight, McKnight, Sidani, & Figueredo, 2007).

Instruments

Decoding inventory. A decoding inventory (Hosp, Hosp, & Howell, 2010) consisting of one- to two-syllable pseudo-words based on the most common grapho-phonemic patterns in written English was used to assess decoding performance. Pseudowords were used on all subtests except one (Contractions) to ensure readers did not read the words from memory or sight. The words in the decoding inventory (Hosp, Hosp, & Howell, 2010) were 104 common word patterns identified from research-based lists of frequently occurring grapho-phonemic word patterns (Fry, 2004; Fry & Kress, 2006; Vaughn & Linan-Thompson, 2004). Each word pattern was presented in five different words, resulting in 597 items. These word patterns were consolidated into the following subtests: Consonant–Vowel–Consonant (CVC), Consonant–Vowel–Consonant–Consonant (CVCC), Consonant–Vowel–Consonant–Silent e (CVC-e), R-controlled, Blends, Digraphs, Vowel Teams, Two Closed Syllable (CVCCVC), Prefixes, Suffixes With Short Vowel in Base Word, Suffixes With Long Vowel in Base Word, and Contractions. In an attempt to control for an order effect, subtests were randomly ordered to create four separate testing protocols (Protocols A–D) using a random numbers table. The participants were randomly divided so that 25% were tested in each group to make up the 100% of participants tested.

To determine internal consistency of the decoding inventory, the internal reliability coefficient alpha was calculated for each subtest (Cronbach, 1951). Correlations ranged from .857 to .963. To determine test–retest reliability of the decoding inventory, 25% of the students (those tested with Protocol B) were retested 1 to 3 weeks from their initial test date. Correlations were calculated between scores for each specific grapho-phonemic pattern from the initial test session and the second test session. Reliability ranged from .748 to .951 (n = 59). Criterion validity of the decoding inventory was also assessed. Correlations were calculated between the Word Attack subtest of the WJIII (Woodcock et al., 2001) and each grapho-phonemic pattern on the decoding inventory. Correlations ranged from .500 to .752.

Spelling inventory. The spelling inventory (Hosp, Hosp, & Robbins, 2010) consisted of short real words with the same grapho-phonemic pattern represented in the decoding inventory. Words were ordered from simple to complex (beginning with the CVC pattern and ending with prefix or suffix patterns), and each word pattern was presented once, resulting in 104 items. Criterion-related validity was assessed for the spelling inventory. Correlations were calculated between the spelling inventory and the WTW spelling inventory (Bear et al., 2008). Correlations ranged from .379 to .872. To determine internal consistency of the spelling inventory, the internal reliability coefficient alpha (Cronbach, 1951) was calculated for each grapho-phonemic pattern. Correlations ranged from .487 to .897.

Procedures

Administration. Examiners included one author of the inventories, one associate professor, and seven doctoral research assistants. The decoding inventory (Hosp, Hosp, & Howell, 2010) was individually administered in a quiet study carrel in the school library. The following scripted instructions were read at the beginning of each test session:
I want you to read some words to me. These are not real words, they are made-up words. I want you to try your best and read each made-up word. Point to each word as you read it. Start with the first word here (point to the first word) and read across the page. Be sure to do your best reading. Any questions? (Answer any questions). Put your finger on the first word. Please begin.

Other standardized administration procedures were also followed, including the provision of (a) a 3-s wait time prior to prompting the student to continue, (b) a specific prompting technique (i.e., pointing to the next word and asking, “What word?”), and (c) start and discontinue rules (i.e., the starting point was the first word on the page and testing was to continue for 1 min, or until the last stimulus item had been read). Testing sessions ranged from 20 min to 45 min. If students required multiple sessions to complete the test items, no more than 5 days separated administration sessions to control for potential learning effects.

The spelling inventory was group administered to all participants in their classrooms with two examiners present. Standardized administration procedures were followed, including (a) a statement to the students that they were not being tested on the words but helping people understand how they can help all students learn to spell, (b) an oral presentation of each spelling word, (c) an oral presentation of the word in a sentence, (d) repeated oral presentation of the spelling word, and (e) a 10-s wait time between words. Examiners were also instructed to “test under typical testing conditions according to the classroom (allow carols/study screens/offices) if they are normally used for testing” (Hosp, Hosp, & Robbins, 2010). Each testing session lasted between 25 and 30 min.

