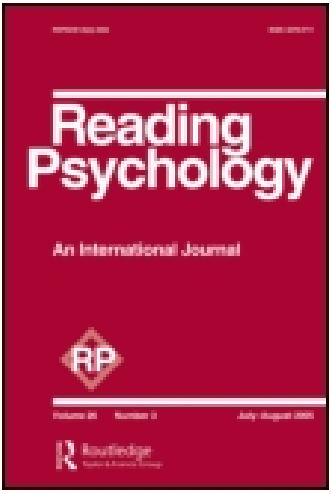
This article was downloaded by: [University of California Santa Cruz]
On: 31 October 2014, At: 14:49
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Reading Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/urpy20

Predicting Kindergarteners' Response to Early Reading Intervention: An Examination of Progress-Monitoring Measures

Eric L. Oslund [a] , Shanna Hagan-Burke [a] , Aaron B. Taylor [a] , Deborah C. Simmons [a] , Leslie Simmons [a] , Oi-Man Kwok [a] , Caitlin Johnson [a] & Michael D. Coyne [b]

[a] Department of Educational Psychology , Texas A&M University , College Station , Texas

[b] Department of Educational Psychology , University of Connecticut , Storrs , Connecticut

Published online: 03 Feb 2012.

To cite this article: Eric L. Oslund , Shanna Hagan-Burke , Aaron B. Taylor , Deborah C. Simmons , Leslie Simmons , Oi-Man Kwok , Caitlin Johnson & Michael D. Coyne (2012) Predicting Kindergarteners' Response to Early Reading Intervention: An Examination of Progress-Monitoring Measures, Reading Psychology, 33:1-2, 78-103, DOI: 10.1080/02702711.2012.630611

To link to this article: http://dx.doi.org/10.1080/02702711.2012.630611

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Reading Psychology, 33:78–103, 2012
Copyright © Taylor & Francis Group, LLC
ISSN: 0270-2711 print / 1521-0685 online
DOI: 10.1080/02702711.2012.630611

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

PREDICTING KINDERGARTENERS' RESPONSE TO EARLY READING INTERVENTION: AN EXAMINATION OF PROGRESS-MONITORING MEASURES

ERIC L. OSLUND, SHANNA HAGAN-BURKE, AARON B. TAYLOR, DEBORAH C. SIMMONS, LESLIE SIMMONS, OI-MAN KWOK, and CAITLIN JOHNSON
Department of Educational Psychology, Texas A&M University, College Station, Texas

MICHAEL D. COYNE
Department of Educational Psychology, University of Connecticut, Storrs, Connecticut

This study examined the predictive validity of combinations of progress-monitoring measures: (a) curriculum-embedded phonemic awareness and alphabetic/decoding measures, and (b) Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002) nonsense word fluency and phoneme segmentation fluency on reading outcomes of kindergarten students in a tier 2 intervention.

Results of multiple-regression analyses indicated that curriculum-embedded mastery checks and DIBELS measures each explained a significant amount of variance on the outcome measure. However, curriculum-embedded measures explained statistically significantly more variance at each time point supporting their utility in documenting progress of kindergarten students receiving intervention.

Response to intervention (RTI) involves the strategic and dynamic integration of assessment and instruction to identify and support students who are at risk of significant academic difficulty (Gersten et al., 2009; Lembke, McMaster, & Stecker, 2010; Mellard, McKnight, & Woods, 2009). A typical RTI model involves screening assessments to identify students whose performance levels indicate a need for additional and more intensive intervention. Following screening, and throughout supplemental intervention, *progress*

Address correspondence to Eric L. Oslund, Department of Educational Psychology, Texas A&M University, 704 Harrington Tower, MS 4225, College Station, TX 77843-4225. E-mail: eoslund@neo.tamu.edu

monitoring is used to evaluate a student's response to intervention (Fuchs & Fuchs, 2006).

In early literacy, progress monitoring involves frequent measurement of reading and reading-related skills to gather information about a student's level and rate of improvement. In the recent RTI practice guide commissioned by the Institute of Education Sciences, panel members recommended using mastery checks from intervention programs to monitor the progress of students receiving Tier 2 interventions (Gersten et al., 2009). The panel noted that when programs do not include mastery checks, school personnel should use technically adequate measures that are reliable and valid for monitoring student progress. The present study examined the predictive validity of kindergarten progress-monitoring measures within an intervention framework.

Schools use a variety of measures to monitor reading progress. In a recent summary of progress-monitoring practices in 41 elementary schools, Mellard et al. (2009) documented that 33% of participating schools monitored progress using published reading assessments that were not tied to a curriculum or intervention. Nineteen percent reported using curriculum-based measures, which they defined as "assessments of a skill (e.g., oral reading fluency) tied to the curriculum of instruction that directly and repeatedly measures student performance" (p. 189). Thirteen percent used DIBELS, and 12% relied on assessments provided by the reading curricula being used. At issue is whether measures that are specific to particular kindergarten reading interventions can reliably predict performance relative to measures such as DIBELS that are designed to be curriculum independent.

Predictive validity involves the accuracy of measurement and the degree to which student performance on progress-monitoring measures corresponds to performance on end-of-year criterion measures. Ideally, progress-monitoring measures enable educators to have confidence in their decisions by estimating both current and future academic performance (Grigorenko, 2009), which makes predictive validity an important aspect of any progress-monitoring tool. When measures indicate that student progress is adequate, decisions may involve maintaining the same type and amount of intervention or even discontinuing an intervention. Conversely, if measures indicate that student response

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

is inadequate, decisions may involve intensifying intervention through curricular and instructional modifications that alter programs, grouping, and time allocations (Fuchs & Fuchs, 2006).

While significant research has been conducted on the oral reading fluency measures used with grades one and above, there is considerably less research on measures that predict kindergarten outcomes. In this study, we compared the predictive validity of two forms of progress-monitoring measures and various combinations of those measures in kindergarten: DIBELS curriculum-based measures and published curriculum-embedded mastery measures. Following, we summarize the relevant literature as context for this study.

