Journal of School Psychology 52 (2014) 49–61

Contents lists available at ScienceDirect

Journal of School Psychology

journal homepage: www.elsevier.com/locate/jschpsyc

Assessing spelling in kindergarten: Further comparison of scoring metrics and their relation to reading skills ☆

Nathan H. Clemens [a,*], Eric L. Oslund [b], Leslie E. Simmons [a], Deborah Simmons [a]

[a] Department of Educational Psychology, Texas A&M University, USA
[b] The Meadows Center for Preventing Educational Risk, The University of Texas at Austin, USA

ARTICLE INFO

Article history:
Received 9 January 2013
Received in revised form 21 December 2013
Accepted 23 December 2013

Keywords:
Early literacy
Spelling
Reading
Assessment

ABSTRACT

Early reading and spelling development share foundational skills, yet spelling assessment is underutilized in evaluating early reading. This study extended research comparing the degree to which methods for scoring spelling skills at the end of kindergarten were associated with reading skills measured at the same time as well as at the end of first grade. Five strategies for scoring spelling responses were compared: totaling the number of words spelled correctly, totaling the number of correct letter sounds, totaling the number of correct letter sequences, using a rubric for scoring invented spellings, and calculating the Spelling Sensitivity Score (Masterson & Apel, 2010b). Students ($N = 287$) who were identified at kindergarten entry as at risk for reading difficulty and who had received supplemental reading intervention were administered a standardized spelling assessment in the spring of kindergarten, and measures of phonological awareness, decoding, word recognition, and reading fluency were administered concurrently and at the end of first grade. The five spelling scoring metrics were similar in their strong relations with factors summarizing reading subskills (phonological awareness, decoding, and word reading) on a concurrent basis. Furthermore, when predicting first-grade reading skills based on spring-of-kindergarten performance, spelling scores from all five metrics explained unique variance over the autoregressive effects of kindergarten word identification. The practical advantages of using a brief spelling assessment for early reading evaluation and the relative tradeoffs of each scoring metric are discussed.

1. Introduction

Although the interrelations between spelling skills and reading development have been acknowledged for quite some time (e.g., Hoillingworth, 1923), educators and practitioners often consider spelling and reading skills as distinct (Cooke, Slee, & Young, 2008; Foorman, Schatschneider, Eakin, Fletcher, Moats & Francis, 2006; Johnston, 2001). Spelling assessments are often considered separately from reading assessments and are typically underutilized within reading evaluations. Norm-referenced reading assessments may omit spelling assessment, seldom extend downward to kindergarten, or may not be scored in ways that are sensitive to the approximate or invented spellings typical of early learners.

There is good reason to attend to the spelling skills of beginning readers (e.g., kindergarten). Spelling and reading rely on a foundation of common skills, and these processes may develop concurrently and reinforce each other (Adams, 1990; Caravolas,

☆ This research was supported in part by Grant Number R324E06067, National Center for Special Education Research, from the Institute of Education Sciences (IES), U.S. Department of Education. The content and positions in this article do not necessarily represent those of the funding agency. We thank Brenda Gamez, Joslynn Benitez, Brittany Sanchez, and Terez Turner for the assistance with spelling scoring, and Kirsten McBride and Kimberly Williams for the editorial assistance.
* Corresponding author at: Texas A&M University, Department of Educational Psychology, 4225 TAMU, College Station, TX 77843-4225, USA.
E-mail address: nclemens@tamu.edu (N.H. Clemens).
ACTION EDITOR: Patti Manz.

0022-4405/$ – see front matter © 2014 Society for the Study of School Psychology. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jsp.2013.12.005

Hulme, & Snowling, 2001; Ehri, 1997, 1998, 2000; Hecht & Close, 2002; Shankweiler & Lundquist, 1993; Treiman, Berch, Tincoff, & Weatherston, 1993). A synergistic relation between these skills is evidenced by gains that are observed when instruction integrates spelling within reading instruction (Santoro, Coyne, & Simmons, 2006; Treiman, 1998; Uhry & Shepherd, 1993; Weiser & Mathes, 2011). The assessment of beginning spelling skills can serve as a window into students' early reading development (Ouellette & Sénéchal, 2008b; Treiman, 1998) and may provide information on students' knowledge and ability to apply foundational reading skills in ways that conventional reading assessments may not (Apel, Thomas-Tate, Wilson-Fowler, & Brimo, 2012). As part of a more comprehensive approach to reading assessment, spelling may help identify specific skill deficits and inform intervention strategies for struggling students (Al Otaiba & Hosp, 2010; Masterson & Apel, 2010a; Moats, 1993; Robbins, Hosp, Hosp, & Flynn, 2010).

Spelling assessment with students at early stages of reading development can provide important information on their knowledge of and ability to apply phonological awareness and alphabetic knowledge. Phonological awareness is critical for early spelling, as beginning spellers must retain an oral representation of a word in memory and then represent phonological units with letters, thus implicating skills in phoneme isolation and phonemic segmentation (Bryant, MacLean, Bradley, & Crossland, 1990; Caravolas et al., 2001; Goswami & Bryant, 1990; Muter, Hulme, Snowling, & Taylor, 1997; Nation & Hulme, 1997; Treiman et al., 1993). Alphabetic knowledge is also closely related to early spelling, and research has established the importance of an understanding of knowledge of letter sounds and patterns on beginning spelling (Caravolas et al., 2001; Ehri, 1997, 1998; Hecht & Close, 2002; Muter et al., 1997; National Early Literacy Panel (NELP), 2008; Ouellette & Sénéchal, 2008b; Puranik, Lonigan, & Kim, 2011; Ritchey, 2008; Ritchey & Speece, 2006). Students' spelling of words, even when not spelled correctly (e.g., spelling "was" as [wuz]), can provide insight on their knowledge of letter sounds (Ahmed & Lombardino, 2000; Ouellette & Sénéchal, 2008a, 2008b). Thus, the spelling responses of children in kindergarten and first grade can provide important information on their knowledge of and ability to apply phonemic awareness and alphabetic knowledge to print-related tasks.

Spelling assessment may also provide insight on word reading skills. Spelling skills are associated with decoding and pseudoword reading (Furnes & Samuelsson, 2009; Lombardino et al., 1999; NELP, 2008; Ritchey, Coker, & McCraw, 2010; Robbins et al., 2010), and spelling assessment can function as an index of students' knowledge of grapho-phonemic patterns, especially when those patterns are more complex (e.g., CVCe, r-controlled, blends, and CVCCVC; Robbins et al., 2010). Spelling skills have also demonstrated moderate to strong relations with skills in reading real words and text, both concurrently and as predictors of these skills in the future (Caravolas et al., 2001; Furnes & Samuelsson, 2009; Lombardino et al., 1999; Morris, Bloodgood, & Perney, 2003; Muter et al., 1997; Nation & Hulme, 1997; NELP, 2008; Ritchey, 2008; Ritchey et al., 2010; Stage & Wagner, 1992).

The assessment of spelling skills may also provide insight into more sophisticated skills related to word reading. When words do not completely conform to phonological regularity, knowledge of word spellings or reoccurring letter patterns (i.e., orthographic knowledge) is required to be able spell the word correctly. For example, knowledge of the correct sequence of letters is needed in order to spell "was" correctly, as opposed to relying solely on phonological and alphabetic information (which may yield [wuz]). Research supports the relation between spelling skills and orthographic knowledge (Lombardino et al., 1999; Ouellette, 2010; Ouellette & Sénéchal, 2008a, 2008b; Stage & Wagner, 1992) and that kindergarten students make use of basic knowledge of orthographic patterns when reading or spelling words (Wright & Ehri, 2007). Similarly, morphological awareness, which is implicated in word recognition and vocabulary knowledge (see Carlisle, 2003), is involved in spelling as well (Sénéchal & Kearnan, 2007; Treiman & Bourassa, 2000; Treiman, Cassar, & Zukowski, 1994).

In short, learning to spell is intertwined with learning to read. In early stages, both involve a strategic integration of skills in phonemic awareness and alphabetic knowledge. As spelling skills become more sophisticated, orthographic knowledge and morphological awareness are utilized to spell words correctly. Assessing spelling skills with beginning readers, therefore, may provide important information about their level of development and their ability to flexibly apply important skills to print-related tasks.

