Published in final edited form as:
J Learn Disabil. 2011 ; 44(4): . doi:10.1177/0022219411407863.

The Use of a Dynamic Screening of Phonological Awareness to Predict Risk for Reading Disabilities in Kindergarten Children

Mindy Sittner Bridges[1] and **Hugh W. Catts**[1]
[1]University of Kansas, Lawrence, KS, USA

Abstract

This study examined the usefulness and predictive validity of a dynamic screening of phonological awareness in two samples of kindergarten children. In one sample (n = 90), the predictive validity of the dynamic assessment was compared to a static version of the same screening measure. In the second sample (n = 96), the dynamic screening measure was compared to a commonly used screening tool, Dynamic Indicators of Basic Early Literacy Skills Initial Sound Fluency. Results showed that the dynamic screening measure uniquely predicted end-of-year reading achievement and outcomes in both samples. These results provide preliminary support for the usefulness of a dynamic screening measure of phonological awareness for kindergarten students.

Recently, response to intervention (RTI) has been proposed as a model for the early identification and prevention of reading disabilities (RD; D. Fuchs & Fuchs, 2005; Haager, Klingner, & Vaughn, 2007). According to this multitiered model, children can be identified as having a learning disability or RD if their response to scientifically based instruction, including targeted intervention, is substantially below that of their peers. A central component of RTI is universal screening. This screening serves to identify children who show inadequate response to classroom instruction. In an RTI framework, all children participate in periodic universal screening, and those who “fail” the universal screening may subsequently be placed in short-term intervention programs (Tier 2). Because of the expense and effort involved in the latter intervention, it is critical for universal screening instruments to measure the most appropriate reading-related skills, in the most accurate manner possible.

Research suggests that an RTI model can be effective as early as kindergarten. Studies show that if children are identified as at risk in kindergarten, intervention can significantly improve their reading outcome (Denton & Mathes, 2003; Simmons et al., 2008; Vellutino, Scanlon, Small, & Fanuele, 2006). To provide early intervention, many school districts have implemented universal screening at the beginning of kindergarten. Universal screening at this time presents a significant challenge for it requires measuring a skill (i.e., reading) that is not yet present. However, screening instruments can assess prereading skills that, if impaired, can signal a potential RD. One such skill is phonological awareness. Numerous studies demonstrate that measures of phonological awareness uniquely predict reading outcomes (Catts, Fey, Zhang, & Tomblin, 2001; O’Connor & Jenkins, 1999). One limitation to using phonological awareness measures for universal screening at the beginning of kindergarten is that many measures designed to be used at that time may be characterized by floor effects. For example, Catts, Petscher, Schatschneider, Bridges, and Mendoza (2009) recently reported that the Initial Sound Fluency subtest of the Dynamic Indicators of Basic

Corresponding Author: Mindy Sittner Bridges, University of Kansas, 3001 Dole Building, 1000 Sunnyside Avenue, Lawrence, KS 66045 msittner@ku.edu.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Early Literacy Skills (DIBELS; Good & Kaminski, 2003) can have a substantial floor effect in kindergarten and that this floor effect can limit its predicative validity. Other common measures of phonological awareness also may have floor effects when administered to kindergarten children (Rathvon, 2004).

The floor effects in measures of phonological awareness likely arise from a large number of children lacking literacy experience at the beginning of kindergarten. Such experience can affect performance on measures of phonological awareness (Castles & Coltheart, 2004; Foy & Mann, 2003). The Initial Sound Fluency subtest of DIBELS and most other measures of phonological awareness are static in nature (see Note 1). That is, they measure a learned product and, in doing so, provide children with little or no feedback (Lidz,1991). As a result, static assessments may not allow children to demonstrate their learning potential. Dynamic assessment, on the other hand, provides such an opportunity. Dynamic assessment refers to a variety of procedures that embed interaction with a child as part of the assessment process. In dynamic assessment, the examiner takes an active role by teaching a task or providing explicit prompts. Success is measured by both a student's level of independent performance as well as a student's assisted performance (i.e., progress). Dynamic assessment takes into account both the process and the product of learning (for a review, see Sternberg & Grigorenko, 1998). Because dynamic assessment may provide information about a child's ability to respond to instruction that is not obtainable through more traditional assessment sources, it may play a useful role in RTI. Specifically, the use of dynamic assessment may allow us to differentiate between students who initially fall at the lower end of the distribution because of limited literacy experience or poor instruction and those who truly are at risk for RD (L. S. Fuchs et al., 2008). In other words, dynamic assessment might assist in limiting a floor effect and improving the predicative validity of a screening instrument. Last, phonological awareness also has the advantage of being a learnable skill and one that could respond well to the instruction provided in dynamic assessment.

