Thank you for your request to our REL Reference Desk regarding research that has been conducted on literacy assessments and their direct correlation to student achievement. Ask A REL is a collaborative reference desk service provided by the ten regional educational laboratories (REL) that, by design, functions much in the same way as a technical reference library. It provides references, referrals, and brief responses in the form of citations on research based education questions.

The information below represents the most rigorous research available. Researchers consider the type of methodology and give priority to research reports that employ well described and thorough methods. The resources were also selected based on the date of the publication with a preference for research from the last ten years. Additional criteria for inclusion include the source and funder of the resource.

**Question:** Is there research addressing literacy assessments and their direct correlation to student achievement?

**Key words and search strings used in the search:** literacy assessments AND relation to student achievement; academic achievement AND literacy assessments; relationship between literacy assessments and academic achievement; predictive validity AND reading assessments AND student achievement

**Search databases and websites:**
2. JSTOR: [http://www.jstor.org/action/showAdvancedSearch](http://www.jstor.org/action/showAdvancedSearch)
3. Google Scholar: [www.google.com/scholar](http://www.google.com/scholar)

**Citations Retrieved:** (NOTE: Abstracts and executive summaries are copied directly from the reports when possible to ensure accuracy):


**Abstract/Summary:** The purpose of this study was to examine oral reading fluency (ORF) in the context of a large-scale federal reading initiative conducted in low performing, high poverty schools. The objectives were to (a) investigate the relation between ORF and comprehensive
reading tests, (b) examine whether slope of performance over time on ORF predicted performance on comprehensive reading tests over and above initial level of performance, and (c) test how well various models of ORF and performance on high stakes reading tests in Year 1 predicted performance on high-stakes reading tests in Year 2. Subjects were four cohorts of students in Grades 1-3, with each cohort representing approximately 2,400 students. Results support the use of ORF in the early grades to screen students for reading problems and monitor reading growth over time. The use of ORF in reading reform and implications for school psychologists are discussed. (Contains 3 tables and 1 figure.)


**Abstract/Summary:** The assessment of early literacy skills during the kindergarten year can provide useful information about student performance in prereading skills, which are predictors of later reading achievement. This study examined the use of fluency-based prompts of student phonemic awareness, alphabetic principle, and oral reading at the end of kindergarten for predicting later reading achievement at the end of second grade. Predictive validity and bias studies were undertaken with respect to English-language learners (ELLs) and four selected ethnic subgroups: European American (EA), African American (AA), Asian American (AsA), and Hispanic American (HA). Results indicated that the predictive validity of the early literacy measures was strong, and no evidence of predictive bias for ELL and non-ELL groups was found. However, evidence of a small amount of predictive bias was found between the EA and HA students with respect to intercept differences. (Contains 3 tables.)


**Abstract/Summary:** 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. (Contains 1 note and 7 tables.)

**Abstract/Summary:** Current empirical evidence indicates poor learning trajectories for students with early literacy skill deficits. As such, reliable and valid detection of at-risk students through regular screening and progress monitoring is imperative. This study investigated the predictive validity of scores on the Dynamic Indicators of Basic Early Literacy Skills (DIBELS). Logistic regression analyses were used to test the utility of the DIBELS first grade indicators for predicting reading proficiency on TerraNova California Achievement Test (CAT) Assessment and Pennsylvania System of School Assessment (PSSA) in second and third grade, respectively. Results suggest that students’ first grade Oral Reading Fluency (ORF) DIBELS risk category scores were the only significant predictor of future TerraNova and PSSA reading proficiency. Although the current data present encouraging results for the predictive validity of ORF as a screening tool for early readers, further investigations of the utility of the remaining indicators (Letter Naming Fluency, Nonsense Word Fluency, and Phonemic Segmentation Fluency) are warranted. (Contains 5 tables.)