1.1. Special considerations for assessing spelling with early learners

A lack of attention to the spelling skills of early readers may be due to several reasons, including a lack of awareness of the close relations between learning to read and learning to spell, the idea that young readers should not be expected to spell until later grades, and perhaps most likely, a reliance solely on conventional scoring of spelling responses (i.e., words are only scored as correct or incorrect). The latter point underscores a challenge to assessing spelling at early grade levels: Beginning spellers seldom spell words correctly, and conventional scoring usually results in scores of zero. Understanding the utility of spelling assessment with beginning spellers requires scoring methods that are sensitive to students' approximate or invented spellings.

Several metrics for scoring beginning spelling have been developed. Some award points based on the number of sounds represented by correct letters or letter combinations (i.e., letter sounds correct or "sound spelling") that are phonologically consistent (i.e., writing "f" for [ph]; Lombardino, Bedford, Fortier, Carter, & Brandi, 1997; Richgels, 1995; Ritchey et al., 2010; Stage & Wagner, 1992). Other metrics award credit for pairs of letters that are correctly sequenced, such as the correct letter sequences metric that is recommended within curriculum-based measurement (Hosp, Hosp, & Howell, 2007). More sophisticated metrics consist of rubrics or rating scales, in which each word is scored based on predetermined rules according to spelling accuracy or how successfully students represent letters with phonologically consistent spellings (De Graff & Torgesen, 2005; Hecht & Close, 2002; O'Connor & Jenkins, 1995; Rao, Prakash, & Joshi, 2006; Ritchey et al., 2010; Tangel & Blachman, 1992). An example of a rating scale approach is the Spelling Sensitivity Score (Masterson & Apel, 2010b), which awards points based on a student's accuracy in spelling elements (i.e., phonemes) of a word.

Limited research has compared the validity of various spelling scoring metrics. Comparing conventional scoring, letter sounds, letter sequences, and rubric scoring with a sample of general education kindergarten students, Ritchey et al. (2010) found that the four approaches were highly correlated ($rs = .84$ to .98). Further, the scoring metrics differed little in the degree to which they

were associated with phonological awareness, letter naming fluency, word reading, and writing skills. These findings have been instructive in establishing the relative equivalence of spelling scoring strategies; however, additional work is needed regarding their relations to reading subskills and their potential application to students with low reading and spelling skills.

Insight provided by spelling assessment on students' early literacy skills may be particularly useful for the assessment of struggling or at-risk readers. Spelling assessment could provide a format in which to observe students' knowledge and proficiency in skills such as segmenting a word by its component sounds, coding sounds with letters, or recalling letter combinations or word spellings in ways that add meaningfully to other assessments of these skills. Robbins et al. (2010) demonstrated that correlations between spelling performance and decoding skills tended to be stronger for students with reading achievement in the lower half of the distribution, and research has also revealed the benefits of integrating spelling with reading intervention for students struggling with reading (Weiser & Mathes, 2011). However, more research is needed, particularly in comparing metrics for scoring beginning spelling, to help better elucidate their potential for informing reading assessment with a population more likely to be involved in reading evaluations.

1.2. Study purpose

The purpose of this study was to extend current knowledge of kindergarten spelling assessment by further exploring the relations between spelling scoring metrics and early reading skills. We also sought to extend recent findings (e.g., Ritchey et al., 2010) by using a latent variable approach and an expanded set of reading skills measures, adding an additional metric (Spelling Sensitivity Score), investigating the predictive validity of the spelling metrics, and examining relations among a sample of students who received intervention to target reading difficulties. We investigated the following research questions: (1) What differences are observed between kindergarten spelling scoring metrics (Word Spelling, Sound Spelling, Letter Sequences, Rubric scoring, and the Spelling Sensitivity Score) in their relations to reading subskills (phonological awareness, pseudoword decoding, and word reading) on a concurrent basis in kindergarten? (2) What differences are observed across the kindergarten spelling metrics in predicting reading skills one year later?

2. Method

2.1. Participants and study context

Analyses utilized data from a sample of monolingual English-speaking students ($N = 287$) who participated in a longitudinal investigation of a kindergarten reading intervention (see Coyne et al., 2013; Simmons et al., 2011). The mean age of the sample at the end of kindergarten (when the data used in this study were collected) was approximately 6:0. The sample was 49.8% girls as well as 60.3% White, 18.5% Hispanic, and 18.1% Black; 12.9% of students were eligible for special education services. Students were recruited from 22 schools in Texas, Connecticut, and Florida. The majority (86%) of the schools received Title 1 funding. The percentage of students school-wide who qualified for free or reduced-price lunch ranged from 69% to 81% in Texas, 50% to 81% in Connecticut, and 32% to 80% in Florida.

All students were participating in a longitudinal investigation of kindergarten intervention for at-risk readers (see Coyne et al., 2013; and Simmons et al., 2011 for a full description). Data utilized in the current study were collected after students had completed the intervention (spring of kindergarten, and spring of first grade). All students had been previously identified as being at risk for reading difficulty during the third week of kindergarten (using a combination of teacher nomination and scores below the 33rd percentile on measures of alphabetic knowledge and phonological awareness).

The intervention study contrasted the Early Reading Intervention (ERI; Pearson/Scott Foresman, 2004) with the supplemental instruction implemented by the schools (i.e., typical practice intervention). In both conditions, groups of three to five students met for 30 min 5 days per week across the kindergarten school year. The ERI program consists of explicit instruction targeting letter names and sounds, phonemic awareness, decoding, word recognition, and reading short sentences. Writing and spelling are integrated in each lesson. Interventionists in the typical practice condition provided school-designed beginning reading intervention, and the content and instructional approaches were allowed to vary naturally. Across this condition lessons targeted letter and sounds, decoding skills, phonological awareness, writing, spelling, vocabulary, and comprehension activities.

Following the intervention, at least 40% of the sample scored within the "some risk" or "at-risk" range on the Nonsense Word Fluency and Phoneme Segmentation Fluency measures from the Dynamic Indicators of Basic Early Literacy Skills (6th Edition; Good & Kaminski, 2007). Although over half of the sample now scored out of standard risk ranges, the sample was still of interest because all students had been found to be at-risk for reading difficulty at kindergarten entry (a risk factor for later reading difficulty; NELP, 2008; Santi, York, Foorman, & Francis, 2009) and received intervention through the research study across the kindergarten year. Additionally, evidence suggests that students initially identified as at-risk earlier in kindergarten but make considerable growth across kindergarten may still be at risk for reading difficulties in later grades (Al Otaiba et al., 2011).

No statistically significant differences on any of the five spelling scoring metrics at the end of kindergarten were observed between students in the two intervention conditions. Additionally, we investigated group differences in the correlations between the spelling metrics and each of the reading skill variables in kindergarten and first grade. Comparisons of these coefficients using Fisher's z transformation revealed high similarity across the intervention groups, as 84% of the spelling–reading skills relations investigated in this study did not significantly differ between the intervention conditions. Given the similarities in spelling scores and the spelling–reading relations across the intervention groups, analyses were conducted with the sample as a whole.

2.2. Measures

The following measures are organized according to their respective skill domain. Raw scores were used in all analyses. Validity estimates reported here are specific to the present sample.

2.2.1. Spelling assessment

Spelling skills were assessed using the Test of Written Spelling — 4 (TWS-4; Larsen, Hammill, & Moats, 1999), in which students write words orally dictated by the examiner. The TWS-4 demonstrates an alternate-form reliability of .86, and Cronbach's alpha of .87 for 6-year-olds (Larsen et al., 1999). The TWS-4 was administered according to standard procedures, which includes discontinuing the test following five consecutive incorrect spellings. Poorly formed, reversed, and inverted letters were accepted provided that the letter could not be confused with another (e.g., *b*/*d*, *p*/*q*, and *m*/*w*). All students' responses were scored using the following five metrics (see additional description and examples in Appendix A). In our sample, correlations between all the spelling metrics ranged from .94 to .99.

2.2.2. Words spelled correctly (Word Spelling)

Word Spelling scores consisted of the total number of words spelled correctly on the TWS-4, which is consistent with the standard scoring on the TWS-4 and akin to the conventional scoring of a spelling list. Cronbach's alpha for Word Spelling scores with the current sample was .76.

2.2.3. Correct letter sounds (Sound Spelling)

The Sound Spelling metric was scored as described by Ritchey et al. (2010). Within in each word spelled, 1 point was awarded for every letter sound represented by the correct letter or letter that can represent the same sound (i.e., "phonologically legal" letters), such as spelling *k* in place of *c*. This rule included letter combinations if the letters could represent the same phoneme (e.g., spelling *ph* in place of *f*). The total number of correct letter sounds in each word was totaled as the student's Sound Spelling score. Cronbach's alpha with this sample was .93. In prior work with kindergarten students, Sound Spelling scores have demonstrated correlations with measures of word reading and phonological awareness of .73 (Ritchey et al., 2010).

