

Approved Literacy Assessments

meeting the requirements of 279.68/ELI
for universal screening and progress monitoring
2016-2017 school year

August 2016



Overview

This document contains information about assessments approved for use as universal screening and/or progress monitoring measures to meet the requirements of 279.68/ELI.

Background

In the winter of the 2012-2013 school year the Department requested and reviewed information and Proposals for a statewide universal screener (US) and progress monitoring (PM) assessment system for preschool through 6th grade literacy. There were two outcomes from this process. First, the Department reviewed all submitted assessments, and identified the Formative Assessment System for Teachers (FAST) and the Individual Growth & Development Indicators (IGDIs) as approved assessments to be purchased and supported across the state for all schools interested in participating in an early warning system for literacy to support implementation of a Multi-Tiered System of Supports (MTSS). Second, the results of the reviews were published to inform users of the relevant technical qualities of other literacy measures that might be used in local schools as they made selections for their own universal screening and progress monitoring measures.

Legislation funded at the end of the 2013 legislative session caused the Department of Education to revisit these reviews in early 2014 for the purposes of setting minimum standards for assessments approved for use in universal screening and progress monitoring of K-3 literacy as required by Iowa Code section 279.68 (review information for grades 4-6 is included to provide continuity). Each subsequent year the Department has requested information from vendors to update the list of assessments approved to meet requirements of Iowa Code section 279.68. This document represents the updated approved list for the 2016-2017 year with additional measures meeting the established 2016-2017 criteria added. The criteria used for the current evaluation are provided in the appendix.

Each submitted assessment was evaluated separately at each grade level and at for each purpose (i.e., universal screening and progress monitoring) for which information was submitted. To be approved at a grade level, the measure needed to meet all criteria for the assessment purpose as well as two general technical adequacy criteria.

Check marks indicate where an assessment met the minimum requirements. Assessments marked with an asterisk (*) have ambiguous, incomplete, or no established benchmarks at one or more screening window (i.e., fall, winter, and spring). These assessments may not be adopted without consultation with Iowa Department of Education staff to clarify the benchmarks to be used.

2016-2017	Universal Screening							Progress Monitoring							Additional Information			
State Supported Assessments	K	1	2	3	4	5	6	K	1	2	3	4	5	6	Approx. Admin. (min.)	Time to Access to Data	Training Required (days)	Seasonal Benchmarks Available?
FAST Adaptive Reading (aReading)	✓	✓	✓	✓	✓	✓	✓								6-20	Instant	0.5	Yes
FAST CBMreading (CBMR)		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	1-3	Instant	0.5	Yes
FAST earlyReading 1st Composite		✓													3-5	Instant	0.5	Yes
FAST earlyReading K Composite	✓														5-7	Instant	0.5	Yes
FAST earlyReading Decodable Words								✓							1-3	Instant	0.5	Yes
FAST earlyReading Letter Naming								✓							1-3	Instant	0.5	Yes
FAST earlyReading Letter Sounds								✓							1-3	Instant	0.5	Yes
FAST earlyReading Nonsense Words								✓	✓						1-3	Instant	0.5	Yes
FAST earlyReading Onset Sounds								✓							1-3	Instant	0.5	Yes
FAST earlyReading Sight Words 150									✓						1-3	Instant	0.5	Yes
FAST earlyReading Word Blending								✓	✓						1-3	Instant	0.5	Yes
FAST earlyReading Word Segmenting								✓	✓						1-3	Instant	0.5	Yes
Additional Approved Assessments	Universal Screening							Progress Monitoring							Additional Information			
AIMSweb – Letter Sound Fluency	✓							✓							1	Instant	0	Yes
AIMSweb – Letter Naming Fluency	✓							✓							1	Instant	0	Yes
AIMSweb – Maze				✓	✓	✓	✓			✓					3	Instant	0	Yes
AIMSweb – Reading CBM*		✓	✓	✓	✓	✓	✓		✓	✓	✓				3	Instant	0	No fall 1 st
DIBELS Next Composite		✓	✓	✓	✓	✓		✓	✓	✓					1-6	Instant	1-2	Yes
easyCBM*				✓	✓	✓	✓	✓	✓	✓	✓		✓		45-60	Same Day	1	Multiple
Edcheckup Maze Reading Passages*			✓		✓										3	Instant	0.5	Multiple
Edcheckup Standard Reading Passages*			✓	✓	✓			✓							3-5	Instant	0.5	Multiple
Gates MacGinitie Reading Tests, 4 th Edition*				✓	✓										75-100	>5 Days	0	Same F, W, S
i-Ready Diagnostic			✓	✓	✓	✓	✓								25-46	Instant	0.5	Yes
i-Ready Growth Monitoring								✓	✓	✓	✓	✓	✓	✓	6-15	Instant	0.5	Yes
mCLASS:Reading 3D Text, Reading, and Comprehension*		✓	✓	✓											5-8	Instant	1-2	Unclear
Observation Survey of Early Lit. Achievement		✓													>15	Same Day	2	Yes
Phonological Awareness and Lit. Screening (PALS 1-3)*		✓		✓											25	Instant	1	No Winter
Phonological Awareness and Literacy Screening (PALS K)*	✓														30	Instant	0.5	No Winter
STAR Early Literacy	✓	✓	✓	✓				✓	✓	✓	✓				13	Instant	0	Yes
STAR Reading		✓	✓	✓	✓	✓	✓		✓		✓	✓	✓		11	Instant	0	Yes
Texas Primary Reading Inventory (TPRI) *	✓	✓	✓	✓											1-5	Same Day	0.5	Multiple

Appendix

The following definitions are intended to provide a general explanation of the meaning and application for each of the required statistics and/or reporting elements for early literacy assessments. The minimum and desired values from the evaluation are reported.

Reliability

Reliability is a common measure of an important quality of an assessment: consistency. Common methods to describe reliability include: internal consistency (i.e., do the parts of the test work together to measure the same thing?), consistency over time (i.e., can we trust that the test will measure consistently over time?), and consistency across testers (can the test be administered and scored to get consistent results?). *Using a scale from 0.0 to 1.0, a reliability value of at least 0.70 is required.*

Validity

Validity statistics are used to help understand if the test results will allow users to make appropriate decisions. Many things can go into this understanding. For example, we ask how well the test results compare to another known measure of reading (this is called criterion validity). A universal screening or progress monitoring test that compares favorably with another measure of reading increases the confidence that the results of the screening test are related to the student's reading ability and that our decisions about that student's skills are accurate. *Using a scale from 0.0 to 1.0, a validity coefficient of at least 0.30 was required.*

Area Under the Curve (AUC)

Area under the curve (or AUC) is a statistical calculation that represents the relative value of a test for accurately classifying outcomes. The closer to 1.0 the AUC value, the better the test at predicting student success. A test with an AUC value of 0.50 predicts at the same rate as chance – in other words, the test is no better than flipping a coin. *Using a scale from 0.0 to 1.0, an AUC value of at least .70 is required.*

Sensitivity/Specificity

Sensitivity and Specificity are statistics that represent the ability of the test to correctly identify students. Sensitivity represents the ability of the test to correctly identify the positive cases (students predicted on track for success). Specificity represents the ability of the test to correctly identify the negative cases (students predicted not on track for success). In the case of universal screening, the aim is for high sensitivity for a prediction of students on track to be successful readers. A test with a high value for sensitivity (approaching 1.0) will rarely miss identifying students who are on track to be successful readers. *Using a scale from 0.0 to 1.0, a sensitivity/specificity value of at least .70 is required.*

Number of Administrations per Year

For the purposes of universal screening it is ideal that a screenings can be completed at least three times a year. This will help school adhere to the 279.68/ELI, but most importantly, enable schools to periodically check to ensure students are on track for later reading success.

Benchmarks Available for Fall, Winter and Spring Screening Windows

Benchmarks are established to identify a level of performance which identifies students to be on track to be successful readers. These benchmarks are used to make universal screening decisions and to set progress monitoring goals. Assessments should have valid benchmarks established for each universal screening window: fall, winter and spring. *While some assessments reviewed previously do not have benchmarks for each grade and season, for the most recent review it was required that each measure have established benchmarks for each of the three screening windows.*

Standard Setting Methodology

Seasonal benchmarks can be set in a number of different ways including normative and criterion-referenced studies, expert opinions, and alignment studies to other instruments, to name a few. *For purposes of this review, it was required that each benchmark be established using sound criteria.*

Number of Forms of Demonstrated Equivalence

When using an assessment to monitor progress weekly it is important to make sure that there are enough forms to avoid a practice effect (i.e., any increases in score are due to student skills gain rather than something else). It is also important to reduce any variation in test results over time caused by forms that are not of similar difficulty. *At least ten forms are required.*

Reliability of Slope

Reliability of slope is a statistic that represents the ability of the test to produce a consistent measure of student growth over time. A test with a less-reliable slope will do a poor job of accurately reflecting student improvement. A test with a very reliable slope will show results that best represent the student's improvement over time. *Using a scale from 0.0 to 1.0, a reliability of slope value of at least 0.60 is required.*

Administration Time

It is important to find tests that are efficient, especially when used for weekly progress monitoring purposes. Since assessment takes away from instructional time, it is a good to minimize the amount of time spent testing. If two tests are otherwise similar (AUC, reliability, etc.), the test that takes less time may be preferred. *Although no time maximum was established in previous reviews, for purposes of this review, administration time for progress monitoring assessments were required to be less than 15 minutes, although even shorter would be highly preferred.*

ADDITIONAL INFORMATION (not used in evaluation of measures)

Accessibility of Student Data

For the purposes of universal screening and monitoring progress it is important for teachers to gain access to results quickly in order to begin using the data. A lag between testing and availability of data will cause the system to be less responsive to student needs. It is preferred to be able to receive and use results very quickly after testing.

Teacher Training

The amount of training needed to reliably administer the tests and use the results is important for planning and resource allocation.