DIBELS Fluency-Based Progress Measures

DIBELS are standardized, fluency-based measures that are widely used to formatively monitor early reading development. For example, Hoffman, Jenkins, and Dunlap's (2009) recent survey of members of a state International Reading Association indicated that 60% of respondents used DIBELS for intervention development and 57% used DIBELS for progress monitoring. The DIBELS Phoneme Segmentation Fluency (PSF) and Nonsense Word Fluency (NWF) subtests are particularly relevant for monitoring kindergarten reading progress. PSF, which is designed for use in the middle and end of kindergarten (Good & Kaminski, 1996), measures phonemic awareness, specifically, how well a child can fluently separate individual phonemes in words presented orally (Schilling, Carlisle, Scott, & Ji, 2007). Phonological measures, including PSF, have been shown to be strong indicators of reading outcomes (Good, Simmons, & Smith, 1998; Rouse & Fantuzzo, 2006; Schatschneider, Francis, Carlson, Fletcher, & Foorman, 2004). In a stratified sample of 330 kindergarten children, Rouse and Fantuzzo (2006) found that evidence for the predictive validity of PSF, as measured by bivariate correlations, was .49 with the TerraNova (CTB/McGraw Hill, 2000) reading subtest and .55 with the Developmental Reading Assessment (Beaver, 1997). Both are in the low moderate range.

NWF, which is designed for use in the middle of kindergarten (University of Oregon Center on Teaching and Learning, 2010),

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

measures a student's ability to produce the individual sounds or provide the phonetically correct pronunciation (i.e., letter sound fluency) of an entire word presented visually (Good & Kaminski, 2002). Letter sound fluency has been demonstrated to predict reading outcomes (Bishop & League, 2006; Schatschneider et al., 2004). Recent studies have documented the predictive validity of NWF (Burke, Hagan-Burke, Kwok, & Parker, 2009; Fien et al., 2008; Ritchey, 2008). Burke et al. found statistically significant path coefficients from kindergarten measures of NWF to first grade measures of a phonemic decoding strategy and to oral reading fluency. Ritchey found NWF had a positive predictive power of 47.83% and negative predictive power of 97.78% in a sample of 91 kindergarten students. Fien et al. found that measures of NWF in winter-of-kindergarten explained 53% of the variance on reading tests administered in the spring-of-kindergarten. Likewise, Rouse and Fantuzzo (2006) found that the predictive validity of NWF was .50 and .63 on the TerraNova (CTB/McGraw Hill, 2000) and Developmental Reading Assessment (Beaver, 1997), respectively, among a sample of 330 kindergarten students.

To date, investigations of the predictive validity of PSF and NWF have primarily been descriptive measurement studies involving the general kindergarten population and a full range of learners. Consequently, less is known about the ability of PSF and NWF to predict reading outcomes among students who participate in reading intervention in kindergarten.

Curriculum-Embedded Mastery Checks

Curriculum-embedded assessments differ from DIBELS in that they directly measure mastery of short-term, specifically taught skills (Shapiro, 1990). Curriculum-embedded mastery checks (hereafter referred to as mastery checks) are typically administered at specific points in the curriculum, usually the point at which one skill set is expected to be mastered and another is introduced. Instructional decisions are based on the mastery of skills within blocks of instruction, which progress sequentially (Fuchs, 2004). Teachers may compose their own tests covering the curriculum, but many published programs also include measures to assess student progress. In this study, the mastery check assessments corresponded to designated sections of the curriculum and measured phonemic and alphabetic/decoding skills that increased in complexity over the intervention period.

Research comparing the predictive utility of CBMs to mastery checks is lacking. One of the few investigations in this area, conducted by Olinghouse, Lambert, and Compton (2006), involved 40 students in grades three through five. The authors compared the predictive validity of an Oral Reading Fluency (ORF) measure to a measure tied directly to the Phonological and Strategy Training Program, which they described as an *intervention aligned word list*. Findings indicated that each type of measure accounted for unique and differential variance depending on the outcome measure. Specifically, the mastery check explained variance on word reading, decoding, and passage accuracy outcomes, whereas the ORF measure accounted for variance on passage reading fluency. The authors underscored a need for further research on measures from other interventions. Outside of the Olinghouse et al. study, we are not aware of any other direct comparison between mastery checks and CBMs.

The purpose of the present study was to examine the comparative predictive validity of two forms of progress-monitoring measures. We were interested in the predictive validity of individual and combinations of measures that assessed two important constructs in kindergarten—phonemic awareness and alphabetic/decoding. Specifically, we compared the predictive validity of DIBELS specific skill curriculum-based measures (PSF and NWF) and program-specific mastery checks with a sample of kindergarten students who were receiving Tier 2 intervention. The predictive validity of both forms of measures was evaluated at two time points: in December/January (Dec/Jan) and February/March (Feb/Mar).

This study extends the extant research in three important ways. First, it adds to the literature by comparing two prominent forms of measures and their ability to predict end-of- kindergarten outcomes: DIBELS and mastery checks. These measures were examined both individually and in combination. Few studies have investigated the predictive validity of mastery checks, and even fewer have compared the predictive validity of these measures to more general CBMs such as DIBELS (Olinghouse et al., 2006).

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Second, the study explored the variance explained by measures of phonemic awareness, as well as alphabetic/decoding performance, and investigated whether the predictive strength of these fundamental skill domains changed over the course of kindergarten. Contrasts included within-measure comparisons (e.g., DIBELS PSF to NWF, and curriculum-embedded phonemic to alphabetic), as well as between-measures comparisons (DIBELS PSF to curriculum-embedded phonemic, and DIBELS NWF to curriculum-embedded alphabetic).

Finally, the study investigated the predictive validity of both measurement types within an intervention context involving students with identified risks of reading difficulty. To date, a majority of predictive validity studies with kindergarten participants have focused on the general population. Based on our review of the literature, no study has compared the influence of PSF/NWF versus mastery checks on kindergarten outcomes for students receiving a supplemental reading intervention.