2.2.4. Correct letter sequences (Letter Sequences)

The Letter Sequences metric was scored as described in the curriculum-based measurement literature (see Hosp et al., 2007). For each word, 1 point was awarded for each correctly sequenced pair of letters, including points for correct initial and final letters. Consistent with this strategy, only correct letters were accepted; phonologically legal letters were not accepted. Cronbach's alpha for Letter Sequences scores with this sample was .92. Letter Sequences scores have demonstrated correlations with kindergarten word reading and phonological awareness ranging from .73 to .78 (Ritchey et al., 2010).

Table 1
Descriptive statistics of the measures assessed in kindergarten and first grade.

Grade	Measure	*M*	*SD*	Range	Skewness	Kurtosis
Kindergarten	Spelling					
	Word Spelling	2.33	2.17	0–10	0.60	0.32
	Sound Spelling	16.85	11.81	0–63	0.72	0.99
	Letter Sequences	18.38	13.47	0–73	0.69	0.60
	Spelling Rubric	31.57	20.12	0–93	0.31	−0.26
	Spelling Sensitivity Score	54.14	31.33	0–178	0.58	0.74
	Phonological Awareness					
	Sound Matching	9.77	4.94	0–20	0.16	0.70
	Blending Words	7.44	3.78	0–17	0.21	0.21
	Phoneme Segmentation Fluency	36.63	19.47	0–72	−0.33	−0.86
	Pseudoword Decoding					
	Word Attack	5.06	4.93	0–27	1.07	1.37
	Phonemic Decoding Efficiency	3.39	3.59	0–14	0.87	0.25
	Nonsense Word Fluency	27.09	14.31	0–80	0.42	1.00
	Word Reading					
	Word Identification	10.51	9.14	0–50	0.97	1.01
	Sight Word Efficiency	7.12	6.08	0–36	1.10	2.35
	Oral Reading Fluency	10.76	8.63	0–74	2.37	11.91
First grade						
	Word Identification	34.50	14.50	0–67	−0.58	−0.03
	Word Attack	13.79	8.48	0–38	0.48	−0.40
	Oral Reading Fluency	44.48	27.69	1–134	0.61	−0.26
	Passage Comprehension	15.57	7.93	0–36	0.02	−0.62

Note. $N = 287$. Word Spelling = Number of words spelled correctly; Sound Spelling = number of letter sounds correct in each word spelled; Letter Sequences = number of correct letter sequences in each word spelled; Spelling Rubric = score on the Tangel and Blachman (1992) spelling scoring rubric; Spelling SSS = Spelling Sensitivity Score (total elements correct). Due to an administrative error, Sight Word Efficiency and Phonological Decoding Efficiency were not administered in kindergarten to the Florida subsample; therefore, $n = 151$ of these measures.

2.2.5. Rubric

This method utilized the scoring rubric developed by Tangel and Blachman (1992). Each word was scored on a 6-point scale according to its spelling accuracy or phonological proximity to the correct spelling. Additional description and a scoring example are provided in Appendix A. In short, 6 points were awarded if the word was spelled entirely correctly, 0 points were awarded if the spelling did not include any correct letters or phonologically legal letters, and between 1 and 5 points were awarded depending on the number of correct or phonologically legal letters. The Rubric differs from the other metrics by placing greater emphasis on phonemes and phoneme position (e.g., initial sound), as well as providing a more "global" evaluation of the accuracy and phonological consistency of a student's spelling of a word. Cronbach's alpha for Rubric scores was .93 with the current sample. This scoring rubric has demonstrated correlations with kindergarten word reading and phonological awareness ranging from .74 to .77 (Ritchey et al., 2010).

2.2.6. Spelling Sensitivity Score (SSS)

The SSS (Masterson & Apel, 2010b) was designed to quantify growth in linguistic knowledge and to be sensitive to changes in spelling skills. Under the SSS metric, each word is segmented by "elements" (i.e., phonemes, juncture changes, or affixes), and each element is awarded 0 to 3 points based on its accuracy or phonological plausibility. The total elements score across words was used in the present analyses. As described in Appendix A, SSS differs from the other metrics in that individual points are awarded based on each word element. Cronbach's alpha for SSS scores with this sample was .92. To our knowledge, concurrent or predictive validity of SSS scoring has not been investigated.

2.2.7. Phonological awareness

Phonological awareness was assessed using three measures. The Blending Words subtest from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) was used to assess students' ability to blend individual syllables or phonemes spoken by the examiner into whole words. This test is untimed. The Sound Matching subtest from the CTOPP is an untimed test in which students are instructed to select from three pictures the one that has the same initial beginning or ending sound as a test picture. Cronbach's reliability coefficients for the Blending Words and Sound Matching subtests range from .86 to .93 for children ages 5–7 (Wagner et al., 1999). On the Phoneme Segmentation Fluency measure from the DIBELS (6th edition; Good & Kaminski, 2007), students are instructed to segment a word spoken orally by the examiner, and scores consist of the number of phonological segments produced in 1 min. The one-month alternate-form reliability estimate for kindergarten is .79 (Good & Kaminski). Concurrent intercorrelations between the phonological awareness measures in our sample ranged from .43 to .63.

2.2.8. Pseudoword decoding

Students' skills in decoding non-words were assessed with three subtests. Word Attack from the WRMT-R/NU asks students to read from a list of pseudowords of increasing difficulty. This test is not timed. The split-half reliability estimate for first-grade students is .94 (Woodcock, 1998). The Phonemic Decoding Efficiency subtest of the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) consists of a list of decodable pseudowords of increasing difficulty. Students are instructed to read as many words as possible in 45 s. The alternate-form reliability estimate for 6-year-olds is .97 (Torgesen et al., 1999). The Nonsense Word Fluency measure from the DIBELS 6th edition assesses fluency with identifying letter sound correspondences (in isolation, as part of a sound segment, or as part of a whole word) in reading VC and CVC pseudowords. The one-month alternate-form reliability estimate is .83 (Good & Kaminski, 2007). Concurrent intercorrelations between the decoding measures in our sample ranged from .58 to .72.

2.2.9. Word reading

Skills in reading real words were assessed using three subtests. The Word Identification subtest from the WRMT-R/NU is an untimed test consisting of a list of real words of increasing difficulty. The split-half reliability estimate for first-grade students is .98 (Woodcock, 1998). The Sight Word Efficiency subtest from the TOWRE consists of a list of real words of increasing difficulty; students are scored in terms of the number of words read correctly in 45 s. The alternate-form reliability estimate is .97 (Torgesen et al., 1999). Oral Reading Fluency was assessed using the passage "Mac Gets Well" (Makar, 1995), which consists of a high percentage of high-frequency and wholly decodable words; students are scored in terms of the number of words read correctly in 1 min. Vadasy, Sanders, and Peyton (2006) reported Cronbach's alpha of .93 with kindergarten students. Concurrent intercorrelations between the word reading measures in our sample ranged from .83 to .88.

2.2.10. First-grade reading skills

Our model of first-grade reading skills included the Word Identification, Word Attack, and Oral Reading Fluency measures described previously, as well as the Passage Comprehension subtest from the WRMT-R/NU. Passage Comprehension consists of several tasks, including matching pictures to rebus items, matching pictures to printed words, and a modified cloze procedure whereby the student produces a missing word that best completes each sentence. Internal consistency exceeds .90 through grade 3 (Woodcock, 1998). Intercorrelations between the first-grade variables ranged from .73 to .86 in our sample.

2.3. Procedures

The data utilized in this study were collected at the end of kindergarten and first grade (i.e., in May of each school year). All measures were individually administered by graduate students and research staff who received approximately 2 h of training per measure. All examiners were required to demonstrate 100% accuracy during practice administration sessions, and meet 100% inter-rater agreement while being observed with study participants prior to independently administering assessments. Measures were administered individually with students in quiet locations outside their classrooms (e.g., unused classrooms and offices). Total testing time for each student was approximately 40 min distributed across two sessions and two days.

Two trained individuals independently scored all test protocols, including each spelling scoring metric. Scoring disagreements were reviewed by a larger group of research staff when they occurred to resolve the disagreements.