Several investigators have examined the dynamic assessment of phonological awareness (O'Connor & Jenkins, 1999; Spector, 1992). Spector (1992) administered a dynamic assessment of phonological awareness and several static measures of phonological awareness to a small group of kindergarten children. Results showed that the dynamic assessment accounted for significant variance in end-of- year reading achievement over and above the static measures. O'Connor and Jenkins (1999) also investigated dynamic assessment for the purpose of early identification. In a series of studies, they found that a dynamic phoneme segmentation task significantly improved the predictive validity of a first-grade screening battery.

The present study is an initial report of a line of research designed to examine the validity and usefulness of a dynamic screening of phonological awareness for the early identification of kindergarten children at risk for RD. Data from two samples of kindergarten children are reported. In one sample, we administered a dynamic screening of phonological awareness and a static version of the same test. In a second sample, we administered a dynamic screening of phonological awareness and a commonly used screening measure of phonological awareness. For both samples, measures of phonological awareness were given at the beginning of the year and performance on these measures was used to predict end-of-year reading outcomes. We were particularly interested in the extent to which the dynamic screening instrument predicted reading achievement or RD over and above a static version of the same test or a commonly used screening measure of phonological awareness.

[1]The term dynamic in the Dynamic Indicators of Basic Early Literacy Skills refers to the fact that the assessments in the battery have multiple forms that can be administered repeatedly over time. This usage differs from the way the term is used in this study and in other lines of research concerning dynamic assessment (Sternberg & Grigorenko, 1998).

Method

Participants

Two samples of students participated in this study. Sample 1 (n = 90) included all kindergarten children (except those meeting exclusion criteria below) from three small schools in Iowa. These three schools included approximately 95% Caucasian students, with between 15% and 20% of the students receiving free or reduced-price lunch. Sample 2 (n = 96) included students from a large school district in Kansas. This school district was somewhat diverse in terms of ethnicity (approximately 63% Caucasian, 11% African American, 6% Hispanic, 7% American Indian or Alaskan Native, 6% Asian or Pacific Islander, and 7% multiracial). Students in Sample 2 were also participants in a larger investigation of early identification of RD. Approximately half of the students in this sample (n = 47) were selected to be "at risk" for RD as designated by scoring in the "At-Risk" or "Some-Risk" range on both the Initial Sound Fluency and the Letter Name Fluency subtests of DIBELS, which were administered by school personnel at the beginning of kindergarten. The remaining students (n = 49) were randomly selected from among classmates of at-risk children and failed to meet the above risk status. The kindergarten language arts curriculum for all schools in both samples included daily components of phonological awareness, phonics, and/or word recognition instruction. In addition, in all schools, students deemed to be at risk by school personnel typically received some sort of additional small group or individual instruction, with an emphasis on alphabetic knowledge.

Exclusionary criteria in both samples included the following: a designation of "nonverbal" on an individualized education program (IEP), limited English proficiency as indexed by a score of 1 or 2 on the Oral Language subtest of the Pre-LAS (Duncan & DeAvila, 1998), or significant health or cognitive impairment (e.g., mental retardation, hearing impaired, autism). In addition, a limited number of students were excluded from Sample 1 by their teachers because of one of the following: (a) limited English proficiency levels per teacher report, (b) behavior concerns, or (c) an IEP designating a cognitive delay or severe speech impairment.

Measures

Students in Sample 1 were administered the Static Deletion Task and the Dynamic Screening of Phonological Awareness (Bridges & Catts, 2008). Students in Sample 2 were administered the Dynamic Screening of Phonological Awareness and the Initial Sound Fluency subtest of DIBELS. In addition, the Word Identification and the Word Attack subtests from the Woodcock Reading Mastery Tests-Revised/NU (WRMT-R/NU; Woodcock, 1998) were administered to students in both samples.

Static Deletion Task (SDT)—This task required a student to say a word after deleting a syllable or phoneme specified by the experimenter (e.g., "Say cowboy without cow."). The SDT consisted of four sets of words; each set increased in complexity. The first set (8 words) required that the first syllable from a two-syllable word (either a compound word or a word with a prefix) be deleted. The second set (3 words) required the deletion of a syllable from a noncompound, nonprefix two-syllable word. The third set (5 words) involved the deletion of the initial consonant from a single-syllable CVC word. Finally, the fourth set (4 words) involved the deletion of the first consonant from a singlesyllable CCVC word. There were 20 items, and each item was worth 1 point. Administration of this assessment was discontinued when students failed to provide the correct response on five consecutive items.