The following research questions were addressed:

1. What is the predictive validity of individual DIBELS and curriculum-embedded mastery checks of phonemic and alphabetic/decoding performance in Dec/Jan and Feb/Mar of kindergarten?
2. What combinations of DIBELS and curriculum-embedded phonemic and alphabetic/decoding measures have the strongest predictive validity in Dec/Jan and Feb/Mar?

Method

The study was part of a larger investigation of the effects of a Tier 2 intervention on kindergarten reading achievement (Coyne et al., 2009). Kindergarten children identified to be at greatest risk for reading problems were randomly assigned to one of two intervention groups: modified intervention condition and standard implementation condition. Both groups used the Early Reading Intervention (ERI; Pearson/Scott Foresman, 2004). In one group, the program was implemented using standard implementation procedures provided by the curriculum; in the other group, implementation was adjusted in response to students' performance on the mastery checks. In both conditions, children received explicit, systematic, early reading intervention in small groups for 30 minutes daily for 21 weeks. The primary difference in the conditions was that at 4-week intervals, interventionists who taught in the responsive condition met with the researchers to adjust instruction based on student performance on the mastery checks. The original study focused on the differential effects of intervention. The focus of the current study was the comparative validity of progress-monitoring measures collected during the intervention.

Setting and Participants

SETTING AND INTERVENTIONISTS

The study took place in five elementary schools in eastern Connecticut and in two schools in south-central Texas. A criterion for participation was that the school typically provided some type of supplemental reading instruction for kindergarten children with identified reading risk. With the exception of one school in Connecticut, all schools received Title I funding and had high percentages of children who qualified for free or reduced-cost lunch services. Participating schools' enrollment ranged from 287 to 739 students in Connecticut and from 279 to 889 in Texas. In schools that received Title I funding, the percentage of students qualifying for free and reduced-cost lunch ranged from 69.6% to 82.3% in Connecticut schools and from 81% to 82% in Texas schools. In all, the study involved 26 (14 in Connecticut and 12 in Texas) instructional groups of three to five children each. The interventionists consisted of nine certified kindergarten teachers and four paraprofessionals.

STUDENTS

Participating students were identified through a combination of school and research team screening procedures. First, schools used their existing literacy screening data to identify children who were at risk of reading difficulty. Next, researchers administered the letter naming fluency (LNF) subtest of the DIBELS (Good & Kaminski, 2002) and sound matching subtest (SM) of the Comprehensive Test of Phonological Processing (CTOPP; Wagner,

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Torgesen, & Rashotte, 1999). Students who scored at/below the 36th percentile on the LNF and at/below the 37th percentile on the CTOPP-SM were administered the CTOPP rapid object naming subtest (RON) and the letter identification (Letter ID) subtest from the Woodcock Reading Mastery Test—Revised/Normative Update (WRMT-R-NU; Woodcock, 1987, 1998). Students obtaining a standard score of 7 or less on the CTOPP-RON *or* a standard score of 80 or less on Letter ID were eligible to participate in the study and were administered the remaining pretests. Additional criteria for inclusion consisted of parental consent (obtained before researcher screening) and student participation in all assessments (i.e., progress monitoring and posttesting).

In all, 64 students met the criteria for inclusion in this study. Their average age at the beginning of the study was 5.36 years ($SD = .38$), and 18 (28.1%) were female. The sample of student participants was 53.1% Hispanic or Latino, 25.0% white, 20.3% black or African American, and 1.6% Alaskan native or American Indian. Eleven (19.3%) of the students were English language learners, and five received special education services under the category of speech/language impairment. Of the 64 students in this study, 53 met the selection criteria for the Dec/Jan analyses and 57 for the Feb/Mar analyses.

Progress-Monitoring Measures

DIBELS PSF AND NWF

Student progress was measured using the DIBELS PSF and NWF subtests. Alternate forms for each measure were administered throughout the year at 4-week intervals. Each measure is standardized and measures a child's fluency with a given task during a 1-minute period.

The PSF subtest measures a child's ability to orally segment the initial sound in words. Students are asked to reproduce individual letter sounds from words presented orally by the examiner. Each word contains three to four phonemes, and the student has 1 minute to identify as many phonemes as possible. Alternate-form reliability for the PSF subtest as reported in technical manuals is .88 at 2-week intervals (Good et al., 2004).

The NWF subtest measures a child's ability to decode nonwords. Students are presented with either vowel-consonant (VC)

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

or vowel-consonant-vowel (VCV) nonsense words and asked to produce either the individual letter sounds or the total blended word. Alternate-form reliability with a 1-month interval between administrations is .83 (Good et al., 2004).

CURRICULUM-EMBEDDED MASTERY CHECKS

The intervention curriculum used in this study has four curriculum-embedded assessments that correspond to the four parts of the program. Each assessment is administered once a specified number of lessons have been taught. The first, second, third, and fourth mastery checks are administered after lessons 42, 72, 96, and 126, respectively, and reflect the developmental sequence of reading and reading-related skills taught in the ERI program. More like a curriculum-based assessment, ERI mastery checks measure a student's mastery of the content taught in the preceding set of 30 to 40 lessons. Because students progressed through the program at different rates, the first two mastery checks were the only ones that all students completed. Thus, only scores from the first two mastery checks were used in this study.

Each mastery check contains a combination of subtests that evaluate alphabetic, phonemic, reading, and spelling tasks. For the purpose of this study, we referred to all tasks that involved reading or writing printed letters or words as alphabetic. Tasks that required students to isolate, blend, or segment words in the absence of printed text were classified as phonemic. Both a phonemic and alphabetic composite score can be derived from each assessment. Table 1 summarizes the composition of the ERI mastery checks curriculum-embedded assessments, highlighting the phonemic and alphabetic subtests. DIBELS PSF and NWF subtests are fluency-based; each focuses on one primary task and is completed within a 1-minute period. NWF has a total of 50 items and PSF varies around 80 items. The ERI mastery checks are untimed and include multiple indicators of progress by alphabetic and phonemic construct. Their administration time ranges, on average, from 2 to 5 minutes. At the Dec/Jan time point, the ERI phonemic mastery check has 22 items and the ERI Alphabetic mastery check is composed of 33 items. ERI Phonemic and ERI Alphabetic mastery checks have 23 and 15 items respectively at the Feb/Mar time point.

TABLE 1 Alphabetic and Phonemic Subtests Used in Corresponding ERI Mastery Checks

	ERI Mastery Check			
Subtests	1	2	3	4
Alphabetic				
Letter names and sounds	√	√	√	√
Spelling words with letter tiles[1]	√	√	√	√
Beginning word reading test/regular words	—	—	√	√
Irregular words	—	—	—	√
Phonemic				
First and last sounds in words (viewing picture displays)[2]	√	√	√	√
Whole word segmentation (no picture display)	—	√	√	√

[1]Assessments 1 and 2 probe only initial and ending sounds, whereas 3 and 4 probe all sounds.

[2]Assessments 1 and 2 probe initial and ending sounds, whereas 3 and 4 probe only initial sounds.

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Two alphabetic subtests are used in ERI mastery checks 1 and 2. The Letter Names and Sounds subtest requires a child to name printed letters and say their corresponding letter sounds. In the Spelling Words with Letter Tiles subtest, the complexity of spelling requirements increases across mastery checks. During the first two mastery checks, this subtest probes the first and last sounds of words. A child is provided with seven to eight unique letter tiles, and the examiner presents a page with six pictures and three empty boxes below each picture. After the examiner points to a picture and reads it aloud, the student is asked to put the letter tiles for the first and last sounds into their corresponding (i.e., first and third) boxes.

Two phonemic subtests are also used in ERI mastery checks 1 and 2. The First and Last Sounds in Words subtest measures a child's ability to produce beginning and ending sounds in words that are presented orally with an accompanying picture. The other phonemic subtest, Whole Word Segmentation, is incorporated in ERI mastery checks 2 to 4. For this subtest, the child is presented oral CVC and VC words with no accompanying pictures and is then asked to segment each word into its respective phonemes.

Outcome Measure

At the end of the intervention, the Word Attack (WA) and Word Identification (WI) subtests of the WRMT-R-NU were administered. Using standardized scores, individual student scores on these subtests were combined to create a Basic Skills Cluster (BSC) composite score, according to publisher directions. The Word Identification subtest contains 106 items and measures the ability to read single words. As the examinee proceeds, the words become increasingly more difficult and consist of words that are used less frequently in written English. Items are scored correctly if the words are read with the correct pronunciation. The basal is reached when six items in a row are correct, and the ceiling is reached when six items in a row are incorrect. This basal and ceiling rule is consistent for the Word Identification and Word Attack subtests. The median split-half reliability is .97; concurrent validity is not reported for kindergarten.

For the WA subtest, the examinee reads pseudowords. This requires the examinee to apply word-reading skills instead of relying on sight word recognition. The subtest consists of 45 progressively longer and more difficult words. Correct scores are given for whole-word pronunciation or specific phonemes. Concurrent validity is not reported for kindergarten, and the median split-half reliability is .87. WA is not fluency based, and the pseudowords become progressively more difficult, whereas NWF is based on fluency and uses words designed to be of comparable difficulty.

Assessment Procedures

Assessments (i.e., pretests, progress-monitoring measures, posttests) were conducted with individual students who were seated next to an examiner at a desk or table in quiet environments outside of their classrooms. All assessments were administered by carefully trained graduate students and members of the research team. Assessors received a minimum of two 4-hour training sessions consisting of a review of general assessment procedures, modeling of the specific test protocols, paired practice, and

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

supervised independent practice of each test. Each data collector met a 100% accuracy criterion in both administering and scoring measures. After data collection, all test protocols were independently scored by at least two trained individuals.

ERI mastery checks were administered according to when lessons were completed (i.e., after lessons 42 and 72), as opposed to regular 4-week intervals. Further, the intervention groups progressed at different rates depending on mastery, so children took specific mastery checks at different points in time. Because all students completed mastery checks 1 and 2, these were used in our analyses; their modal dates of administration were December 9 and February 10, respectively.

The DIBELS measures were administered every fourth week starting the 10th week of school (which corresponded with the fourth complete week of intervention). The modal dates for the DIBELS assessments were December 2, January 16, February 10, and March 11. Comparing mastery checks and DIBELS measures at the same time point required that measures be matched separately for each child. To accomplish this, the assessment date for an individual child's mastery check was matched to that child's closest DIBELS administration. In all instances, the most proximal DIBELS measurement was within 2 school weeks before or after the mastery-check administration.

Analyses

To answer the research questions for this study, we examined mastery checks and DIBELS uniquely as well as in combination and tested the statistical significance of differences for each of the two time periods. The first and second ERI mastery checks (administered in Dec/Jan and Feb/Mar, respectively) and matched DIBELS measures (i.e., given within two weeks of the corresponding ERI mastery check) were used as the predictors.

To compare the predictive validity of the ERI mastery checks and the DIBELS measures at both time points, two separate sets of models were estimated. The first set of models regressed the BSC scores on the Dec/Jan ERI mastery-check measures and corresponding DIBELS measures, and the second set used the Feb/Mar mastery checks and corresponding DIBELS measures. Within each set of models, all possible subsets (i.e., phonemic and alphabetic subtests of ERI mastery checks and PSF and NWF from DIBELS) were entered in separate models to determine which predictor(s) accounted for the most variance in BSC scores.

All predictors were entered in a single model; then, individually, they were subtracted from the model and added back in to see how much the multiple R^2 changed due to any one variable's unique contribution. This analysis provides the portion of variance explained by any one unique predictor above and beyond the other predictors, because it controls for the common variance between all other predictors besides the one of interest. This approach was taken because, in practice, it is possible that either the ERI mastery checks or the DIBELS measures would be used alone. Therefore, we deemed it valuable to know how much variance on BSC outcomes either type of measure could account for, independent of the other predictive measures.