Training To ensure decoding and spelling administration fidelity, examiners were trained and given feedback on the standardized administration procedures outlined above. Ongoing checks of examiner pronunciation were conducted throughout data collection to attempt to control for examiner drift (i.e., movement away from standardized administration and scoring procedures). To ensure accurate decoding scoring, examiners were trained on scoring procedures including correct pronunciation of all decoding items and a standardized method for indicating errors (i.e., a slash through the exact grapheme that the student mispronounced). In addition, digital recordings of the pronunciation of each decoding inventory item were provided to all examiners. During this training process, all examiners also participated in the development of scoring decision rules. Furthermore, to ensure accurate spelling scoring, examiners were trained on scoring procedures including a standardized method for indicating errors.

Scoring All words read on the decoding inventory were scored during the test administration for whole word accuracy. Standardized scoring rules were followed, including the following: (a) mispronunciation of any sound of the word resulted in an error, (b) skipped words counted as errors, (c) and all errors were marked with a slash through the word. All testing sessions were recorded on digital audio recorders and checked for reliability of scoring. Examiners rescored 20% of the tests from the digital recordings; 96% interrater reliability was determined.

All words spelled on the spelling inventory were scored for whole word accuracy. Standardized scoring rules were followed, including the following: (a) to be counted as correct the student response had to include all necessary letters, in the correct order, (b) capitalized letters did not count as errors, and (c) if two letters were overlapping (one over the other), the darkest letter was considered the student’s intended response. Spelling scoring fidelity checks were conducted, and specific decision rules were developed. Examiners rescored 20% of the tests; 98% interrater reliability was determined.

Data Management and Analysis

Data were initially entered into an Excel spreadsheet and then transferred into Statistical Package for the Social Sciences (SPSS) Version 13 (http://www.spss.com) and Statistical Analysis Software (SAS) Version 9.1 to analyze student data. All statistical analyses of these data were run using SPSS and SAS. Data were entered by two researchers to increase accuracy of data entry. Of all participant scores entered, 20% were rechecked for accuracy of entry. Reliability averaged 99% for both decoding and spelling entry.

The Pearson product–moment correlation coefficient (r) was used to calculate correlation coefficients between student performance on each of the 13 decoding and 13 spelling variables (one for each of the 12 word patterns [e.g., CVC, CVCC] and one for the Total Test score). Bivariate analyses assume a normal distribution, or a linear relationship, among the variables being compared (Hays, 1994). All variable distributions were inspected for normality based on skewness and kurtosis values. To meet the linear assumption of the bivariate analysis, variables with abnormal distributions were transformed using Tukey’s ladder of re-expression (Mosteller & Tukey, 1977).

Quantile regression using the SAS Quantreg procedure was used to calculate a regression curve illustrating the relation between the predictor (i.e., spelling performance) and criterion variables (i.e., decoding performance; see Hao & Naiman, 2007, for an explanation of this statistical procedure). In general, quantile regression is useful for examining the variation, or lack thereof, in regression between two variables across a quantile distribution (see Catts, Petscher,
Schatshneider, Bridges, & Mendoza, 2009, for a discussion of quantile regression in literacy research). Quantile regression is useful in analyzing extreme portions of a distribution (higher and lower skill readers) and measuring the rate of change (i.e., slope) in student performance across a distribution (Cade & Noon, 2003). The analysis for this study involved standardizing the variables and calculating quantile regression coefficients based on observed scores (i.e., nontransformed scores) for each of the 12 decoding and spelling word patterns (e.g., CVC) as well as the total test scores for 13 total comparisons. This procedure produced 21 correlation coefficients for each comparison, one for each quantile (i.e., .01, .05 . . . .99). Quantile plots (see Figure 1) were created by transferring the coefficients onto line graphs, along with the mean correlation coefficients determined by the bivariate analysis. The decoding and spelling performance at the lower (i.e., .01–.25 quantiles) and upper (i.e., .70–.99 quantiles) tails of the distribution were then examined.

In this study, the shape of the quantile regression curves reflects the magnitude of the relation between student decoding and spelling performance. Therefore, the shapes of the quantile curves are directly related to the magnitude of the relation between student decoding and spelling performance. For example, if the relation between decoding and spelling performance was high across all students, the quantile curves would be flat (indicating a similar strength of relation between decoding and spelling performance across the distribution), with a high magnitude (indicating a strong relation between decoding and spelling performance). However, if the relation between decoding and spelling performance differed across student performance levels, the quantile curve would reflect this in variation of magnitude across the curve. For example, if students with low and high performance levels had weaker relations between their decoding and spelling performance whereas students with midrange performance levels had strong relations between decoding and spelling performance, the quantile plot would fluctuate from low to high, with a peak in the middle and low levels on both tails of the distribution.