Dynamic Screening of Phonological Awareness (DSPA)—DSPA (Bridges & Catts, 2008) consisted of the exact same items included in the SDT, but it was administered in a dynamic manner, with graduated prompts provided by the examiner. Because the items were the same as those for the SDT, the DSPA was administered immediately after the SDT in Sample 1. In addition, to reduce redundancy the examiner started the DSPA on the first item that the child had missed on the SDT. Items prior to the starting item were given full credit. The protocol for the prompts on the DSPA was the following. When students gave a correct response, the response was acknowledged as such. Alternatively, when students gave an incorrect response to a target item, the examiner provided a series of prompts until the item was answered correctly or the answer was given. The prompts included the following: (a) the examiner repeated the initial example and then repeated the question ("Remember the word sailboat without sail is boat. Now say ___without."); (b) the examiner repeated the test item, pausing after the first syllable or phoneme and stressing the remaining portion of the word; (c) the examiner repeated the question and used a picture (e.g., dog) or a blank square to denote the first syllable or phoneme, paused, and used a picture (e.g., house) or blank square to denote the remaining portion of the word. Each item was assigned 0 to 4 points, with a higher score on an item indicating the need for fewer examiner prompts. Therefore, the procedure for scoring reflected the degree of independence that a student achieved in performing the task. There were 20 items, and the highest score possible was 80 points. Administration of the DSPA was discontinued if the child received a score of 0 on five consecutive test items. An earlier version of the DSPA was found to have an interrater reliability of .98 and a test-retest reliability of .86 (Bridges, 2009). Although administration time varied because of the number of prompts provided to an individual student, the average time of administration was 8 to 10 minutes.

Initial Sound Fluency (ISF)—This subtest of the DIBELS is a measure of a student's ability to recognize and produce the initial sound in an orally presented word. This subtest is commonly used as a universal screening tool in kindergarten to identify children who are not showing adequate progress in the general curriculum (Catts et al., 2009). For this measure, the examiner showed the student four pictures, named each picture, and asked students to identify (i.e., either point or say) the picture that began with the sound produced by the examiner. Students were also asked to pro-duce the beginning sounds of some words presented orally by the examiner. The subtest was discontinued if a student answered the first five items incorrectly. The amount of time taken to identify or produce the correct sounds was used to convert the raw score into the number of initial sounds correct per minute. Alternate-form reliability of this measure is .72 (Good et al., 2004).

WRMT-R/NU Word Identification—The Word Identification subtest (Woodcock, 1998) was a measure of untimed real-word reading in isolation. Students were required to read a list of words that gradually decreased in frequency of occurrence. The WRMT-R/NU manual reports the split-half reliability is .98.

WRMT-R/NU Word Attack—The Word Attack subtest (Woodcock, 1998) was a measure of untimed nonsense- word reading in isolation. This measure assessed a student's ability to apply grapheme-phoneme correspondence rules to pronounce unfamiliar printed words (i.e., pronounceable nonwords varying in complexity). The manual reports the split-half reliability is .94.

Procedures

All testing sessions were administered in the schools by trained examiners. The SDT and DSPA were administered to Sample 1 and the ISF and DSPA to Sample 2 in mid- to late September. The Word Identification and Word Attack subtests were given at the end of April or beginning of May of the following year (i.e., the end of kindergarten).

Results

Descriptive statistics for measures administered to Sample 1 are displayed in Table 1. Of particular relevance is the skewness statistic. The skewness statistic, an index of the symmetry of the distribution, can provide information regarding the presence of floor effects in the data. The farther the absolute value of the skewness statistic is from zero, the greater the skew of the distribution. Positive numbers, typically seen in a distribution with a floor effect, are indicative of a mean that is closer to the lower end of the distribution and an asymmetric tail extending toward the higher end of the scale. These data showed that the distribution of scores from the DSPA had a skewness index near zero (i.e., −.062), whereas the skewness statistic associated with the SDT measure (.990) suggested that a floor effect was present. Table 1 also shows a positive skewness was present in the end-of-year reading achievement data. Although this is not optimal for an outcome measure, it is likely an accurate reflection of end-of-year reading achievement in many kindergarten classrooms.

The Predictive Validity of DSPA as Compared to the SDT

One of the aims of this study was to determine if a dynamic assessment of phonological awareness added to the prediction of reading outcomes over and above a similar static assessment. To this end, hierarchical linear regression was employed to investigate if the DSPA explained variance over and above the SDT in end-of-year performance on the Word Identification and Word Attack subtests. In these analyses, the variables were entered in a sequential manner, with the SDT entered first and the DSPA entered next. Results indicated that the SDT accounted for 28% of the variance in Word Identification and 23% of the variance in Word Attack. When entered as the second variable, the DSPA accounted for a significant amount of unique variance in Word Identification (4%) and Word Attack (9%).

Because the dynamic assessment was designed as a screening assessment, it was also important to evaluate its accuracy relative to the SDT in the identification of children at risk for RD. To do this, we first classified children as good or poor readers based on their reading achievement at the end of kindergarten. Logistic regression was then employed to examine the classification accuracy of the DSPA relative to the SDT. In these analyses, students were identified as a poor reader if their score on the reading achievement outcome measure (i.e., Word Identification or Word Attack) was at or below the 25th percentile. Children scoring above the 25th percentile were considered good readers. Furthermore, it was necessary to use sample statistics to identify the 25th percentile rather than the normative data provided in the WRMT-R/NU manual. This manual was revised more than 10 years ago prior to the influence of the No Child Left Behind Act on early reading achievement. Although this legislation is formally confined to the upper elementary grades, instructional requirements associated with the new standards have affected kindergarten curricula. As a result, many kindergarten students are reading at levels well beyond those reported in the normative data provided by the WRMT-R/NU. Therefore, cutoff scores for outcome measures were based on sample characteristics. The cutoff score was 5 for Word Identification and 2 for Word Attack. These procedures identified 28 and 27 students from Sample 1 as poor readers based on the Word Identification and Word Attack subtests, respectively.

In the logistic analyses, the SDT and the DSPA were first entered separately as single predictors and then were entered together in a sequential manner, the SDT followed by the DSPA. As seen in Table 2, the DSPA was a significant single predictor of both measures of end-of-year reading achievement. In addition, it proved to be the better predictor in sequential models. In each of these models, the DSPA was a significant predictor and the SDT was nonsignificant. Classification information obtained from these logistic regressions was also used to plot a receiver operating characteristic (ROC) curve for each predictor as well as the combination of predictors. Area under the ROC curve (AUC) provides an estimate of the overall predictability of a screening measure. Generally, values greater than .70 are considered to be fair and values greater than .80 are considered to be good (Swets, 1988). Based on these guidelines, the AUCs associated with the DSPA alone or in combination with the SDT showed it was a fair predictor of reading outcome when measured by Word Attack performance. The DSPA alone or in combination with the SDT failed to reach an acceptable level of prediction when the Word Identification subtest was used to index reading outcome.

DSPA as Compared to the ISF

Sample 2 was used to compare the DSPA to the ISF, which is a commonly used universal screening instrument (Good & Kaminski, 2003). The distributional characteristics of all variables from Sample 2 are presented in Table 3. Results concerning skewness again indicated that the DSPA did not appear to have a floor effect, whereas the ISF and both reading achievement outcome measures were characterized by a positive skew indicative of a floor effect. The absence of a floor effect for the DSPA is particularly notable as the majority of students in this sample were deemed to be at risk for RD.

A series of hierarchical linear regression analyses examined the amount of variance the DSPA accounted for over and above the ISF when predicting reading achievement as indexed by the Word Identification and Word Attack subtests at the end of kindergarten. Again, the variables were entered in a sequential manner, with the ISF entered first and the DSPA entered second. Results indicated that the ISF accounted for 28% of the variance in both Word Identification and Word Attack. When entered as the second variable, the DSPA accounted for a significant amount of unique variance (5%) for the Word Attack subtest but a nonsignificant amount of unique variance (2%) for the Word Identification subtest.

A series of logistic regression analyses was also conducted to examine the relationship between the predictor variables (ISF and DSPA) and end-of-year reading outcomes (see Table 4). For each measure, the first two models included the ISF or the DSPA as a single predictor of reading outcome. In the final models, variables were entered in a sequential manner with ISF entered first and DSPA entered second. In these analyses, we again identified good and poor readers using a 25th percentile cutoff score on reading outcome measures. As was done in Sample 1, we used sample statistics to determine the 25th percentile score. For this sample, the cutoff score was 3 for Word Identification and 1 for Word Attack. These procedures identified 26 children who were poor readers on the Word Identification subtest and 28 poor readers on the Word Attack. As seen in Table 4, both ISF and DSPA were significant single predictors of reading outcomes. Furthermore, DSPA was found to significantly predict both reading outcomes over and above the ISF. Data from the AUCs showed that the DSPA was a good predictor of reading outcome on the Word Attack subtest when entered alone or in combination with the ISF. The AUCs associated with the DSPA's prediction of reading outcome on the Word Identification subtest was in the "fair" range.