2.4. Data analyses

2.4.1. Overview

Analyses investigated the degree to which each spelling scoring metric was associated with reading subskills, including phonological awareness, pseudoword decoding, and word reading. We began by specifying a kindergarten measurement model containing a Phonological Awareness factor (measured by Sound Matching, Blending Words, and Phoneme Segmentation Fluency) and a Pseudoword Decoding factor (measured by Word Attack, Phonemic Decoding Efficiency, and Nonsense Word Fluency). This portion of the model is displayed as the dependent variables in Fig. 1.

The word reading measures (Word Identification, Sight Word Efficiency, and Oral Reading Fluency) were highly correlated with the pseudoword decoding measures, and including a Word Reading factor within the same model as the Pseudoword Decoding factor introduced problems of multicollinearity. Therefore, a separate measurement model (see Fig. 2) was specified that contained a Word Reading factor measured by Word Identification, Sight Word Efficiency, and Oral Reading Fluency. This model was just identified (i.e., equal number of known and unknown parameters).

A final measurement model was specified to capture reading skills at the end of first grade. As depicted on the right side of Fig. 3, the First-Grade Reading factor was measured by the spring administrations of Word Identification, Word Attack, Oral Reading Fluency, and Passage Comprehension.

We used structural equation modeling (SEM) to investigate relations between the spelling metrics and reading skills. Analyses were conducted using Mplus version 7 (Muthén & Muthén, 1998–2012). Across the SEM analyses, missing data were handled using full information maximum likelihood which uses all available data when estimating parameters for variables with missing data. The only appreciable missing data occurred for the kindergarten administrations of Sight Word Efficiency and Phonemic Decoding Efficiency,

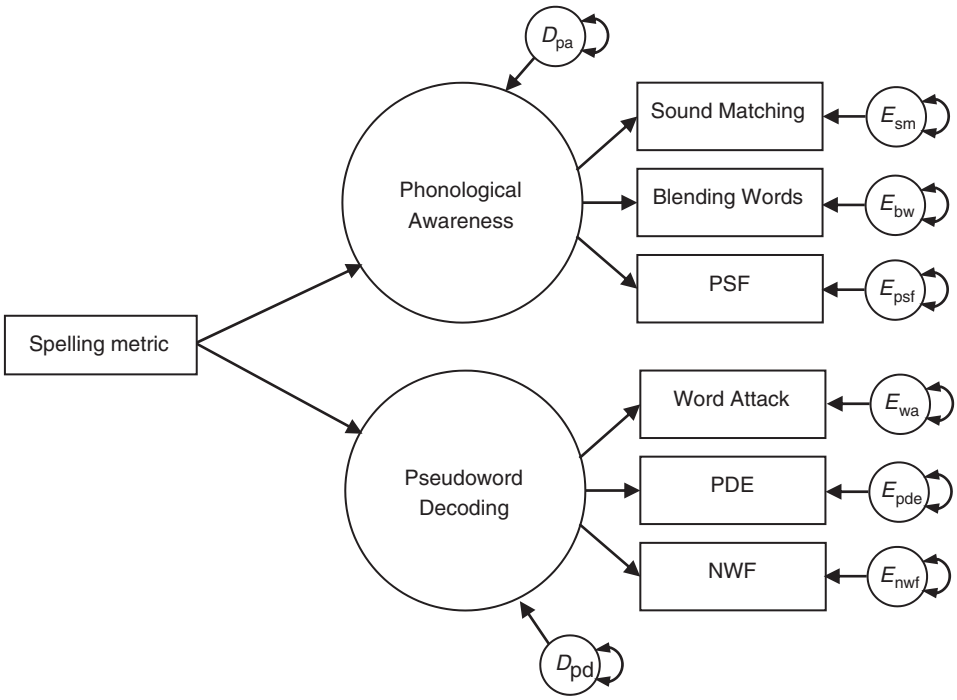

Fig. 1. Models examining the concurrent validity of the kindergarten spelling scoring metrics with kindergarten phonological awareness and pseudoword decoding skills. *Note.* Models were run separately for each spelling scoring metric due to the high degree of correlation between each metric. PSF = Phoneme Segmentation Fluency; NWF = Nonsense Word Fluency; PDE = Phonemic Decoding Efficiency. Models that included the Word Reading factor were run separately.

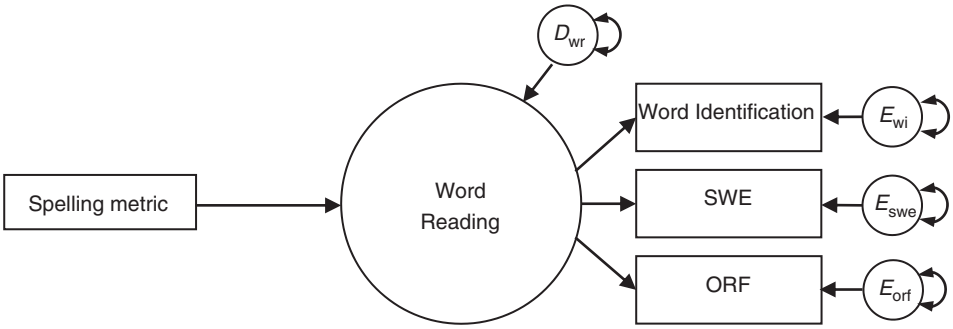

Fig. 2. Models examining the concurrent validity of the kindergarten spelling scoring metrics with kindergarten word reading skills. *Note.* Models were run separately for each spelling scoring metric due to the high degree of correlation between each metric. SWE = Sight Word Efficiency; ORF = Oral Reading Fluency.

which were not administered to students in Florida (n = 136; 47% of the sample) due to an administrative error, and 50 students (17% of total sample) did not have available data for the analyses predicting first-grade reading.[1] To control for the multilevel nature of our data (i.e., 287 students from three states, nested within three schools, further nested within 94 intervention groups), the "TYPE = COMPLEX" option was used with maximum likelihood with robust estimators (MLR) as the estimation method. Because variance at the individual student level was of primary interest, intervention groups were specified as the "CLUSTER" variable. The MLR estimation procedure accounts for potential dependency in the data by adjusting the standard errors, which reduces Type 1 error and protects against spurious findings. Model fit was evaluated using the following fit indices and criteria for adequate fit (Kline, 2005; Sharma, Mukherjee, Kumar, & Dillon, 2005): Comparative fit index (CFI; ≥.90), Tucker–Lewis index (TLI; ≥.90), root mean square error of approximation (RMSEA; ≤.08), and standardized root mean square residual (SRMR; ≤.10). When interpreting model results, we relied primarily on standardized path coefficients (β) and proportion of explained variance in the dependent variables (R^2). Standardized path coefficients can be interpreted as effect sizes (Kline, 2005), and indicate the standard deviation change in a dependent variable for every standard deviation change in the predictor. An alpha level of .05 was used for all statistical significance tests across the study.

2.4.2. RQ1: What differences are observed between kindergarten spelling scoring metrics in their relations to reading subskills on a concurrent basis in kindergarten?

To address Research Question 1, concurrent validity of the spelling metrics was evaluated by analyses that used each spelling metric as a predictor of the Phonological Awareness, Pseudoword Decoding, and Word Reading factors (see Figs. 1 and 2). Scores on each of the spelling metrics were highly intercorrelated (rs = .94–.99), and to include highly correlated predictors in the same model would introduce problems related to multicollinearity and complicate interpretation. One method for addressing multicollinearity is to eliminate one or more variables (Kline, 2005), and because we were primarily interested in the degree to which each spelling scoring metric was individually associated with the reading skills measured in this study, we investigated each spelling metric in separate models. For example, the Word Spelling metric was modeled as a predictor of the Phonological Awareness and Pseudoword Decoding factors. Next, in separate models, each of the remaining spelling metrics were modeled as predictors of the Phonological Awareness and Pseudoword Decoding factors. This process was repeated with the Word Reading factor (Fig. 2).

2.4.3. RQ2: What differences are observed across the kindergarten spelling metrics in predicting reading skills one year later?