**Results**

**Correlations for All Participants**

To answer the first research question (What is the relation between elementary students’ decoding and spelling performance on tasks that represent identical specific grapho-phonemic patterns?), descriptive statistics including values of skewness and kurtosis were calculated for each variable (see Tables 1 and 2), followed by a bivariate analysis. One decoding variable (i.e., Contractions-D) and six spelling variables (i.e., CVC-S, CVCC-S, CVCe-S, R-controlled-S, Blends-S, and Digraphs-S) had values of skewness or kurtosis that indicated a skewed and/or significantly peaked distribution (greater than 1 or less than –1). These were normalized using Tukey’s ladder of re-expression.

A wide range of correlations were identified among the variables (see Table 3). Moderate to strong correlations ranging from .490 to .790 were identified between 9 out of 13 variables (i.e., CVCe, Blends, Digraphs, Vowel Teams, CVCCVC, Prefixes, Short Vowel Suffixes, Long Vowel Suffixes, Total Test), with the highest correlation being between Total Test scores. CVC, CVCC, R-controlled, and Contractions had weaker correlations, ranging from .173 to .377.

**Correlations for Participants With Differing Reading Performance Levels**

Quantile plots based on observed, or nontransformed scores, were created to answer the second research question (What is the relation between decoding and spelling performance on tasks that represent identical specific grapho-phonemic patterns in elementary students of differing reading performance levels?). In all, 13 quantile plots (one for each of the 12 variables or word patterns [e.g., CVC, CVCC] and one for the Total Test comparison) each with 21 distinct quantile regression estimates ranging from .01 to .99 with a dashed line overlaying each plot to indicate the Pearson correlation (r) between the decoding and spelling variables for all students tested in the study. The quantile regression curves fell into three groups, based on the general shape of their curves. The first general shape included the least complex patterns assessed as well as the isolated pattern that included real words (i.e., CVC, CVCC, and Contractions). The second general shape included the more complex patterns assessed (CVCe, R-controlled, Blends, Digraphs, Vowel Teams, CVCCVC, Prefix, and Short Vowel Suffix) as well as the Total Tests, and the third group included only the Long Vowel Suffix pattern.

The first general quantile regression curve shape was observed in the least complex patterns assessed (i.e., CVC, CVCC) as well as the single pattern that included real words (i.e., Contractions). This group had quantile regression curves that fluctuated from low to midrange in height, peaked at the lower end of the distribution, and gradually declined through the remainder of the distribution (see Figure 1). An examination of correlations between student decoding and spelling performance on these less complex skills indicated that correlations for students in the lower half of the distribution, except for the extreme low end of the distribution, exceeded the mean correlation reported for all students on these less complex patterns.
Assessment for Effective Intervention 36(1)
The second general quantile regression curve shape was observed in the more complex patterns assessed (CVCe, R-controlled, Blends, Digraphs, Vowel Teams, CVCCVC, Prefix, Short Vowel Suffix, and Total Test). This group had quantile regression curves that fluctuated from low to midrange in height, peaked in the middle of the distribution, and had low levels on both ends of the distribution (see Figure 1). An examination of the correlations between student decoding and spelling performance on these more complex patterns indicated that correlations for students in the lower half of the distribution, with the exception of those in the extreme low end of the distribution, exceeded the mean correlation reported for all students (ranging from $r = .372$ for R-controlled to $r = .790$ for the Total Test), whereas the correlations for students in the upper half of the distribution fell below the mean correlation.

The third general quantile regression curve shape was observed in the Long Vowel Suffix grapho-phonemic pattern (see Figure 1). The shape of this curve began low,
ascended gradually, peaked in the upper half of the distribution, and then descended slightly at the upper end of the distribution. An examination of the correlations between student decoding and spelling performance on this pattern indicated that correlations for students in the upper half of the distribution exceeded the mean correlation reported for all students (r = .534), whereas correlations for students in the lower half of the distribution fell below the mean correlation.