DSPA as a Supplemental Screening Measure

The DSPA was developed in part to serve as a supplemental measure to reduce false positives associated with many early screening measures. To assess its usefulness in this regard, additional analyses were carried out. In these analyses, the ISF was used as the initial measure to identify children at risk. The DSPA was then used as a supplemental screening measure to reclassify students deemed to be at risk by the ISF. Classifications were categorized in the following manner: (a) true positives = poor readers identified as at risk, (b) true negatives = good readers identified as not at risk, (c) false positives = good readers identified as at risk, and (d) false negatives = poor readers identified as not at risk. The sensitivity and specificity of prediction were then used as indices for the evaluation of this screening approach. Sensitivity refers to the percentage of true positives among those identified as poor readers, whereas specificity refers to the percentage of true negatives among those identified as good readers.

In the first set of analyses, the cutoff score for ISF (< 8) that is recommended for the beginning of kindergarten was used to identify poor readers (Good & Kaminski, 2003). At-risk students were then reclassified using the mean score on the DSPA as a cutoff score. Other cutoff scores on the DSPA were also utilized, but the mean score provided the best trade-off between sensitivity and specificity.

Using the recommended cutoff score on the ISF resulted in 17 false positives when predicting outcomes on the Word Identification subtest and 16 false positives when predicting outcomes on the Word Attack subtest (see Table 5). By using the DSPA as a secondary measure, we were able to reduce these false positives by 35% to 40% while increasing false negatives in only a limited manner (10%). Although these results suggest that the DSPA has potential as a secondary measure, this conclusion is tempered by the poor sensitivity shown in these data. These results indicate that more than 40% of the children who become poor readers are missed using the above procedure. Debate continues on what may be an acceptable level of sensitivity; however, many agree with Jenkins (2003) that values close to .9 are desirable to minimize the number of students missed by the screening measure. In the above analyses, the overall sensitivity was influenced considerably by the use of the recommended cutoff score for the ISF. To address this problem, we selected a cutoff score on the ISF that provided a sensitivity level greater than .9. As seen in Table 6, this led to very high false positive rates, especially for the prediction of Word Identification. However, by selecting the appropriate cutoff score on the DSPA (i.e., 50) to predict outcome in Word Identification, we were able to reduce the false positive rate by 45%, with only a limited change in sensitivity.

For outcome in Word Attack, false positives could be reduced by 25% with only a limited change in sensitivity.

The above analyses provide initial evidence for the use the ISF and the DSPA in a two-step screening process. However, it is important to investigate whether this approach would result in higher classification indices than those associated with administering all students both measures or each individually. To investigate this possibility, the specificity values associated with the supplemental models (Table 6) were compared to those of the combined screening, the DSPA alone, and the ISF alone, when sensitivity was held constant to that found with the supplemental model. Table 7 shows that for Word Identification, the supplemental model has better specificity (fewer false positives) than the combined or ISF alone models. However, the DSPA alone model has better specificity than the supplemental model. For Word Attack, the supplemental model again is better than the combined and the ISF alone model, and it also results in fewer false positives than the DSPA alone model.

Discussion

The purpose of this study was to investigate the validity and usefulness of a dynamic screening of phonological awareness. There has been a resurgence of interest in dynamic assessment with the recent consideration of RTI as a framework for early identification (Caffrey, 2007; L. S. Fuchs et al., 2008; Grigorenko, 2009). In this study, we administered a dynamic screening of phonological awareness to two samples of kindergarten children. In the first sample, we compared the dynamic measure (i.e., DSPA) to a static version of the same test. Our results showed that the DSPA predicted reading achievement and outcome over and above the static version. These results provide initial evidence that a dynamic component can add to the predictive validity of a static phonological awareness screening measure.