Research Question 2 involved investigating the predictive validity of the spelling scoring metrics. For this step, we used the previously described First-Grade Reading factor as the dependent variable. At this stage, it was critical to evaluate whether kindergarten spelling is statistically significant in the prediction of first-grade reading over and above reading skills measured in kindergarten. Otherwise, the use of spelling assessment would be of little predictive utility because it would add nothing over simply assessing reading skills at an earlier point in time. Accounting for the autoregressive effects of prior reading achievement is important and often overlooked in predictive validity studies of reading (Gollob & Reichardt, 1987; Wagner, Torgesen, & Rashotte, 1994; Wagner et al., 1997). Therefore, in order to evaluate whether kindergarten spelling was a unique predictor of first-grade reading skills over and above the predictive effects of kindergarten reading skills, kindergarten Word Identification was included as a predictor in all five models predicting first-grade outcomes (see Fig. 3).

3. Results

Descriptive statistics for the variables assessed in kindergarten and first grade are displayed in Table 1.

[1] Missing data were treated as missing at random because data were missing due to (a) a procedural oversight in the case of Sight Word Efficiency and Phonemic Decoding Efficiency, or (b) students that had moved from the districts prior to data collection in first grade; data were not missing as a function of the values that would have been present in those occasions.

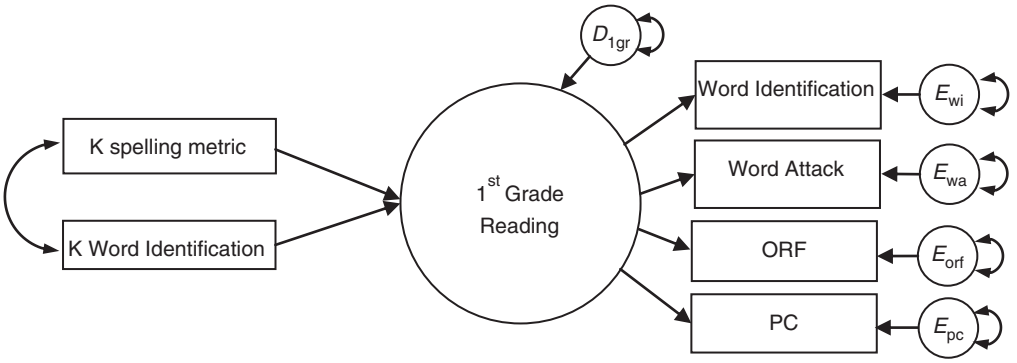

Fig. 3. Models investigating the validity of the kindergarten spelling scoring metrics predicting first grade reading skills, while controlling for the autoregressive effects of kindergarten word reading. *Note.* Models were run separately for each spelling scoring metric due to the high degree of correlation between each metric. ORF = Oral Reading Fluency; PC = Passage Comprehension; K = kindergarten.

3.1. RQ1: What differences are observed between kindergarten spelling scoring metrics in their relations to reading subskills on a concurrent basis in kindergarten?

Fit statistics from the analysis of the two factor measurement model (i.e., Phonological Awareness and Pseudoword Decoding) displayed in Fig. 1 supported the two-factor solution, as adequate fit was observed (CFI = .98, TLI = .97, RMSEA = .08, 90% CI [.04–.12], SRMR = .03). The Word Reading factor was just identified (i.e., equal number of known and unknown parameters). We next added each of the spelling scoring metrics as predictors of the Phonological Awareness, Pseudoword Decoding, and Word Reading factors in the spring of kindergarten. As described earlier, models were run separately for each spelling metric to avoid issues of multicollinearity and to examine the ability of any single metric to explain variance on the dependent variables.

Table 2 displays the results of models depicted in Figs. 1 and 2, and includes the standardized path coefficients from the spelling scoring metrics that served as predictors of the reading subskill factors (β), explained variance on the reading subskill factors (R^2), and fit indices for each model. A high level of consistency was observed across obtained values associated with the spelling scoring metrics, both with regard to the path coefficients with the factors and the proportion of explained variance. High path coefficients were observed for all metrics across each reading subskill factor, which ranged from .80 to .89, indicating large effects. Explained variance on the factors ranged from 64% to 79%. The scoring metrics that provided credit for partial or approximate spellings (Sound Spelling, Letter Sequences, Rubric, and SSS) accounted for 6% to 14% more variance than Word Spelling across the factors. For the Phonological Awareness factor, the Rubric metric accounted for the highest proportion of variance (76%). Very minor differences in explained variance were observed across the Pseudoword Decoding or Word Reading factors for the Sound Spelling, Letter Sequences, Rubric, or SSS metrics, which ranged from 74% to 79%.

Table 2
Concurrent validity of kindergarten spelling scoring metrics in relation to kindergarten Phonological Awareness, Pseudoword Decoding, and Word Reading.

spelling scoring metric	Factor	β	R^2	Model fit			
				CFI	TLI	RMSEA (90% CI)	SRMR
Word Spelling	Phonological Awareness	.80	.64				
	Pseudoword Decoding	.84	.71	.99	.98	.06 (.02–.09)	.02
	Word Reading	.84	.70	.99	.96	.13 (.06–.20)	.01
Sound Spelling	Phonological Awareness	.85	.72				
	Pseudoword Decoding	.88	.78	.98	.97	.07 (.04–.10)	.02
	Word Reading	.87	.75	1.00	1.00	.05 (.00–.14)	.01
Letter Sequences	Phonological Awareness	.84	.70				
	Pseudoword Decoding	.88	.78	.98	.97	.07 (.04–.10)	.02
	Word Reading	.87	.76	1.00	.99	.06 (.00–.15)	.01
Rubric	Phonological Awareness	.87	.76				
	Pseudoword Decoding	.89	.79	.98	.96	.08 (.05–.11)	.03
	Word Reading	.86	.74	1.00	.99	.05 (.00–.14)	.01
Spelling Sensitivity Score	Phonological Awareness	.84	.71				
	Pseudoword Decoding	.88	.77	.98	.97	.08 (.04–.11)	.02
	Word Reading	.86	.75	1.00	1.00	.03 (.00–.12)	.01

Note. Models are depicted in Figs. 1 and 2. All path coefficients and R^2 values were significant at $p < .05$. The Phonological Awareness factor was measured by Sound Matching, Blending Words, and Phonemic Segmentation Fluency; the Pseudoword Decoding factor was measured by Word Attack, Phonemic Decoding Efficiency, and Nonsense Word Fluency; and the Word Reading factor was measured by Word Identification, Sight Word Efficiency, and Oral Reading Fluency.

3.2. RQ2: What differences are observed across the kindergarten spelling metrics in predicting reading skills one year later?

We next examined the degree to which each spelling scoring metric explained variance in reading skills on a predictive basis. The First-Grade Reading factor (see Fig. 3) fits the data well (CFI = 1.00, TLI = 1.00, RMSEA = .03, 90% CI [.00–.14], SRMR = 0.01). Each kindergarten spelling scoring metric was then used as a predictor of this factor in separate individual models. Good fit was observed for each predictive model, and the results are displayed in Table 3.

As reported in the first five rows of Table 3, when modeled individually as predictors of the First-Grade Reading factor, the spelling metrics each demonstrated similar relations to the First-Grade Reading factor. Path coefficients indicated large effects and ranged from .62 to .68.

The remaining rows of Table 3 report the results of models (depicted in Fig. 3) that controlled for the autoregressive effects of kindergarten word reading. In each model, the correlation between the predictors (i.e., each spelling metric and Word Identification) ranged from .80 to .83. As would be expected, kindergarten Word Identification demonstrated stronger path coefficients in the prediction of first grade reading in contrast to spelling scores. However, across all scoring metrics, while simultaneously accounting for kindergarten word identification skills, kindergarten spelling was a statistically significant predictor of first-grade reading skills. Scores on the Sound Spelling, Letter Sequences, Rubric, and SSS metrics played a slightly greater role in this prediction, as the path coefficients for these metrics indicated moderate effects and ranged from .28 to .33, compared to .19 for Word Spelling.

4. Discussion

Despite the established interrelation between spelling and early reading development, and calls for greater attention to spelling assessment that is not restricted to grade or ability levels (Al Otaiba & Hosp, 2010), the potential for spelling assessment to inform the evaluation of reading skills in practice has yet to be fully realized. To advance understanding in this area, we compared five metrics for scoring the spelling responses of kindergarten students who had received supplementary intervention to target reading difficulties, on the degree to which the metrics were associated with subskills of early reading development.

Overall, the results of this study are consistent with recent research on the validity of various spelling scoring metrics (Ritchey, 2008; Ritchey et al., 2010). Although the metrics varied in their complexity and the manner in which they awarded credit for spelling responses, a high degree of similarity was observed in the degree to which they were associated with reading subskills. Ritchey et al. (2010) found that correct words, letter sounds, letter sequences, and rubric-based metrics showed a high degree of similarity in relation to basic reading (phonemic awareness, letter naming, nonsense word fluency, and word reading) and writing skills. Our results confirmed and extended this work through the use of a latent variable approach with an expanded set of reading skills measures, extended the results to a predictive basis, utilized a sample of students who received intervention to target reading difficulties, and added an additional metric (SSS).