Discussion
This study adds to past research by investigating how closely decoding and spelling relate at the level of specific grapho-phonemic patterns as well as this relation across students with differing reading performance levels. The findings of this study indicate that decoding and spelling have a moderate to strong relation not only at a general level, as already indicated by prior research (Juel et al., 1986; Morris & Perney, 1984; Nagy et al., 2003), but also at a grapho-phonemic level. Overall, more complex grapho-phonemic patterns (e.g., Vowel Teams, Prefix, Short Vowel Suffix) nearly always had higher correlations than those of less complex patterns (e.g., CVC, CVCC, CVCe). In addition, the findings of this study indicate that students whose scores were represented in the tails of the distribution demonstrated weaker relations between decoding and spelling at the grapho-phonemic level for nearly all patterns than students whose scores were represented in the middle of the distribution.
Table 3. Decoding-Spelling Correlations Among Specific Grapho-Phonemic Patterns

<table>
<thead>
<tr>
<th></th>
<th>CVC-D</th>
<th>CVCC-D</th>
<th>CVCe-D</th>
<th>Rcont-D</th>
<th>Blnds-D</th>
<th>Digraph-D</th>
<th>Volteam-D</th>
<th>CVCCVC-D</th>
<th>Prefix-D</th>
<th>Shtvolsufx-D</th>
<th>Lgvolsufx-D</th>
<th>Contract-D</th>
<th>Total-D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CVC-S</strong></td>
<td>.282**</td>
<td>.235**</td>
<td>.214**</td>
<td>.282**</td>
<td>.318**</td>
<td>.266**</td>
<td>.350**</td>
<td>.318**</td>
<td>.294**</td>
<td>.230**</td>
<td>.013</td>
<td>.314**</td>
<td></td>
</tr>
<tr>
<td><strong>CVCC-S</strong></td>
<td>.415**</td>
<td>.377**</td>
<td>.321**</td>
<td>.433**</td>
<td>.407**</td>
<td>.390**</td>
<td>.404**</td>
<td>.396**</td>
<td>.365**</td>
<td>.407**</td>
<td>.300**</td>
<td>.121</td>
<td>.438**</td>
</tr>
<tr>
<td><strong>CVCe-S</strong></td>
<td>.480**</td>
<td>.451**</td>
<td>.490**</td>
<td>.479**</td>
<td>.506**</td>
<td>.449**</td>
<td>.544**</td>
<td>.501**</td>
<td>.525**</td>
<td>.512**</td>
<td>.392**</td>
<td>.047</td>
<td>.549**</td>
</tr>
<tr>
<td><strong>Rcont-S</strong></td>
<td>.337**</td>
<td>.358**</td>
<td>.452**</td>
<td>.370**</td>
<td>.341**</td>
<td>.330**</td>
<td>.443**</td>
<td>.339**</td>
<td>.393**</td>
<td>.379**</td>
<td>-0.021</td>
<td>.424**</td>
<td></td>
</tr>
<tr>
<td><strong>Blnds-S</strong></td>
<td>.562**</td>
<td>.472**</td>
<td>.486**</td>
<td>.572**</td>
<td>.564**</td>
<td>.515**</td>
<td>.527**</td>
<td>.542**</td>
<td>.497**</td>
<td>.514**</td>
<td>.412**</td>
<td>.118</td>
<td>.582**</td>
</tr>
<tr>
<td><strong>Digraph-S</strong></td>
<td>.575**</td>
<td>.545**</td>
<td>.560**</td>
<td>.584**</td>
<td>.614**</td>
<td>.586**</td>
<td>.616**</td>
<td>.594**</td>
<td>.572**</td>
<td>.604**</td>
<td>.431**</td>
<td>.119</td>
<td>.649**</td>
</tr>
<tr>
<td><strong>Volteam-S</strong></td>
<td>.633**</td>
<td>.635**</td>
<td>.593**</td>
<td>.656**</td>
<td>.680**</td>
<td>0.674**</td>
<td>.747**</td>
<td>.646**</td>
<td>.670**</td>
<td>.712**</td>
<td>.564**</td>
<td>.139</td>
<td>.748**</td>
</tr>
<tr>
<td><strong>CVCCVC-S</strong></td>
<td>.523**</td>
<td>.514**</td>
<td>.469**</td>
<td>.502**</td>
<td>.538**</td>
<td>0.487**</td>
<td>.543**</td>
<td>.534**</td>
<td>.494**</td>
<td>.514**</td>
<td>.421**</td>
<td>.038</td>
<td>.569**</td>
</tr>
<tr>
<td><strong>Prefix-S</strong></td>
<td>.651**</td>
<td>.582**</td>
<td>.517**</td>
<td>.655**</td>
<td>.681**</td>
<td>0.637**</td>
<td>.681**</td>
<td>.632**</td>
<td>.624**</td>
<td>.697**</td>
<td>.519**</td>
<td>.0187**</td>
<td>.715**</td>
</tr>
<tr>
<td><strong>Shtvolsufx-S</strong></td>
<td>.646**</td>
<td>.622**</td>
<td>.546**</td>
<td>.672**</td>
<td>.673**</td>
<td>0.660**</td>
<td>.736**</td>
<td>.649**</td>
<td>.708**</td>
<td>.736**</td>
<td>.585**</td>
<td>.181**</td>
<td>.751**</td>
</tr>
<tr>
<td><strong>Lgvolsufx-S</strong></td>
<td>.605**</td>
<td>.604**</td>
<td>.558**</td>
<td>.663**</td>
<td>.649**</td>
<td>0.627**</td>
<td>.689**</td>
<td>.620**</td>
<td>.645**</td>
<td>.696**</td>
<td>.534**</td>
<td>.145**</td>
<td>.715**</td>
</tr>
<tr>
<td><strong>Contract-S</strong></td>
<td>.512**</td>
<td>.526**</td>
<td>.449**</td>
<td>.561**</td>
<td>.545**</td>
<td>.519**</td>
<td>.567**</td>
<td>.511**</td>
<td>.579**</td>
<td>.573**</td>
<td>.459**</td>
<td>.173**</td>
<td>.606**</td>
</tr>
<tr>
<td><strong>Total-S</strong></td>
<td>.696**</td>
<td>.671**</td>
<td>.612**</td>
<td>.714**</td>
<td>.730**</td>
<td>.693**</td>
<td>.765**</td>
<td>.759**</td>
<td>.759**</td>
<td>.788**</td>
<td>.162**</td>
<td>.793**</td>
<td></td>
</tr>
</tbody>
</table>

