Extensive Assessment Rubric

Name of Assessment:

Criterion	Indicators	Rating	Notes
Assessment Validity: The Early Literacy Assessment Action Team views the assessment to have, in general, acceptable levels of validity as defined by the Joint Committee on Standards for Educational and Psychological Testing (2014)	For all of the below, priority is given to studies published in peer-reviewed research journals. **Evidence Includes:** **[overall] validity**: The degree to which accumulated evidence and theory support a specific interpretation of test scores for a given use of a test. If multiple interpretations of a test score for different uses are intended, validity evidence for each interpretation is needed. **predictive validity evidence**: Evidence indicating how accurately test data collected at one time can predict criterion scores that are obtained at a later time. **[avoidance of] predictive bias**: The systematic under- or over-prediction of criterion performance for people belonging to groups differentiated by characteristics not relevant to the criterion	**DOES NOT MEET**-information is not available for this criteria or information does not demonstrate sufficient evidence. **PARTIALLY MEETS-**partial information is available related to the criterion and/or data demonstrates weak evidence. **MEETS OR EXCEEDS** –most information for the criterion is available and information and data suggest acceptable or strong evidence.	

performance.

consequences: The outcomes, intended and unintended, of using tests in particular ways in certain contexts and with certain populations.

content-related validity evidence: Evidence based on test content that supports the intended interpretation of test scores for a given purpose. Such evidence may address issues such as the fidelity of test content to performance in the domain in question and the degree to which test content representatively samples a domain, such as a course curriculum or job.

construct equivalence: 1. The extent to which a construct measured by one test is essentially the same as the construct measured by another test. 2. The degree to which a construct measured by a test in one cultural or linguistic group is comparable to the construct measured by the same test in a different cultural or linguistic group.

[avoidance of] construct underrepresentation: The extent to which a test fails to

	capture important aspects of the construct domain that the test is intended to measure, resulting in test scores that do not fully represent that construct. **classification accuracy**: Degree to which the assignment of test takers to specific categories is accurate; the degree to which false positive and false negative classifications are avoided. **convergent evidence**: Evidence based on the relationship between test scores and other measures of the same or related construct. **fairness**: The validity of test score interpretations for intended use(s) for individuals from all relevant subgroups. A test that is fair minimizes the construct-irrelevant variance associated with individual characteristics and testing contexts that otherwise would compromise the validity of scores for some individuals.		
Assessment Reliability: The Early Literacy Action Team views the assessment to have, in general, acceptable levels of reliability as defined by the Joint Committee on Standards for Educational and Psychological	**Evidence includes:** **Evidence Includes:** Studies that demonstrate reliability across the entire population for whom the assessment is designed (e.g.,	**DOES NOT MEET**-information is not available for this criteria or information does not demonstrate sufficient evidence. **PARTIALLY MEETS-**partial information is available related	

Testing (2014)	grade levels, nationally representative sample) For all of the below, priority is given to studies published in peer-reviewed research journals. **reliability/precision**: The degree to which test scores for a group of test takers are consistent over repeated applications of a measurement procedure and hence are inferred to be dependable and consistent for an individual test taker; the degree to which scores are free of random errors of measurement for a given group. **internal consistency coefficient**: An index of the reliability of test scores derived from the statistical inter-relationships among item responses or scores on separate parts of a test	to the criterion and/or data demonstrates weak evidence. **MEETS OR EXCEEDS** –most information for the criterion is available and information and data suggest acceptable or strong evidence.	
Constructs: The Early Literacy Assessment Action Team views the assessment to address key areas of literacy development	**Evidence includes**: Assessment may include at least one of the following areas of literacy development **or** other important contributors to literacy development: Phonics, Phonemic Awareness,	**DOES NOT MEET**-information is not available for this criteria or information does not demonstrate sufficient evidence. **PARTIALLY MEETS-**partial information is available related to the criterion and/or data	

MICHIGAN
Department of Education

	Fluency, Comprehension, Vocabulary. Aligned to the Michigan ELA Standards: Reading Comprehension (Literature and Informational Texts), Reading Foundations, Writing (Text Types and Purposes), Listening, Language (Conventions, Knowledge of Language, Vocabulary Acquisitions).	demonstrates weak evidence. **MEETS OR EXCEEDS** –most information for the criterion is available and information and data suggest acceptable or strong evidence.	
Efficiency of Administration: The Early Literacy Assessment Action Team views the time required to conduct the assessment as reasonable and exerting minimum impact on instructional time.	**Evidence Includes:** Assessment administration time for students Assessment administration time for adults	**DOES NOT MEET**-information is not available for this criteria or information does not demonstrate sufficient evidence. **PARTIALLY MEETS-**partial information is available related to the criterion and/or data demonstrates weak evidence. **MEETS OR EXCEEDS** –most information for the criterion is available and information and data suggest acceptable or strong evidence.	

Strengths:

Weaknesses:

Recommend: Not Recommended: