Early Literacy Screening Assessments
Selection Criteria for the KDE Approved List

The Kentucky Department of Education (KDE) has adopted the following minimum criteria for the selection of valid and reliable Universal Screening tools guided by the National Center on Intensive Intervention at American Institutes for Research Screening Tools Chart Rating System and the Institute for Education Sciences:

Technical Adequacy (Based on statistical analysis)	➢ Classification Accuracy – The lower band of the confidence interval of the area under the curve (AUC) must meet or exceed .70 or sensitivity must meet or exceed .70 and specificity must meet or exceed .70. ➢ Reliability – Reliability measures (usually reported as internal consistency or Cronbach's alpha) must meet or exceed .70. ➢ Validity – Predictive validity must meet or exceed .60. ➢ Representative Sample – The normative group closely matches the characteristics of its population as a whole.

Administration Usability and Support	➢ **Brief** – Administration time 60 minutes or less;
	➢ **Scoring criteria** (includes a minimum of the following):
	    o Provides a percentile rank that compares the students' results to a nationally representative group;
	    o Includes a pre-determined (external) benchmark score that represents levels of proficiency;
	    o Provides percentile ranks and benchmark scores in narrowly defined skills (e.g., phonemic awareness, comprehension, etc.);
	➢ **Able to be administered more than once per year** – Screening three times per year is necessary to evaluate effectiveness of the program; and
	➢ **Vendor training and support.**

Districts are advised to perform internal reviews of the early literacy screening materials to make assessment decisions that are an appropriate fit for their local context. These tools are intended to be administered as part of a comprehensive screening and assessment system within a multi-tiered system of supports.

Definitions:

Benchmark/Cut-Score means a score on a screening measure that divides students who are considered potentially at risk from those who are not at risk.

Classification Accuracy means how well the scores on a screening assessment correctly identify students at risk versus those not at risk. Area under the curve (AUC) values closer to 1 indicate the screening measure reliable distinguishes students at risk from those not at risk.

Reading diagnostic assessment means an assessment that measures a student's skills against established performance levels in essential components of reading and identifies students that require intervention in at least one (1) of those components to accelerate the student's progress toward proficient performance in reading.

Reliability means the consistency of a set of scores that are designed to measure the same thing. It is important to ensure the type of reliability is appropriate for the type of screener being used.

Representative Sample means a group that closely matches the characteristics of its population as a whole.

Sensitivity means the probability of correctly identifying a problem. Sensitivity rates help determine which tests can achieve a high rate of accurate classification.

Specificity means the probability of correctly identifying that there is not a problem. Specificity rates help determine which tests can achieve a high rate of accurate classification.

Universal screener for reading means a brief assessment given to all students within a grade level to assess the students' performance on the essential components of reading.

Validity broadly defined means how well something measures what it is supposed to measure or how well a set of scores reflect the intended construct or domain being assessed.