D = decoding; S = spelling; CVC = consonant–vowel–consonant; CVCC = consonant–vowel–consonant–consonant; CVCe = consonant–vowel–consonant–silent e; Rcont = R-controlled; Blnds = blends; Digraph = digraphs; Volteam = vowel teams; CVCCVC = two closed syllable; Prefix = prefixes; Shtvolsufx = short vowel suffix; Lgvolsufx = long vowel suffix; Contract = contraction.

Bold values indicate correlations between student decoding and spelling performance on identical grapho-phonemic patterns.

*p < .05, **p < .01
Assessment Factors and Weaker Relations

One factor related to assessment features that may help explain weaker relations in student decoding and spelling performance is a small number of test items. A small number of items can potentially decrease correlations (Fitz-Gibbon & Morris, 1987). This happens in part because of a restriction of range in test items that can manifest as a ceiling effect. The ceiling effect may offer explanation for the weaker relations between decoding and spelling performance on the less complex grapho-phonemic patterns. Evidence of this can be seen in that the grapho-phonemic patterns with the lowest number of items (i.e., 35 decoding and 4–7 spelling items) had more skewed distributions and weaker relations. Furthermore, the items on the decoding and spelling assessments were designed to target very discrete skills. In turn, some of the items representing the less complex patterns were kept short (i.e., all items were restricted to one syllable). The limited range of difference between items may have resulted in a ceiling effect. Evidence for this was found in an examination of the modes of the distributions. The mode was at or near the maximum score possible in the less complex pattern distributions demonstrating most students decoded and spelled nearly all items for these patterns accurately. In contrast, the mode for the less common patterns was closer to the middle of the distribution, demonstrating a more balanced number of students who decoded and spelled the items correctly or incorrectly. This assessment factor may have resulted in weaker relations for the less complex grapho-phonemic patterns.

Another factor related to the assessments that may help explain weaker correlations between student decoding and spelling performance is the frequency of the grapho-phonemic patterns being tested. The items on both the decoding and spelling assessments for the less complex patterns were high frequency patterns such as CVC, CVCC, and R-controlled. Based on the frequency and exposure to these less complex patterns, students likely had more practice with them compared to more complex patterns. This may have resulted in highly accurate decoding and spelling of the less complex patterns which resulted in a ceiling effect and weaker relations. The frequency of these less complex patterns is supported by a grapho-phonemic analysis of the English language that indicated the short vowel closed syllable pattern, which includes CVC and CVCC patterns, occurs in 15,921 words out of a 17,310-word vocabulary (Fry, 2004). This same analysis reported that 2,791 words out of a 17,310-word vocabulary had an R-controlled pattern (Fry, 2004). Furthermore, research indicates that words with high frequency grapho-phonemic sequences are decoded and spelled more accurately than words with less frequent grapho-phonemic sequences (Apel et al., 2006). Although weaker relations were found between student decoding and spelling performance on less complex patterns, the explanations offered here indicate that a coordinated decoding and spelling assessment is still likely to be useful for identifying strengths and weaknesses in student grapho-phonemic knowledge. Although factors associated with the assessments themselves offer some explanation for weaker relations in specific grapho-phonemic patterns, factors associated with student performance levels should also be explored.