Results

Table 2 summarizes the correlations between the predictor variables and the standard scores of the outcome measure (i.e., WRMT-R-NU) for each assessment period. The values below the diagonal are correlations for measures administered during the Dec/Jan period ($n = 53$), and those above the diagonal are those for the Feb/Mar period ($n = 57$). At either time point, we used only those students who had scores on all measures, which

TABLE 2 Correlations Among Study Variables

Measure	1	2	3	4	5
1. ERI Phonemic		.725	.362	.738	.629
2. ERI Alphabetic	.716		.538	.577	.832
3. DIBELS NWF	.535	.577		.364	.563
4. DIBELS PSF	.510	.615	.338		.588
5. WRMT-R BSC	.751	.615	.564	.428	

Note. Values below the diagonal are correlations involving variables measured at the first ERI part-test (Dec/Jan) and matched DIBELS measures ($n = 53$). Values above the diagonal are correlations involving variables measured at the second ERI part-test (Feb/Mar) and matched DIBELS measures ($n = 57$). NWF = Nonsense Word Fluency; PSF = DIBELS Phonemic Segmentation Fluency; BSC = Basic Skills Cluster standard score.

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

resulted in different sample sizes for the Dec/Jan and Feb/Mar measures.

In the Dec/Jan measurement period, correlations ranged from .338 (PSF with NWF) to .751 (ERI Phonemic with BSC). Comparing the ERI Phonemic measure with PSF resulted in a correlation of .510. The correlation between NWF and ERI Alphabetic was .577. For the Feb/Mar measurements, there was a similar pattern for the NWF and ERI Alphabetic measures ($r = .538$). A stronger correlation emerged for the phonemic measures (PSF with ERI Phonemic) during this later time period ($r = .738$). The range for correlations in Feb/Mar was .364 (NWF with PSF) to .832 (ERI Alphabetic with BSC). All values were statistically significant ($p < .05$, two tailed).

We compared R^2 values for the two intervention groups prior to combining them for analysis. The groups were compared on each predictor and combination of predictors for both time periods, for a total of 12 comparisons. Of the 12 comparisons, only 2 were statistically significant, both in the Feb/Mar measurement period. For the modified intervention group, PSF and the combined DIBELS measures were better predictors than they were for the standard implementation group. Correlations between intervention groups were moderate to strong on the majority of measures.

December/January Measurements

INDIVIDUAL PREDICTORS

Table 3 shows the amount of variance the Dec/Jan predictor measures explained on end-of-kindergarten BSC performance. The first column lists the predictor(s) entered in the model. The R^2 column indicates the amount of variance the predictor(s) explained on the BSC. Finally, the columns to the right of the R^2 column report how much additional variance was explained when the variable listed at the top of the column was added to the predictor(s) in the far-left column.

In the Dec/Jan model examining the contributions of individual measures, ERI phonemic was the strongest predictor, explaining 56.4% of the variance on BSC outcomes; the next best single predictor was ERI Alphabetic, which explained 37.8%. As

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

TABLE 3 Unique Contribution of Individual and Combinations of Predictors to End-of-Year Basic Skills Cluster in Dec/Jan

Unique Explained Variance		Added Explained Variance			
		DIBELS		ERI	
Combinations of Predictor Variables	R^2	NWF	PSF	Phonemic	Alphabetic
NWF	.318	—	.064	.283	.125
PSF	.183	.199	—	.384	.199
ERI Phonemic	.564	.037	.003	—	.012
ERI Alphabetic	.378	.066	.004	.198	—
NWF + PSF	.382	—	—	.221	.067
NWF + ERI Phonemic	.601	—	.000	—	.001
NWF + ERI Alphabetic	.444	—	.005	.160	—
PSF + ERI Phonemic	.567	.035	—	—	.009
PSF + ERI Alphabetic	.382	.067	—	.195	—
ERI Phonemic + ERI Alphabetic	.577	.027	.000	—	—
NWF + PSF + ERI Phonemic	.603	—	—	—	.001
NWF + PSF + ERI Alphabetic	.449	—	—	.155	—
NWF + ERI Phonemic + ERI Alphabetic	.604	—	.000	—	—
PSF + ERI Phonemic + ERI Alphabetic	.577	.027	—	—	—
All	.604				

Note. $n = 53$; NWF = Nonsense Word Fluency; PSF = Phoneme Segmentation Fluency.

sole predictors, the DIBELS NWF subtest explained 31.8% of the BSC variance and the PSF explained 18.3%.

COMBINATIONS OF PREDICTOR VARIABLES

All possible combinations were examined (see Table 3) to compare the amount of unique variance an individual predictor accounted for when added to other predictors. ERI Phonemic accounted for the largest amount of unique variance explained

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

when added to all other predictors, in both two- and three-measure combinations. The unique variance that ERI Phonemic explained above and beyond other predictors ranged from 15.5% (when added to NWF, PSF, and ERI Alphabetic) to 38.4% (when added to PSF).

The greatest amount of variance on BSC outcomes was explained by the model containing all four predictor variables *and* the model containing NWF, ERI Phonemic, and ERI Alphabetic. Each of these Dec/Jan models explained 60.4% of the variance in end-of-kindergarten reading outcomes. PSF explained no additional variance when combined with the other three predictors. The model containing only NWF and ERI Phonemic was almost as strong; it explained 60.1% of the variance in outcomes.

R^2 DIFFERENCES BETWEEN CURRICULUM-EMBEDDED AND CURRICULUM-BASED MEASURES

We examined whether the difference in R^2 values for mastery checks versus DIBELS were statistically significant by creating asymptotic confidence intervals around their difference scores. The formula used, which was based on the work of Alf and Graf (1999), tested the null hypothesis that multiple R^2 values of the ERI measures as predictors model minus the DIBELS measures as predictors model equal zero. If the confidence interval contains the value zero (which indicates the possibility that the R^2 values are identical), the R^2 values are not assumed to be statistically significantly different from one another. Ninety-five percent intervals were used to create the confidence intervals, as this is analogous to using the commonly accepted p value of .05.

Table 4 contains the comparisons of R^2 values between the mastery checks and DIBELS measurements for each time period. Comparisons were tested between phonemic (ERI Phonemic vs. PSF), alphabetic (ERI Alphabetic vs. NWF), and the combined predictor models (Combination of ERI Alphabetic and Phonemic vs. DIBELS). At the Dec/Jan assessment point, ERI Phonemic explained statistically significantly more variance than PSF on the BSC. The difference in R^2 values was .381. However, the difference between the alphabetic measures was not statistically significant during this time period. The combination of mastery checks was a stronger predictor than the combined DIBELS measures,

TABLE 4 Asymptotic 95% Confidence Intervals for R^2 Comparisons

	Basic Skills Cluster Percentile Scores			
			95% C.I.	
Comparisons	R^2 diff.	*SE*	Lower	Upper
Dec/Jan[a] measurements				
ERI Phonemic versus PSF	.381*	.105	.175	.587
ERI Alphabetic versus NWF	.06	.113	−.161	.281
ERI Phonemic + ERI Alphabetic versus NWF + PSF	.195*	.099	.0002	.390
Feb/Mar[b] measurements				
ERI Phonemic versus PSF	.051	.105	−.154	.256
ERI Alphabetic versus NWF	.375*	.103	.172	.578
ERI Phonemic + ERI Alphabetic versus NWF + PSF	.208*	.091	.030	.386

[a]Dec/Jan: $n = 53$; [b]Feb/Mar: $n = 57$.
$^{*}p < .05$.

and the R^2 difference score (.195) between the two was statistically significant.

February/March Measurements

INDIVIDUAL PREDICTORS

We also examined the contributions of individual measures for the Feb/Mar assessment period. Table 5 shows the amount of variance the predictors explained on end-of-kindergarten BSC performance. For individual measures, ERI Alphabetic emerged as the strongest individual predictor, explaining 69.2% of the variance. ERI Phonemic explained the next greatest amount of variance, at 39.6%. Finally, the DIBELS PSF subtest explained 34.5%, followed by NWF, which accounted for 31.7% of the outcome variance.

COMBINATIONS OF PREDICTOR VARIABLES

In addition to being the strongest individual predictor, ERI Alphabetic explained the greatest amount of unique variance when added to all other possible combinations of predictors (see Table 5). The amount of unique additional variance explained

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

TABLE 5 Unique Contribution of Individual and Combinations of Predictors to End-of-Year Basic Skills Cluster in Feb/Mar

Unique Explained Variance		Added Explained Variance			
		DIBELS		ERI	
Combinations of Predictor Variables	R^2	NWF	PSF	Phonemic	Alphabetic
NWF	.317	—	.169	.208	.394
PSF	.345	.140	—	.084	.364
ERI Phonemic	.396	.129	.033	—	.298
ERI Alphabetic	.692	.019	.017	.001	—
NWF + PSF	.485	—	—	.056	.240
NWF + ERI Phonemic	.525	—	.017	—	.188
NWF + ERI Alphabetic	.710	—	.015	.002	—
PSF + ERI Phonemic	.429	.112	—	—	.282
PSF + ERI Alphabetic	.709	.016	—	.002	—
ERI phonemic + ERI alphabetic	.693	.019	.018	—	—
NWF + PSF + ERI Phonemic	.541	—	—	—	.185
NWF + PSF + ERI Alphabetic	.725	—	—	.001	—
NWF + ERI Phonemic + ERI Alphabetic	.712	—	.014	—	—
PSF + ERI Phonemic + ERI Alphabetic	.711	.015	—	—	—
All	.726				

Note. $n = 57$; NWF = Nonsense Word Fluency; PSF = Phoneme Segmentation Fluency.

by ERI Alphabetic above and beyond all other predictors ranged from 18.5% (when added to NWF, PSF, and ERI Phonemic) to 39.4% (when added to NWF). During the Feb/Mar time period, the most variance explained was 72.6% by a model containing all predictors. The combination of NWF and ERI Alphabetic ($R^2 =$ .710) explained nearly as much variance as a model with all of the predictors combined ($R^2 = .726$).

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

R^2 DIFFERENCES BETWEEN CURRICULUM-EMBEDDED AND CURRICULUM-BASED MEASURES

We compared the R^2 values between the mastery checks and DIBLES measures collected during Feb/Mar (see Table 4). We applied the same approach used with the Dec/Jan data, creating asymptotic confidence intervals around difference scores (Alf & Graf, 1999). Unlike the earlier assessment period, the difference between the ERI Alphabetic measure and NWF was statistically significant; ERI Alphabetic was a stronger predictor of end-of-kindergarten BSC scores. Conversely, the difference between ERI phonemic and PSF was not statistically significant at this later measurement point.

The combination of mastery checks remained a stronger predictor of end-of-kindergarten performance than the combined DIBELS measures, and the R^2 difference score (.208) between the two was statistically significant. Interestingly, while the R^2 values increased between the Dec/Jan and Feb/Mar assessment periods for both sets of progress-monitoring measures, the difference between the two was very similar (.195 in Dec/Jan vs. .208 in Feb/Mar).

Discussion

Monitoring kindergarteners' response to early reading intervention requires valid measures of student progress that enable teachers to make trustworthy instructional decisions (Gersten et al., 2009; Good et al., 1998). In a Response to Intervention paradigm, predictive validity is of particular importance, as it enables educators to make inferences about a learner's future performance. With respect to early literacy, formative measures that predict phonological awareness and alphabetic skills are particularly relevant (Bishop & League, 2006; Rouse & Fantuzzo, 2006). In practice, many of the measures used to inform instructional decisions in kindergarten have not been critically evaluated for their predictive validity.

In this study we were interested in the comparative predictive validity of mastery checks (i.e., those that are often used in interventions) and DIBELS PSF and NWF, commonly used measures in kindergarten. The purpose of this study was to compare the predictive validity of mastery checks and DIBELS measures that assessed phonemic awareness and alphabetic/decoding skills. We

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

examined and compared the variance explained by individual and combinations of measures at two points in time.

Our first principal finding was that both DIBELS and individual mastery-check measures are highly correlated and predicted significant amounts of end-of-year variance. This finding affirms the predictive validity of both types of measures and their use in kindergarten intervention decision making. Although mastery-check measures explained statistically significantly more variance at each time point, large amounts of variance could be explained by a combination of ERI and DIBELS measures (e.g. 60.1% explained by the NWF and ERI Phonemic combination in Dec/Jan). This finding indicates that curriculum-embedded measures may complement information obtained by DIBELS and aid in evaluating student progress and making data-driven instructional decisions.

Regarding the predictive validity of individual measures, the curriculum-embedded mastery checks were better as individual predictors of kindergarten BSC outcomes. At each of the study's progress monitoring time points, an individual mastery-check measure was the strongest individual predictor and explained over twice the amount of variance as its corresponding DIBELS measure. In Dec/Jan, the ERI Phonemic measure was the strongest individual predictor, while no differences were observed between the alphabetic measures (i.e., ERI Phonemic and NWF). Interestingly, at the Feb/Mar time point the ERI Alphabetic measure replaced ERI Phonemic as the strongest predictor, and the difference between ERI Phonemic and DIBELS Phonemic measures was no longer statistically significant.

Another major finding is the change in predictive strength of Phonemic and Alphabetic measures at each time point, suggesting that the skills that best predict end-of-year outcomes changed during the course of the kindergarten year. Thus, in Dec/Jan, phonemic skills were the best predictors, whereas alphabetic skills were most predictive in Feb/Mar. This pattern reflects the developmental progression of skills over time and the increasing importance and alignment of the alphabetic principle to the criterion reading tasks that were measured at the end of the year.

Findings from the combination analyses likewise revealed that the combination of mastery-check measures (i.e., ERI Alphabetic plus ERI Phonemic) was a better predictor of reading outcomes than PSF and NWF combined. This difference was

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

statistically significant at both time points; the combination of mastery-check measures explained roughly 1.5 times the amount of variance as the combined DIBELS.

Another equally important finding concerns the most parsimonious set of predictive measures. At the Dec/Jan time point, the combination of the ERI phonemic measure and the DIBELS Alphabetic measure (NWF) explained 60.1% of the variance on the BSC, which is nearly as much variance explained by all four together (60.4%). However, no paired combination of predictors (including combinations of DIBELS and mastery checks) explained statistically significantly more variance than the combination of the mastery-check measures.

For the Feb/Mar assessments, the combination of all four predictor variables explained the most variance (72.6%); however, this value was not a statistically significantly greater than the variance explained by the combination of ERI Phonemic and Alphabetic (69.3%). Nor was the variance explained by ERI Phonemic plus ERI Alphabetic statistically different from that of ERI Alphabetic plus NWF (71.0%). As with the earlier assessment period, the combined within-curriculum mastery-check measures (alphabetic plus phonemic) accounted for more variance in end-of-kindergarten BSC performance than the DIBELS combination (NWF and PSF). This finding supports the hypothesis that the mastery-check measures not only explained more variance, but also were more parsimonious in doing so.

Results at both time points indicated that mastery-check measures were strong predictors of kindergarteners' Response to Intervention as measured by the word attack and word identification composite measure. Both uniquely and as a set, the ERI mastery checks were reliable predictors of outcomes. While a combination of mastery check and DIBELS measures explained the most variance at any one time point, this additional predictive power was not statistically significant; the mastery-check measures predicted as much as any combination of measures and statistically significantly more than the DIBELS measures. Although mastery-check measures explained the most amount of variance, both individual and combinations of DIBELS were also found to have predictive utility. This finding has important implications indicating that reliable decisions can be made with either form of measurement.

The final important finding is that mastery-check measures can make strong predictions early. Early intervention is critical, and a major goal of progress monitoring is to collect information that can guide instructional decisions. Having a measure that is capable of explaining nearly 60% of the variance on BSC outcomes by midyear (i.e., Dec/Jan) gives a teacher strong evidence-based data to support educational decisions.

Limitations and Future Research Directions

Results of this study must be considered in light of its limitations. First, the study involved a relatively small number of students, which indicates the results cannot be generalized to other populations of students. Second, our findings are limited to the mastery-check measures of the ERI program (Pearson/Scott Foresman, 2004). Future research should investigate whether mastery-check measures from other programs have similar predictive power. Third, the study only considered progress-monitoring measurements from two points in time. For students in a Tier 2 intervention, a minimum of monthly progress monitoring is recommended. Future research should evaluate the predictive validity of more frequently administered mastery-check and progress-monitoring measurements. The reliability of the ERI mastery checks has yet to be investigated. Research is needed to validate the reliability of the ERI mastery checks and allow for future comparisons against measures with demonstrated technical adequacy. Finally, the use of a single outcome measure is a limitation. Investigations are needed that employ additional outcome measures, some of which are fluency-based, to examine whether formative measures of student progress predict performance beyond the basic skills cluster of the WRMT-R-NU. Further, the individual subtests that compose the BSC (WI and WA) could have been used as separate outcome measures to examine the predictive validity on more narrowly focused and specific skills.

Implications and Conclusion

Overall, the results of this study provide potentially important information regarding the predictive validity of measures used to assess kindergarten response to early reading intervention, both

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

DIBELS and mastery checks. Whereas DIBELS' validity have been investigated in previous studies, there have been few investigations of curriculum-embedded mastery measures. Findings indicate that the ERI Phonemic and ERI Alphabetic measures can be an effective element of an RTI approach. The ability of the mastery measures to explain a majority of the variance on end-of-year outcomes by Dec/Jan adds to their usefulness in aiding teacher decisions early in the intervention process. The sooner outcomes can be accurately projected, the earlier teachers can make informed adjustments (e.g., regrouping of students, reteaching, pace of instruction, scope and sequence of critical skills), thereby increasing the potential for stronger Response to Intervention.

While early measurements of phonemic awareness accounted for the greatest amount of variance on end-of-year outcomes, the level of children's alphabetic skills explained more variance later in the year. This underscores the importance of monitoring more than one skill type and indicates that more attention may be warranted for certain skills depending on the time of year, especially given the sequential and often rapid development of reading skills during the first year of reading instruction.