In a second sample of kindergarten children, we administered the DSPA and the ISF subtest of DIBELS. The ISF was chosen because of its widespread use as a universal screening measure in elementary schools. The ISF is commonly used to inform placement of kindergarten students in Tier 2 intervention. We compared the predictive validity of the DSPA to that of the ISF. Our findings indicated that the DSPA accounted for a significant amount of variance over and above the ISF for one of the two measures of reading achievement. The DSPA also provided unique prediction of good versus poor reading outcomes. In fact, by itself or combined with the ISF, the DSPA was found to be a fair predictor of reading outcome measured by the Word Identification subtest and a good predictor for reading outcomes as measured by the Word Attack subtest. In general, the predictive validity was better when predicting Word Attack outcomes than Word Identification across all analyses. This is expected given that phonological awareness should be more directly related to phonological decoding (measured by the Word Attack subtest) than to sight word reading (measured by the Word Identification subtest;Ehri, 1998; Wagner & Torgesen, 1987).

Of particular interest was the use of the DSPA as a supplementary screening measure. It has been suggested that dynamic assessment might be most useful as a secondary measure that is employed to reduce false positives introduced by a highly sensitive screening instrument (O’Connor & Jenkins, 1999). We tested this approach in several analyses. First, we used the recommended cutoff scores for the ISF to identify children at risk. The DSPA did reduce false positives, but models were characterized by poor sensitivity. In a second set of analyses, we chose an ISF cutoff score that resulted in at least 90% sensitivity, the level recommended by Jenkins (2003) as an appropriate level for a universal screening instrument or process. As expected, such a cutoff score led to a large number of false positives. These false positives could be reduced by 25% to 45% by using the DSPA as a supplemental measure, with only a small cost to sensitivity. However, further data also indicated that nearly similar results could be obtained by using the DSPA alone to predict reading outcome. Thus, although our findings provide some initial evidence that a dynamic screening of phonological awareness may contribute to the screening process, it is not clear whether it may best be used as a primary measure of phonological awareness or as a supplementary measure.

Dynamic Screening

Our findings are consistent with those of others who have reported that a dynamic screening of phonological awareness can add to the prediction of early reading achievement and outcome (O’Connor & Jenkins, 1999; Spector, 1992). Specifically, dynamic screening may be able to reduce the false positives associated with universal screening in the early school grades. Although preschool opportunities provide many children with literacy experience and instruction, a large number of children continue to enter kindergarten with limited literacy knowledge. A lack of literacy knowledge at the beginning of kindergarten can affect performance on measures of phonological awareness and lead to children being identified as at risk for RD. In fact, it is not uncommon for early screening measures to be associated with an overrepresentation of children at the bottom of the distribution (Catts et al., 2009). Such a floor effect in turn can lead to a high false positive rate. Dynamic tasks may allow some students with partially developed knowledge the opportunity to demonstrate their potential for reading success while at the same time identifying other children who may have more trouble learning to read. Such an assessment should act to better separate student ability levels and lead to less of a floor effect. Indeed, our results showed that the dynamic screening measure had a much less positively skewed distribution than the static assessments.

An added advantage of dynamic assessment is the information it may provide the teacher or practitioner concerning a student's modifiability as related to instruction. By measuring how well children respond to targeted instruction and feedback concerning reading-related skills, the examiner may be able to gauge how difficult it will be for children to learn to read over a more extend period of instruction. Such information should be valuable for not only determining risk for RD but also designing treatment and education plans for children.

Our results provide some preliminary evidence that dynamic screening of phonological awareness may be useful in universal screening within a RTI framework. Although useful, such a measure may need to be combined with other screening tools that assess reading-related skills. As a single predictor, our dynamic screening measure explained only a limited amount of variance in reading achievement and had only fair to good predictability of reading outcome. Most previous research has shown that multivariate screening approaches have more validity than univariate approaches (Catts et al., 2001; Compton, Fuchs, Fuchs, & Bryant, 2006; O'Connor & Jenkins, 1999). In addition, this research has shown that multivariate screening batteries that are composed of various component skills (e.g., letter identification, rapid naming) improve classification accuracy. Therefore, future research should examine the predictive validity and usefulness of a dynamic screening of phonological awareness within the context of a multivariate screening approach.

Acknowledgments

Funding

This study was supported in part by a National Institutes of Health Institutional Traineeship (T32 DC000052) and an Institute of Education Sciences grant (R324 A080118).

References

Bridges, MS. Unpublished doctoral dissertation. Lawrence: University of Kansas; 2009. The use of a dynamic screening of phonological awareness to predict reading outcomes in kindergarten children.

Bridges, MS.; Catts, HW. Dynamic Screening of Phonological Awareness (Prepublication version). Moline, IL: Linguisystems; 2008.

Caffrey, E. Unpublished doctoral dissertation. Nashville, TN: Vanderbilt University; 2006. A comparison of dynamic assessment and progress monitoring in the prediction of reading achievement for students in kindergarten and first grade.