**Abstract/Summary:** Because of the increased emphasis on standardized testing results, scores from a high-stakes, end-of-year test (Tennessee Comprehensive Assessment Program [TCAP] Reading Composite) were used as the standard against which scores from a group-administered, curriculum-based measure (CBM), Monitoring Instructional Responsiveness: Reading (MIR:R), were compared for 448 third-grade students. A zero-order correlation coefficient of 0.58 (p < 0.001) partially defined the relationship between the MIR:R composite score (comprehension rate) and student performance on the TCAP reading composite; a classification analysis yielded the following percentages: sensitivity = 85, specificity = 53. Results from a stepwise multiple-regression equation revealed that the Comprehension score provided moderate predictive validity for TCAP reading composite performance (29% variance accounted for, p < 0.001); the rate (Total Words Read) score was less predictive (1% additional variance accounted for, p < 0.05). Discussion focuses on the implications of using unidimensional versus multidimensional CBMs for early screening and/or progress monitoring within response to intervention.


Abstract/Summary: This study examined the changing role and longitudinal predictive validity of curriculum-embedded progress-monitoring measures (CEMs) for kindergarten students receiving Tier 2 intervention and identified as at risk of developing reading difficulties. Multiple measures were examined to determine whether they could predict comprehensive latent first- and second-grade reading outcomes and whether their predictive validity changed concurrent with reading development. CEMs of phonemic, alphabetic, and integrated tasks were given 3 times during the kindergarten year to 299 students. Structural equation modeling indicates that CEMs explained a significant amount of variance on first- (54%-63%) and second-grade (34%-41%) outcomes. The predictive validity of specific measures varied over the kindergarten year with sound and letter identification measures being predictive early and segmenting and word reading becoming important as reading abilities progressed. Findings suggest that CEMs may be viable and helpful tools for making data-driven instructional decisions in a response to intervention framework.


Abstract/Summary: The factor structure of the CORE Phonics Survey was analyzed using a sample of 165 students in upper elementary school with specific learning disabilities. Confirmatory factor analysis was used to identify the hypothesized constructs of the CORE Phonics Survey and predictive validity of the CORE Phonics Survey to predict students' success in word level reading and oral reading fluency one year later. The results of the analyses indicated the two-factor model representing two latent variables, alphabetic knowledge and decoding skills, provided the most appropriate fit to the given data, indicating that the hypothesized two factors of the CORE Phonics Survey demonstrate construct validity. The two constructs of CORE Phonics Survey significantly predicted students' fluency and decoding abilities as measured by norm-referenced assessments one year later. The practical implications and limitations of the present study are discussed


Abstract/Summary: This study investigated the predictive validity of a dynamic assessment designed to evaluate later risk for reading difficulty in bilingual Latino children at risk for language impairment. During kindergarten, 63 bilingual Latino children completed a dynamic assessment nonsense-word recoding task that yielded pretest to posttest gain scores, residuum gain scores, and modifiability scores. At the end of first grade, the same participants completed
criterion reading measures of word identification, decoding, and reading fluency. The dynamic assessment yielded high classification accuracy, with sensitivity and specificity at or above 80% for all three criterion reading measures, including 100% sensitivity for two out of the three first-grade measures. The dynamic assessment used in this study has promise as a means for predicting first-grade word-level reading ability in Latino, bilingual children.


Abstract/Summary: The study addresses the extent to which subtests on the Dynamic Indicators of Basic Early Literacy Skills Reading Assessment (DIBELS; Good & Kaminski, 2002) predict student success on a measure of reading comprehension and if prediction is consistent for native and second English Language Learners. 2,649 elementary students were assessed on a reading comprehension measure, of which 29.7% were English Language Learners. Descriptive and analytic statistics were generated including bivariate correlation analysis split by language proficiency. Critical measures and suggested cutoff values (Good, Simmons, et al., 2002) were evaluated for predictive utility by visualization of Receiver Operating Characteristic (ROC) curves (Swets, Dawes, & Monahan, 2000), and comparison of the area-under-the-curve (AUC) values. DIBELS better predicts children who are at "low risk" than those "at risk;" however, DIBELS correctly classifies children "at risk" better for ELL than non-ELL students in third grade. (Contains 10 tables and 3 figures.)

Referrals
- Florida Center for Reading Research: http://www.fcrr.org/
- Center on Teaching and Learning: http://ctl.uoregon.edu/

Federally Funded Resources:
- Institute of Education Sciences (IES), public search engine available at: http://ies.ed.gov/pubsearch/

Disclaimer:
This Ask A REL response was developed by REL-SE under Contract ED-IES-12-C-0011 from the U.S. Department of Education, Institute of Education Sciences. The content does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government.