Acknowledgments

This research was supported in part by the 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 our testers and data-entry staff for their contributions and Kimberly Williams for her support in manuscript preparation. Last, we thank the school districts, administrators, teachers, interventionists, and children for their support and participation in this research.

The intervention used in this study was *The Early Reading Intervention* (Pearson/Scot Foresman, 2004). Because two authors of this manuscript are also coauthors of the intervention program (i.e., Simmons and Coyne), the following steps were implemented to ensure objectivity of findings.

1. All data analyses were conducted by statisticians who had no financial interest with the Early Reading Intervention.
2. An external consultant with no financial affiliation with the Early Reading Intervention program independently reviewed the manuscript to ensure that (a) data analyses were appropriate, accurate, and objective, (b) reported findings and discussion were accurate, and (c) interpretations were consistent with data analysis.

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

References

Alf, E. F., & Graf, R. G. (1999). Asymptotic confidence limits for the difference between two squared multiple correlations: A simplified approach. *Psychological Methods, 4*(1), 70–75. doi:10.1037/1082–1989X.1034.1031.1070

Beaver, J. (1997). *Developmental reading assessment, K–3.* Glenview, IL: Celebration Press.

Bishop, A. G., & League, M. B. (2006). Identifying a multivariate screening model to predict reading difficulties at the onset of kindergarten: A longitudinal analysis. *Learning Disability Quarterly, 29*, 235–252. doi:210.2307/30035552

Burke, M. D., Hagan-Burke, S., Kwok, O., & Parker, R. (2009). Predictive validity of early literacy indicators from the middle of kindergarten to second grade. *Journal of Special Education, 42*(4), 209–226. doi:210.1177/0022466907313347

Coyne, M. D., Simmons, D. C., Kwok, O., Kim, M., Oslund, E. L., Simmons, L. S., et al. (2009). Assessing the impact of reading intervention in kindergarten classrooms: A between-contexts comparison. Manuscript in preparation.

CTB/McGraw Hill. (2000). *TerraNova, the second edition: California achievement tests.* Monterey, CA: CTB/McGraw-Hill LLC.

Fien, H., Baker, S. K., Smolkowski, K., Mercier Smith, J. L., Kame'enui, E. J., & Beck, C. T. (2008). Using nonsense word fluency to predict reading proficiency in kindergarten through second grade for English learners and native English speakers. *School Psychology Review, 37*, 391–408.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to Response to Intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*, 93–99. doi:10.1598/RRQ.1541.1591.1594

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*, 188–192.

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., et al. (2009). *Assisting students struggling with reading: Response to Intervention and multi-tier intervention in the primary grades. A practice guide.* (NCEE 2009–4045). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/practice_guides/rti_reading_pg_021809.pdf

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014

Good, R. H. III, & Kaminski, R. A. (1996). Assessment for instructional decisions: Toward a proactive/prevention model of decision-making for early literacy skills. *School Psychology Quarterly, 11*, 326–336. doi:310.1037/h0088938

Good, R. H. III, & Kaminski, R. A. (Eds.). (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Good, R. H. III, Kaminski, R. A., Shinn, M. R., Bratten, J., Shinn, M., Laimon, D., et al. (2004). *Technical adequacy of Dibels: Results of the Early Childhood Research Institute on measuring growth and development.* (Tech. Rep. No. 7). Eugene, OR: University of Oregon.

Good, R. H. III, Simmons, D. C., & Smith, S. B. (1998). Effective academic interventions in the United States: Evaluating and enhancing the acquisition of early reading skills. *The School Psychology Review, 27*, 45–56.

Grigorenko, E. L. (2009). Dynamic assessment and response to intervention: Two sides of one coin. *Journal of Learning Disabilities, 42*, 111–132. doi:110.1177/0022219408326207

Hoffman, A. R., Jenkins, J. E., & Dunlap, S. K. (2009). Using DIBELS: A survey of purposes and practices. *Reading Psychology, 30*, 1–16. doi:10.1080/02702710802274820

Lembke, E. S., McMaster, K. L., & Stecker, P. M. (2010). The prevention science of reading research within a response-to-intervention model. *Psychology in the Schools*, 22–35. doi:10.1002/pits.20449

Mellard, D. F., McKnight, M., & Woods, K. (2009). Response to Intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice, 24*, 186–195. doi:110.1111/j.1540-5826.2009.00292.x

Olinghouse, N. G., Lambert, W., & Compton, D. L. (2006). Monitoring children with reading disabilities' response to phonics intervention: Are there differences between intervention aligned and general skill progress monitoring assessments? *Exceptional Children, 73*, 90–106.

Pearson/Scott Foresman. (2004). *Scott Foresman sidewalks: Early reading intervention.* Glenview, IL: Author.

Ritchey, K. D. (2008). Assessing letter sound knowledge: A comparison of letter sound fluency and nonsense word fluency. *Exceptional Children, 74*, 487–506.

Rouse, H. L., & Fantuzzo, J. W. (2006). Validity of the Dynamic Indicators for Basic Early Literacy Skills as an indicator of early literacy for urban kindergarten children. *School Psychology Review, 35*, 341–355.

Schatschneider, C., Francis, D. J., Carlson, C. D., Fletcher, J. M., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96*, 265–282. doi:210.1037/0022-0663.1096.1032.1265

Schilling, S. G., Carlisle, J. F., Scott, S. E., & Ji, Z. (2007). Are fluency measures accurate predictors of reading achievement? *Elementary School Journal, 107*, 429–448. doi:410.1086/518622

Shapiro, E. S. (1990). An integrated model for curriculum-based assessment. *School Psychology Review, 19*, 331–350.

University of Oregon Center on Teaching and Learning. (2010). DIBELS nonsense word fluency. Retrieved from https://dibels.uoregon.edu/measures/nwf.php

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: Pro-Ed.

Woodcock, R. W. (1987, 1998). *Woodcock Reading Mastery Test—revised/normative update*. Bloomington, MN: Pearson Assessments.

Downloaded by [University of California Santa Cruz] at 14:49 31 October 2014