Castles A, Coltheart M. Is there a causal link from phonological awareness to success in learning to read? Cognition. 2004; 91:77–111. [PubMed: 14711492]

Catts HW, Fey ME, Zhang X, Tomblin JB. Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implementation. Language, Speech, and Hearing Services in Schools. 2001; 32:38–50.

Catts HW, Petscher Y, Schatschneider C, Bridges MS, Mendoza K. Floor effects associated with universal screening and their impact on the early identification of reading disabilities. Journal of Learning Disabilities. 2009; 42:163–176. [PubMed: 19098274]

Compton DL, Fuchs D, Fuchs LS, Bryant JD. Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. Journal of Educational Psychology. 2006; 98:394–409.

Denton, CA.; Mathes, PG. Intervention for struggling readers: Possibilities and challenges. In: Foorman, BR., editor. Preventing and remediating reading difficulties: Bringing science to scale. Timonium, MD: York; 2003. p. 229-251.

Duncan, SE.; DeAvila, EA. Pre-LAS. Monterey, CA: CTB/McGraw-Hill; 1998.

Ehri, LC. Grapheme-phoneme knowledge is essential for learning to read words in English. In: Metsala, JL.; Ehri, LC., editors. Word recognition in literacy. Mahwah, NJ: Lawrence Erlbaum; 1998. p. 3-40.

Foy JG, Mann V. Home literacy environment and phonological awareness in preschool children: Differential effects for rhyme and phoneme awareness. Applied Psycho-linguistics. 2003; 24:59–88.

Fuchs D, Fuchs LS. Responsiveness-to-intervention: A blueprint for practitioners, policymakers, and parents. Teaching Exceptional Children. 2005; 38:57–61.

Fuchs LS, Compton DL, Fuchs D, Hollenbeck KN, Craddock CF, Hamlett CL. Dynamic assessment of algebraic learning in predicting third graders' development of mathematical problem solving. Journal of Educational Psychology. 2008; 100:829–850. [PubMed: 19884957]

Good, RH.; Kaminski, RA. Dynamic Indicators of Basic Early Literacy Skills. Longmont, CO: Sopris West; 2003.

Good, RH.; Kaminski, RA.; Shinn, MR.; Bratten, J.; Shinn, M.; Laimon, D.; Flindt, N. Technical adequacy of DIBELS: Results of the Early Childhood Research Institute on measuring growth and development (Tech. Rep. No. 7). Eugene: University of Oregon; 2004.

Grigorenko EL. Dynamic assessment and response to intervention: Two sides of one coin. Journal of Learning Dis-abilities. 2009; 42:111–132.

Haager, D.; Klingner, J.; Vaughn, S. Evidence-based reading practices for response to intervention. Baltimore, MD: Brookes; 2007.

Jenkins, J. Candidate measures for screening at-risk students; Paper presented at the NRCLD Responsiveness-to-Intervention Symposium; Kansas City, MO. Dec. 2003

Lidz, CS. Practitioner's guide to dynamic assessment. New York, NY: Guilford; 1991.

O'Connor RE, Jenkins JR. The prediction of reading disabilities in kindergarten and first grade. Scientific Studies of Reading. 1999; 3:159–197.

Rathvon, N. Early reading assessment: A handbook for practitioners. New York, NY: Guilford; 2004.

Simmons D, Coyne M, Kwok O, McDonagh S, Harn B, Kame'enui E. Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. Journal of Learning Disabilities. 2008; 41:143–158. [PubMed: 18354934]

Spector JE. Predicting progress in beginning reading: Dynamic assessment of phonemic awareness. Journal of Educational Psychology. 1992; 84:353–363.

Sternberg, RJ.; Grigorenko, EL. Dynamic assessment: The nature and measurement of learning potential. Cambridge, UK: Cambridge University Press; 1998.

Swets J. Measuring the accuracy of diagnostic systems. Science. 1988; 240:1285–1293. [PubMed: 3287615]

Vellutino FR, Scanlon DM, Small S, Fanuele DP. Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. Journal of Learning Disabilities. 2006; 38:157–169. [PubMed: 16583795]

Wagner RK, Torgesen JK. The nature of phonological processing and its causal role in the acquisition of reading skills. Psychological Bulletin. 1987; 101:192–212.

Woodcock, RW. Woodcock Reading Mastery Tests-Revised Normative Update. Circle Pines, MN: American Guidance Service; 1998.