Our first research question concerned the degree to which each scoring metric was associated with reading and reading-related subskills on a concurrent basis at the end of kindergarten. A high degree of similarity was observed across the spelling metrics in their relations to reading skills, consistent with the findings of prior work (Ritchey et al., 2010). Structural equation models indicated that

Table 3
Predictive validity of kindergarten spelling scoring metrics predicting first-grade reading skills, including analyses of prediction when accounting for autoregressive effects of kindergarten word reading.

	1st Grade Reading factor		Model fit			
Kindergarten predictors	β	R^2	CFI	TLI	RMSEA (90% CI)	SRMR
Word Spelling	.62	.39	1.00	1.00	.02 (.00–.09)	.01
Sound Spelling	.68	.46	1.00	1.00	.04 (.00–.10)	.01
Letter Sequences	.67	.45	1.00	1.00	.04 (.00–.10)	.01
Rubric	.68	.46	1.00	.99	.05 (.00–.11)	.01
Spelling Sensitivity Score	.66	.44	1.00	1.00	.05 (.00–.10)	.01
Word Spelling	.19					
Word Identification	.54	.49	.99	.98	.06 (.01–.10)	.02
Sound Spelling	.32					
Word Identification	.43	.51	.99	.99	.06 (.00–.10)	.02
Letter Sequences	.31					
Word Identification	.43	.51	.99	.99	.06 (.01–.10)	.02
Rubric	.33					
Word Identification	.42	.51	.99	.99	.07 (.03–.11)	.02
Spelling Sensitivity Score	.28					
Word Identification	.46	.50	.99	.99	.06 (.01–.10)	.02

Note. Models are depicted in Fig. 3. All path coefficients and R^2 values were significant at $p < .05$. 1st-Grade Reading factor was measured by Word Identification, Word Attack, Oral Reading Fluency, and Passage Comprehension.

kindergarten spelling skills accounted for a very high proportion (64% to 79%) of the variance in the factors measuring phonological awareness, pseudoword decoding, and word reading. Metrics that awarded points for partial or invented spellings (Sound Spelling, Letter Sequences, Rubric, and SSS) explained between 6% and 15% more variance than conventional Word Spelling, consistent with evidence supporting the importance of partial or invented spelling in assessment of early learners (Lombardino et al., 1997; Richgels, 1995; Silva & Alves Martins, 2003).

Our second research question concerned the validity of the kindergarten spelling metrics predicting reading skills one year later. Kindergarten spelling explained a sizeable proportion of variance in reading skills at the end of first grade (ranging from 39% to 46% across metrics). As in the kindergarten concurrent analyses, metrics that were sensitive to partial or invented spellings (Sound Spelling, Letter Sequences, Rubric, and SSS) explained somewhat more variance in subsequent reading skills than conventional scoring (Word Spelling). Moreover, kindergarten spelling scores were a unique predictor of reading skills in first grade even when the autoregressive effects of kindergarten word reading were controlled, indicating that kindergarten spelling assessment added unique information in the prediction of subsequent reading achievement over and above that explained by kindergarten word reading skills. These findings support arguments that spelling assessment may provide information on students' early reading development in unique ways (Ouellette & Sénéchal, 2008b). The present findings are also of note in demonstrating the predictive power of spelling assessment with a sample of students considered at school entry to be at risk for reading difficulty, and add to the body of work suggesting that spelling may serve as a valuable tool for predicting subsequent reading achievement (De Graff & Torgesen, 2005; Lombardino et al., 1999; NELP, 2008).

In summary, all of the kindergarten spelling metrics investigated in this study were associated with reading subskills measured concurrently (kindergarten) as well as one year later (first grade). Metrics that awarded credit for partial or invented spellings captured slightly more variance in the Phonological Awareness, Decoding, and Word Reading factors than simply scoring words as either correct or incorrect. Relative differences among spelling metrics were observed when considering the Phonological Awareness factor. Consistent with its aim and intent (Tangel & Blachman, 1992), the Rubric metric demonstrated somewhat stronger relations and accounted for more variance in phonological awareness skills compared to the other spelling metrics. However, a high degree of similarity was still observed across the scoring approaches.

4.1. Limitations

Several limitations of this study should be acknowledged. Because the spelling metrics were so highly correlated with each other, multicollinearity concerns prevented their simultaneous inclusion as predictors in the same model. Similar issues prevented us from including the Word Reading factor in the same model as the Decoding factor. To address these issues we analyzed separate models, but as a result the analyses were less parsimonious than if we were able to include all predictors and factors together.

Second, the spelling assessment (TWS-4) was designed for students in first grade and up, and was not specifically designed for investigating students' spelling patterns to the extent we did. Although the words on the early portion of the TWS-4 are short and occur frequently in English, analyses of kindergarten students' spelling skills would likely be enhanced by using spelling assessments developed specifically for beginning spellers.

Third, we utilized a sample that had been recruited for an intervention study based on their at-risk status at the beginning of kindergarten. Although this may permit generalization to other kindergarten students considered to be at-risk for reading difficulty, results may not generalize across the wider kindergarten population.

Fourth, we did not formally evaluate children's fine motor skills, although based on observations of students' spelling behavior, we are confident that their grapho-motor skills were sufficient to complete the tasks. Practitioners can consider alternative spelling modalities (e.g., oral spelling or letter tiles) if fine motor skills are underdeveloped; however, the spelling modality should not affect spelling scores once students can write at least 19 letters (Puranik & Apel, 2010).

4.2. Practical implications for assessment: selecting a spelling scoring metric

Although the scoring metrics demonstrated only minor quantitative differences, it is instructive to consider the qualities and characteristics of each that may have implications for their potential for informing reading assessment. As evident in Appendix A, the scoring metrics varied in terms of their complexity and dependence on scoring rules. For example, Correct Words provides the simplest and most straightforward scoring metric; however, it accounted for lower variance in reading skills than the other metrics investigated in this study. This finding is not surprising, as it is not sensitive to approximate and invented spellings and may be the least likely to provide instructionally-relevant information on individual differences in developing reading skills. Therefore, its utility for informing reading assessment and instruction with early learners or low-achieving students may be limited.

Sound Spelling is more sensitive to invented or approximate spellings, because it counts phonetically consistent letter sounds as correct, and thus can provide important information on students' alphabetic knowledge. Additionally, information on students' phonemic awareness is yielded by noting how accurately students can parse and sequence initial, medial, and final sounds. Scoring for letter sounds involves some degree of subjectivity (e.g., can *er*, *ur*, or *ir* stand for *r*?), and it can be difficult to score when going beyond short (i.e., CVC) words that are phonetically regular. Also, information on orthographic or morphological awareness is not readily apparent, because students receive equal credit for misspelled and correctly spelled words, thus potentially obscuring more sophisticated reading and spelling skills. In short, Sound Spelling may be a preferred assessment metric with beginning readers and spellers when their knowledge and application of grapho-phoneme relations are of primary interest.

The Letter Sequences metric introduces very little subjectivity while providing some sensitivity to approximate spellings. Its dependence on correct letters (i.e., phonologically "legal" letter sounds are not counted as correct) might be considered both a strength and a weakness; by counting only correct letters, the metric might provide information on students' orthographic and morphological knowledge, because these skills represent knowledge of letter combinations that may not be phonetically regular. On the other hand, Letter Sequences has the potential to lose sensitivity to individual differences in grapho-phonemic knowledge, because phonetically legal letter sounds are not counted as correct using this approach. Therefore, Letter Sequences may be preferred when one seeks to determine how well students are beginning to develop knowledge of reoccurring letter patterns (i.e., early orthographic awareness).

The Rubric investigated in this study was highly sensitive to phonological awareness skills, which is consistent with its purpose and design (Tangel & Blachman, 1992). Because more points are awarded for spellings that are more orthographically accurate, the scaled nature of the rubric may also make it useful for identifying students with more sophisticated literacy skills while still providing sensitivity to basic phonemic and alphabetic skills. However, rubric-scoring metrics tend to be complex and often include a large number of scoring rules or potential ways in which students might receive points for a particular response. Although prior work has established good inter-rater reliability (Tangel & Blachman, 1992), rubric metrics may require additional training to reduce scoring errors or pre-established frameworks for scoring various responses, which may complicate their application to user-generated spelling lists.