Student Factors and Weaker Relations

One factor related to student performance levels that may offer some explanation for the weaker relations between student performance on less complex grapho-phonemic patterns is consistency of grapho-phonemic knowledge application. There was evidence of less consistent grapho-phonemic knowledge application in decoding and spelling of less complex grapho-phonemic patterns (i.e., CVC, CVCC, and R-controlled) across all student performance levels, meaning some students were able to spell but not decode, or decode but not spell, some grapho-phonemic patterns accurately. This difference indicates that students may have applied their grapho-phonemic knowledge for these patterns in an inconsistent manner when decoding and spelling identical grapho-phonemic patterns.

Research on developmental phase theory supports this explanation of inconsistent grapho-phonemic knowledge (Ehri, 1986; Henderson, 1981). Students in the initial phases of decoding and spelling development often inconsistently apply grapho-phonemic knowledge (Ehri, 1986, 2000; Henderson, 1981). For example, students with beginning skills may at times decode or spell a word by matching every phoneme with a corresponding grapheme, whereas at other times students with beginning skills may match only initial and final phonemes with corresponding graphemes. Moreover, decoding and spelling development may not be synchronized (Ehri, 1986; Frith, 1985). This may provide further explanation for inconsistent application of grapho-phonemic knowledge in decoding compared to spelling tasks. For example, at certain points in development, students may have the ability to recognize grapho-phonemic patterns but not yet produce them in their spelling (Frith, 1985).

In addition to consistency of grapho-phonemic knowledge application, the restricted range in student performance levels within the sample of this study may further explain weaker relations between student performance on the less complex patterns on both decoding and spelling tasks. A restricted range in student performance levels (i.e., the majority of students decoded and spelled the items with less complex patterns correctly, resulting in a lack of differentiation within the scores), specifically at the highest and
lowest levels of performance, may explain the weak relations in the performance of these students. Restriction in range has been shown to contribute to skewed distributions and weaker relations (Fitz-Gibbon & Morris, 1987). However, data from this study are not sufficient to determine if the weaker relations were because of the previously described restricted range of items in the less complex patterns or if they were because of a restricted range of student performance (i.e., many students decoded or spelled the majority of the items correctly), or both. Although weaker relations were found between decoding and spelling performance in lower and higher performing students, the explanations offered here indicate that a coordinated decoding and spelling assessment is still likely to be useful for identifying strengths and weaknesses in student grapho-phonemic knowledge. Although these factors associated with the assessments themselves and factors associated with student performance levels offer explanation for the weaker relations in specific grapho-phonemic patterns, other factors associated with the assessments and student performance levels offer explanation for the stronger relations found in this study.

**Assessment Factors and Stronger Relations**

Multiple factors related to the assessment features as well as those related to the students are addressed. These factors have joint implications and are addressed following the description of the explanations for each factor. One factor related to assessment features is that the spelling assessment was designed to test the same grapho-phonemic patterns as the decoding assessment. This factor provides some explanation for the moderate to strong relation between decoding and spelling performance on most grapho-phonemic patterns. Because of the similarities in the patterns being assessed, one would expect to find a strong relation between the grapho-phonemic patterns. The moderate to strong relations between the decoding and spelling skills indicate the decoding and spelling assessments for these patterns tapped the same grapho-phonemic knowledge base. This is supported by research that has demonstrated the application of grapho-phonemic knowledge by early readers and spellers in both decoding and spelling tasks (Treiman, 1984; Waters et al., 1985; Waters, Seidenberg, & Bruck, 1984). In addition, students with proficient grapho-phonemic knowledge are more likely to be proficient at decoding and spelling (Waters et al., 1985).