Table I

Descriptive Statistics for Measures Administered to Sample 1

	M	*SD*	*Mdn*	Range	Skew	Kurtosis
SDT	4.73	5.43	3.0	0–19	0.990	0.387
DSPA	34.82	22.39	35.0	0–79	−0.062	−0.931
Word Identification	13.81	10.95	12.5	0–48	1.124	0.988
Word Attack	6.67	4.18	7.0	0–24	0.823	2.110

Note: SDT= Static Deletion Task; DSPA = *Dynamic Screening of Phonological Awareness.*

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Classification Indices for Logical Regression Involving the SDT and DSPA in Sample 1

Outcome Measure	*B*	*SE*	Wald	*p* Level	AUC
Word ID					
SDT alone	−0.92	.049	3.51	.061	.605
DSPA alone	−.031	.011	7.37	.007	.667
Combined SDT	.059	.088	0.449	.503	
DSPA	−.42	.020	4.24	.040	.692
Word Attack					
SDT alone	−.182	.063	8.37	.004	.635
DSPA alone	−0.50	.013	14.55	.001	.766
Combined SDT	.012	.099	0.014	.907	
DSPA	−0.52	.021	5.99	.014	.772

Note: SDT= Static Deletion Task; DSPA = *Dynamic Screening of Phonological Awareness*; AUC = area under the curve.

Table 3

Descriptive Statistics for Measures Administered to Sample 2

	M	*SD*	*Mdn*	Range	Skew	Kurtosis
ISF	12.82	9.39	11.15	0–38	0.806	0.894
DSPA	44.59	26.26	48.00	0–78	−0.145	−1.05
Word Identification	14.12	14.14	8.00	0–61	1.303	1.17
Word Attack	6.74	7.10	6.00	0–33	1.679	3.35

Note: ISF = Initial Sound Fluency; DSPA = *Dynamic Screening of Phonological Awareness*.

Table 4

Classification Indices for the Logistic Regression Involving the ISF and DSPA in Sample 2

Outcome Measure	*B*	*SE*	Wald	*p* Level	AUC
Word ID					
ISF alone	−.114	.036	9.980	.002	.742
DSPA alone	−.034	.010	11.878	.001	.755
Combine ISF	−.075	.040	3.425	.064	
DSPA	−.022	.012	3.692	.045	.760
Word Attack					
ISF alone	−.167	.043	15.109	.001	.796
DSPA alone	−.044	.011	17.704	.001	.801
Combined ISF	−.117	.046	6.499	.011	
DSPA	−.029	.012	5.824	.016	.825

Note: ISF = Initial Sound Fluency; DSPA = *Dynamic Screening of Phonological Awareness*; AUC = area under the curve.

Table 5

Classification Indices Across Models Utilizing ISF Recommended Score Alone and Combined With the DSP-R as a Supplemental Screening

	TP	TN	FP	FN	Specificity	Sensitivity
Word Identification						
Model I:ISF < 8 only	16	53	17	10	76	62
Model 2:ISF + DSPA-R ≤ 45	15	59	11	11	84	58
Word Attack						
Model 3:ISF < 8	17	52	16	11	76	61
Model 4:ISF + DSPA-R ≤ 45	16	58	10	12	85	57

Note: ISF = Initial Sound Fluency; DSPA = *Dynamic Screening of Phonological Awareness*; TN= true positive; TN= true negative; FP = false positive; FN = false negative.

Table 6

Classification Indices Across Models Utilizing ISF Recommended Score Alone and Combined With the DSP-R as a Supplemental Screening

	TP	TN	FP	FN	Specificity	Sensitivity
Word Identification						
Model I (ISF < 20)	24	19	51	2	27	92
Model 2 (DSPA-R ≤ 50)	22	42	28	4	60	85
Word Attack						
Model 3 (ISF < 13)	26	39	29	2	57	92
Model 4 (DSPA < 55)	24	46	22	4	67	86

Note: ISF = Initial Sound Fluency; DSPA = *Dynamic Screening of Phonological Awareness*; TN = true positive; TN = true negative; FP = false positive; FN = false negative.

Table 7

Specificity for the Supplemental Model, Combination Model, DSPA Alone Model, and ISF Alone Model When Sensitivity Is Held Constant to That Found With the Supplemental Model

		Specificity			
	Sensitivity	**Supplemental**	**Combined**	**DSPA**	**ISF**
Sample 2					
Word ID	85	60	49	63	30
Word Attack	86	67	60	63	59

Note: DSPA = *Dynamic Screening of Phonological Awareness*; ISF = Initial Sound Fluency.