The SSS metric may provide a good mix of the approaches described above. Each phonemic segment is scored on an objective 0–3 point scale, thus providing a good deal of sensitivity to partial and invented spellings. Further, it retains the advantage of rubrics in awarding more points for correct letters and allows for scoring phonetically legal letters using objective criteria for each segment. Although growth was not investigated in this study, SSS was designed as an index of spelling growth over time (Masterson & Apel, 2010b). Thus, this metric may be preferred if formative spelling assessment is of interest. Additionally, students receive significantly more points for using correct letters and letter combinations; therefore, SSS may be useful for gathering information on more sophisticated orthographic and morphological skills. SSS scoring may be more prone to error than letter sequences, letter sounds, or correct words, but software exists for aiding scoring (Masterson & Hrbec, 2011).

In summary, Sound Spelling and Rubric metrics may be somewhat better suited for beginning readers and spellers, when practitioners are interested in learning how students are applying early skills in phonemic awareness and alphabetic understanding and less concerned about students' use of correct letters. On the other hand, Word Spelling, Letter Sequences, and SSS may be better suited for students for whom use of correct letters is expected, or if practitioners are interested in determining whether students are beginning to demonstrate orthographic or morphological awareness and building a memory for reoccurring letter patterns. The partial and invented scoring metrics, and especially SSS, may be advantageous for monitoring student progress over time, as their sensitivity to small changes in spelling skills may make them well suited for monitoring spelling growth. Research investigating different scoring metrics for monitoring spelling growth is underway (Clemens, Simmons, & Davis, 2013).

4.3. Future directions

Our analyses of validity included investigations of path coefficients and the amount of variance in reading skills explained by the scoring metrics. Although these statistics are important for establishing their relationship to reading skills, it does not fully reveal the degree to which these metrics might be useful for directly informing reading assessment with kindergarten students. Answering this question would involve more in-depth or qualitative analysis of students' spelling responses and their consistency with reading skills.

Additional implications pertain to test development, as well as the ways in which spelling is considered within a reading assessment. The present results reinforce the importance of spelling assessment in early elementary school; thus, standardized assessments should not neglect early grade levels for measuring spelling skills. Consequently, these assessments should include scoring options that are sensitive to partial or invented spelling responses. On a more general level, we advocate that educators and practitioners view spelling not as a discrete skill or something specific to written expression, but as an ability closely tied to reading development and an important component of a comprehensive reading assessment battery.

4.5. Conclusion

Although close interrelations between spelling and reading development are well established, this knowledge is not often extended to the application of spelling assessment with beginning readers. Consistent with prior work, this study demonstrated that kindergarten spelling skills are highly associated with reading subskills, both on a concurrent and a predictive basis, and may provide unique information on students' early literacy functioning and in the prediction of subsequent reading skills over that provided by traditional reading assessments alone. Scoring metrics designed to be sensitive to partial or invented spelling afford some advantages over simply tallying the number of words spelled correctly and are more sensitive to students' phonological awareness skills. Although scoring metrics were relatively similar in their relations to early reading, practitioners should consider their alignment with the goals and purpose of the assessment when selecting a metric for scoring spelling responses. Like scholars who have called for the integration of spelling and reading instruction (e.g., Moats, 2005; Santoro et al., 2006; Treiman, 1998), we advocate for greater utilization of spelling assessment to inform more comprehensive reading evaluations.

Appendix A. Description of spelling scoring metrics and scoring examples

Scoring metric	Description	Example scoring for target word *fan**
Word Spelling	One point awarded for each word spelled correctly and 0 points for all other spellings. No partial credit is given.	fan = 1 any other response = 0
Sound Spelling	One point awarded for each sound spelled with correct letter or with a phonetically consistent substitution. Extra letters are ignored.	fan, phan = 3 fa, pha = 2 n = 1
Letter Sequences	One point awarded each for the correct first and last letters, and 1 point for each pair of correctly sequenced letters between the first and last letters. Phonetically consistent letters are not counted as correct.	fan = 4 fa, fn = 2 f, phn = 1
Rubric	Points awarded on a 0-to-6 scale for all words regardless of length: 6 — word spelled correctly 5 — long vowels attempted and all other sounds spelled correctly 4 — one sound spelled with a phonetically related letter(s) and all other sounds spelled correctly 3 — initial sound spelled correctly and a minimum of one additional sound spelled correctly without penalty for sequence errors 2 — initial sound spelled correctly 1 — initial sound spelled with phonetically related letter(s) or another sound spelled correctly 0 — no sounds represented with correct or phonetically unrelated letters	fan = 6 phan = 4 faq, fqa = 3 f, fq = 2 ph, n = 1
Spelling Sensitivity Score	Word is broken into segments (phonemes, affixes, etc.), and each segment is scored on a 0-to-3 scale according to its orthographic or phonemic consistency: 3 — segment spelled with correct letter 2 — spelled with a phonetically consistent letter(s) 1 — segment spelled with a letter that is neither correct nor phonetically consistent 0 — segment represented by non-alphabetic characters (numbers, scribbles), or segment not attempted.	fan = 9 phan = 8 faq, fqn, qan = 7 phaq, phqn = 6 pha, phn, fqq = 5 phqq, fq, aq = 4 f, a, n, phq = 3 ph, qq = 2 q = 1

Note. * "q" represents any phonologically unrelated letter.

References

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

Ahmed, S. T., & Lombardino, L. J. (2000). Invented spelling: An assessment and intervention protocol for kindergarten children. *Communication Disorders Quarterly*, *22*, 19–28.

Al Otaiba, S., Folsom, J. S., Schatschneider, C., Wanzek, J., Greulich, L., Meadows, J., et al. (2011). Predicting first-grade reading performance from kindergarten response to tier 1 instruction. *Exceptional Children*, *77*, 453–470.

Al Otaiba, S., & Hosp, J. L. (2010). Spell it out: The need for detailed spelling assessment to inform instruction. *Assessment for Effective Intervention*, *36*, 3–6.

Apel, K., Thomas-Tate, S., Wilson-Fowler, E. B., & Brimo, D. (2012). Acquisition of initial mental graphemic representations by children at risk for literacy development. *Applied Psycholinguistics*, *33*, 365–391.

Bryant, P. E., MacLean, M., Bradley, L. L., & Crossland, J. (1990). Rhyme and alliteration, phoneme detection, and learning to read. *Developmental Psychology*, *26*, 429–438.

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language*, *45*, 751–774.

Carlisle, J. F. (2003). Morphology matters in learning to read: A commentary. *Reading Psychology*, *24*, 291–322.

Clemens, N. H., Simmons, L., & Davis, J. L. (February 2013). *Can we identify ideal measures for monitoring kindergarten reading progress?. Poster presented at the annual Pacific Coast Research Conference, San Diego, CA.*

Cooke, N. L., Slee, J. M., & Young, C. A. (2008). How is contextualized spelling used to support reading in first-grade core reading programs? *Reading Improvement*, *45*, 26–45.

Coyne, M. D., Little, M., Rawlinson, D. A., Simmons, D., Kwok, O. M., Kim, M., et al. (2013). Replicating the impact of a supplemental beginning reading intervention: The role of instructional context. *Journal of Research on Educational Effectiveness*, *6*, 1–23.

De Graff, A. J., & Torgesen, J. K. (2005). Measuring alphabetic reading skills in a group: New alternatives. *Assessment for Effective Intervention*, *30*(3), 1–14.

Ehri, L. (1997). Learning to read and learning to spell are one and the same, almost. In C. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 237–269)Mahwah, NJ: Erlbaum.

Ehri, L. C. (1998). Research on learning to read and spell: A personal – historical perspective. *Scientific Studies of Reading*, *2*, 97–114.

Ehri, L. C. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders*, *20*(3), 19–36.

Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty school. *Contemporary Educational Psychology*, *31*, 1–29.

Furnes, B., & Samuelsson, S. (2009). Preschool cognitive and language skills predicting kindergarten and grade 1 reading and spelling: A cross-linguistic comparison. *Journal of Research in Reading*, *32*, 275–292.

Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, *58*, 80–92.

Good, R. H., & Kaminski, R. A. (Eds.). (2007). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* Hillsdale, NJ: Erlbaum.