Another factor related to the assessments was that the grapho-phonemic patterns with the most items had the strongest relation between decoding and spelling performance. Just as a small number of items can potentially decrease correlations, as discussed above, a large number of items can potentially increase the correlations (Fitz-Gibbon & Morris, 1987). The larger number of items may have provided a wider range of difficulty among the items resulting in a broader distribution of scores and higher correlations (Fitz-Gibbon & Morris, 1987). The more complex grapho-phonemic patterns had twice as many decoding and spelling items (i.e., 70–80 decoding items and 16–17 spelling items) as the less complex grapho-phonemic patterns (i.e., 35–36 decoding items and 4–7 spelling items) and also had the strongest relation between decoding and spelling performance. Furthermore, 10 out of 12 decoding variables had the strongest relation with the Vowel Teams, Prefix, and Short Vowel Suffix spelling variables. Again, these spelling patterns had nearly twice the number of items presented (i.e., 16–17) compared to the number of spelling items with less complex patterns (i.e., 4–7). Although factors associated with the assessments themselves offer some explanation for stronger relations in specific grapho-phonemic patterns, factors associated with student performance levels should also be explored.

**Student Factors and Stronger Relations**

Although less consistent grapho-phonemic knowledge application can result in weaker relations between decoding and spelling performance, as previously described, more consistent grapho-phonemic knowledge application can result in stronger relations. Consistency of grapho-phonemic knowledge application may offer some explanation for the stronger relations between student performance on more complex grapho-phonemic patterns. Grapho-phonemic knowledge can be demonstrated by accurate decoding or spelling of words with specific grapho-phonemic patterns (e.g., CVC, CVCe). There was evidence of consistent grapho-phonemic knowledge application (i.e., stronger relations between decoding and spelling performance) on more complex patterns (i.e., CVCCVC, Prefix, and Short Vowel Suffix) across all student performance levels. These findings indicate that students who had difficulty decoding more complex grapho-phonemic patterns also had difficulty spelling these same patterns. These findings are supported by research demonstrating that students who were able to apply grapho-phonemic correspondence knowledge to decode pseudowords also demonstrated the ability to apply knowledge of grapho-phonemic correspondences in their spelling, as evidenced by phonetic spelling errors (i.e., errors that maintain grapho-phonemic correspondences; Gough, Juel, & Griffith, 1992). In addition, students with more developed decoding and spelling skills were more consistent with their application of grapho-phonemic knowledge (Ehri, 1986; Henderson, 1981). For example, students in the higher phases of decoding and spelling development consistently match individual phonemes with correct grapheme sequences.
**Limitations and Directions for Future Research**

Although the findings of this study lend support to a wide range of past research on decoding and spelling, there are limitations that should be considered as well as related directions for important future research. This study included an investigation of the relation between decoding and spelling in early elementary students with differing reading performance levels. Additional data from students with more emergent reading skills (i.e., first-grade students) would enable a better understanding of the relation between decoding and spelling across a broader range of performance levels.

Another limitation is that the spelling assessment used in this study included 5 to 17 items per pattern, potentially resulting in a ceiling effect as well as affecting the internal consistency correlation coefficients for some of the grapho-phonemic patterns (e.g., R-controlled, CVCCCV, and CVCC, which had correlations ranging from .487 to .686). Furthermore, the ceiling effect could have resulted from the fact that the more simplistic patterns were already mastered by the majority of students. Increasing the number of spelling items may increase the range of difficulty across items within a pattern, decrease the ceiling effect, and improve internal consistency of specific grapho-phonemic patterns. A second potential method of decreasing the ceiling effect is to increase the complexity of the spelling items by including multisyllabic words that have several grapho-phonemic patterns embedded within the word. Increasing the range of difficulty of spelling assessment items may provide further understanding of the strength of the relations between student decoding and spelling performance on the level of grapho-phonemic patterns, particularly for the less complex patterns, which had the fewest number of items (e.g., CVC, CVCC, CVCe). Furthermore, spelling performance on grapho-phonemic patterns embedded in longer words could be compared with spelling performance on grapho-phonemic patterns in shorter words, as demonstrated in this study. For example, comparing spelling accuracy of the OU pattern in a multisyllabic word (e.g., shouted) to spelling accuracy of the OU patterns in a single syllable word (e.g., pout). Determining this relation might provide data for informing the design of future spelling assessments and improving the efficiency of grapho-phonemic knowledge assessment.