Hecht, S. A., & Close, L. (2002). Emergent literacy skills and training time uniquely predict variability in responses to phonemic awareness training in disadvantaged kindergartners. *Journal of Experimental Child Psychology*, *82*, 93–115.

Hoillingworth, L. S. (1923). Spelling. In L. S. Hoillingworth (Ed.), *Special talents and defects: Their significance for education* (pp. 98–113)New York, NY: MacMillan.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement.* New York, NY: Guilford Press.

Johnston, F. R. (2001). Exploring classroom teachers' spelling practices and beliefs. *Reading Research and Instruction*, *40*, 143–155.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Larsen, S. C., Hammill, D. D., & Moats, L. C. (1999). *Test of written spelling* (4th ed.). Austin, TX: Pro-Ed.

Lombardino, L. J., Bedford, T., Fortier, C., Carter, J., & Brandi, J. (1997). Invented spelling: Developmental patterns in kindergarten children and guidelines for early literacy intervention. *Language, Speech, and Hearing Services in Schools*, *28*, 333–343.
Lombardino, L. J., Morris, D., Mercado, L., DeFillipo, F., Sarisky, C., & Montgomery, A. (1999). The early reading screening instrument: A method for identifying kindergarteners at risk for learning to read. *International Journal of Language and Communication Disorders*, *34*, 135–150.
Makar, B. W. (1995). *Primary phonics.* Cambridge, MA: Educators Publishing Service.
Masterson, J. J., & Apel, K. (2010). Linking characteristics discovered in spelling assessment to intervention goals and methods. *Learning Disability Quarterly*, *33*, 185–198.
Masterson, J. J., & Apel, K. (2010). The spelling sensitivity score: Noting developmental changes in spelling knowledge. *Assessment for Effective Intervention*, *36*, 35–45.
Masterson, J., & Hrbec, B. (2011). *Computerized Spelling Sensitivity System [computer software].* Springfield, MO: Missouri State University Language-Literacy Lab.
Moats, L. C. (1993). Spelling error interpretation: Beyond the phonetic/dysphonetic dichotomy. *Annals of Dyslexia*, *43*, 174–185.
Moats, L. C. (2005). How spelling supports reading and why it is more regular and predictable than you may think. *American Educator*, *29*(4), 12–22 (42–43).
Morris, D., Bloodgood, J., & Perney, J. (2003). Kindergarten predictors of first- and second-grade reading achievement. *The Elementary School Journal*, *104*, 93–109.
Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1997). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology*, *65*, 370–396.
Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly*, *32*, 154–167.
National Early Literacy Panel (NELP) (2008). *Developing early literacy: Report of the National Early Literacy Panel, a scientific synthesis of early literacy development and implications for intervention.* Washington, DC: National Institute for Literacy.
O'Connor, R. E., & Jenkins, J. R. (1995). Improving the generalization of sound/symbol knowledge: Teaching spelling to kindergarten children with disabilities. *The Journal of Special Education*, *29*, 255–275.
Ouellette, G. (2010). Orthographic learning in learning to spell: The roles of semantics and type of practice. *Journal of Experimental Child Psychology*, *107*, 50–58.
Ouellette, G., & Sénéchal, M. (2008). Pathways to literacy: A study of invented spelling and its role in learning to read. *Child Development*, *79*, 899–913.
Ouellette, G. P., & Sénéchal, M. (2008). A window into early literacy: Exploring the cognitive and linguistic underpinnings of invented spelling. *Scientific Studies of Reading*, *12*, 195–219.
Pearson/Scott Foresman (2004). *Scott Foresman sidewalks: Early reading intervention.* Glenview, IL: Author.
Puranik, C., & Apel, K. (2010). Effect of assessment task and letter writing ability on preschool children's spelling performance. *Assessment for Effective Intervention*, *36*, 46–56.
Puranik, C. S., Lonigan, C. J., & Kim, Y. S. (2011). Contributions of emergent literacy skills to name writing, letter writing, and spelling in preschool children. *Early Childhood Research Quarterly*, *26*, 465–474.
Rao, C., Prakash, P., & Joshi, M. (2006). Development of a scoring system for Kannada spelling assessment. *Psychological Studies*, *51*, 49–51.
Richgels, D. J. (1995). Invented spelling ability and printed word learning in kindergarten. *Reading Research Quarterly*, *30*, 96–109.
Ritchey, K. D. (2008). The building blocks of writing: Learning to write letters and spell words. *Reading and Writing*, *21*, 27–47.
Ritchey, K. D., Coker, D. L., Jr., & McCraw, S. B. (2010). A comparison of metrics for scoring beginning spelling. *Assessment for Effective Intervention*, *35*, 78–88.
Ritchey, K. D., & Speece, D. L. (2006). From letter names to word reading: The nascent role of sublexical fluency. *Contemporary Educational Psychology*, *31*, 301–327.
Robbins, K. P., Hosp, J. L., Hosp, M. K., & Flynn, L. J. (2010). Assessing specific grapho-phonemic skills in elementary students. *Assessment for Effective Intervention*, *36*, 21–34.
Santi, K. L., York, M., Foorman, B. R., & Francis, D. J. (2009). The timing of early reading assessment in kindergarten. *Learning Disability Quarterly*, *32*, 217–227.
Santoro, L. E., Coyne, M. D., & Simmons, D. C. (2006). The reading–spelling connection: Developing and evaluating a beginning spelling intervention for children at risk of reading disability. *Learning Disabilities Research and Practice*, *21*, 122–133.
Sénéchal, M., & Kearnan, K. (2007). The role of morphology in reading and spelling. *Advances in Child Development and Behavior*, *35*, 297–325.
Shankweiler, D., & Lundquist, E. (1993). *On the relations between learning to spell and learning to read. Haskins Laboratories Status Report on Speech Research*, *SR-113*. (pp. 135–144).
Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W. R. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. *Journal of Business Research*, *58*, 935–943.
Silva, C., & Alves Martins, M. (2003). Relations between children's invented spelling and the development of phonological awareness. *Educational Psychology*, *23*, 3–16.
Simmons, D. C., Coyne, M. D., Hagan-Burke, S., Kwok, O., Simmons, L., Johnson, C., et al. (2011). Effects of supplemental reading interventions in authentic contexts: A comparison of kindergarteners' response. *Exceptional Children*, *77*, 207–228.
Stage, S. A., & Wagner, R. K. (1992). Development of young children's phonological and orthographic knowledge as revealed by their spellings. *Developmental Psychology*, *28*, 287–296.
Tangel, D. M., & Blachman, B. A. (1992). Effect of phoneme awareness instruction on kindergarten children's invented spelling. *Journal of Reading Behavior*, *24*, 233–261.
Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency.* Austin, TX: Pro-Ed.
Treiman, R. (1998). Why spelling? The benefits of incorporating spelling into beginning reading instruction. In J. L. Metsala, & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 289–313)Mahwah, NJ: Erlbaum.
Treiman, R., Berch, D., Tincoff, R., & Weatherston, S. (1993). Phonology and spelling: The case of syllabic consonants. *Journal of Experimental Child Psychology*, *56*, 267–290.
Treiman, R., & Bourassa, D. C. (2000). The development of spelling skill. *Topics in Language Disorders*, *20*(3), 1–18.
Treiman, R., Cassar, M., & Zukowski, A. (1994). What types of linguistic information do children use in spelling? The case of flaps. *Child Development*, *65*, 1318–1337.
Uhry, J. K., & Shepherd, M. J. (1993). Segmentation/spelling instruction as part of a first-grade reading program: Effects on several measures of reading. *Reading Research Quarterly*, *28*, 218–233.
Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2006). Code-oriented instruction for kindergarten students at risk for reading difficulties: A randomized field trial with paraeducator implementers. *Journal of Educational Psychology*, *98*, 508–528.
Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, *30*, 73–87.
Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing.* Austin, TX: Pro-Ed.
Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skills readers: A 5-year longitudinal study. *Developmental Psychology*, *33*, 468–479.
Weiser, B., & Mathes, P. (2011). Using encoding instruction to improve the reading and spelling performances of elementary students at risk for literacy difficulties: A best-evidence synthesis. *Review of Educational Research*, *81*, 170–200.
Woodcock, R. W. (1998). *Woodcock Reading Mastery Test — Revised/normative update.* Bloomington, MN: Pearson Education.
Wright, D. M., & Ehri, L. (2007). Beginners remember orthography when they learn to read words: The case of doubled letters. *Applied Psycholinguistics*, *28*, 115–133.