**Summary and Implications for Practice**

This study examined the relation between decoding and spelling on the specific grapho-phonemic level in second- through fourth-grade students. Overall, a moderate to strong correlation between student performance in decoding and spelling tasks was found. Stronger relations were found between more complex grapho-phonemic patterns, whereas weaker relations were found between less complex grapho-phonemic patterns. Furthermore, decoding and spelling of students with lower to middle ability levels was more strongly related than for students with low and high ability levels. This is important as it has the potential to affect how researchers and teachers assess student grapho-phonemic knowledge and ultimately how teachers approach decoding and spelling instruction. One implication of this study is an improvement in the utility of data derived from decoding and spelling assessments. Prior research indicates that researchers and teachers typically use decoding assessments that assess few grapho-phonemic patterns and provide few opportunities for student response to a single pattern (National Institute of Child Health and Human Development [NICHHD], 2000). As a result, teachers and researchers are left with an incomplete assessment of student grapho-phonemic knowledge. Furthermore, grapho-phonemic patterns are rarely isolated; rather, they are typically embedded within longer, more complex words. The results of this study also indicate that it is possible to accurately assess a much wider range of grapho-phonemic patterns. Teachers could then use these data to develop and implement specific interventions targeting identified gaps in student grapho-phonemic knowledge rather than spending instructional time on patterns that students have already mastered.

The results of this study also have implications for improving grapho-phonemic assessment efficiency by linking decoding and spelling assessments. Prior research on reading assessment shows that researchers and teachers are using time-consuming assessments to identify general areas of weakness in reading skill (NICHHD, 2000). Typically, decoding assessments require one-on-one testing as in the Word Attack subtest of the WJIII (Woodcock et al., 2001), the NWF subtest of the DIBELS measure (Good & Kaminski, 2002), and the decoding assessment used in this study. This is a time-consuming process and often reduces the amount of time available for instruction. As the results of this study indicate, student performances on specific grapho-phonemic patterns in decoding and spelling tasks are, in general, moderately to strongly related. This indicates that a spelling assessment that is coordinated with a more in-depth decoding assessment could be used as a screening instrument. Using a group-administered spelling test as a screening instrument could greatly reduce the number of students as well as the number of grapho-phonemic patterns needing to be assessed with an individual decoding assessment. Teachers could use the spelling assessment to determine a range of grapho-phonemic patterns needing further assessment with a more in-depth decoding assessment rather than having to assess every student on all patterns. If decoding and spelling assessment were linked, a great deal of time that is currently used for assessment could be used for instruction.
A final implication of this study involves linking decoding and spelling instruction. Typical instructional practices often involve separating decoding and spelling instruction (Gentry, 2006). Combining these efforts could maximize student learning and improve reading and spelling outcomes by increasing the redundancy of exposure to specific grapho-phonemic patterns. This redundancy is necessary if students are to become able to automatically decode and spell (Ehri, 1980, 1997). Moreover, it has been shown that decoding strategy instruction can improve spelling (Ehri & Wilce, 1987) and spelling strategy instruction can improve decoding (Berninger, Abbott, Thomson, & Raskind, 2001; Graham, Harris, & Chorzempa, 2002; Santoro, Coyne, & Simmons, 2006). If decoding and spelling instruction were linked and based on explicit discussions of orthographic influences (i.e., grapho-phonemic correspondences, etymology, etc.), students would have the opportunity to recognize the systematic and predictable nature of the orthography. These opportunities would support specific grapho-phonemic learning that could readily enhance both decoding and spelling skills.

In conclusion, the results of this study indicate that a spelling assessment with items representing a wide range of grapho-phonemic patterns may be used to screen for specific areas of grapho-phonemic knowledge weakness and provide educators with an efficient method of collecting the data needed to develop targeted interventions, resulting in improved student grapho-phonemic skill.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Financial Disclosure/Funding

The authors received no financial support for the research and/or authorship of this article.

References


About the Authors

Kelly P. Robbins, PhD, is an adjunct professor of special education at the University of Utah and Westminster College. Her current interests include spelling, reading, and learning disabilities.

John L. Hosp, PhD, is an associate professor of teaching and learning at the University of Iowa. His current research interests include curriculum-based measurement and curriculum-based evaluation, particularly in the elementary grades.

Michelle K. Hosp, PhD, is a research associate in the Department of Teaching and Learning at the University of Iowa. Her interests are curriculum-based measurement and curriculum-based evaluation for reading and literacy with elementary students.

Lindsay J. Flynn, PhD, is a postdoctoral research scholar in the Graduate School of Education at the University of California, Riverside. Her current interests include assessment and intervention of reading difficulties, responsiveness to intervention, and positive behavior support.