Formative Assessment System for Teachers

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Criterion | Specific Indicators | Rating | Feedback from Reviewers | Tally of rating |
| Validity, Reliability and Consistency in Scoring |  |  |  |  |
| Evidence of test reliability and consistency in scoring | Results of reliability studies are reported for each grade assessment  **Evidence includes:**  The studies are appropriate given the purpose of the measure.  For each grade-level, studies provide evidence of:   * Split-half reliability * Coefficient alpha * Test-retest reliability * Classification consistency | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS-**partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. Correlations demonstrate ranges of .7 or higher. (2) | Would like to see SEM for cut scores and score ranges.  Some of the items are available upon request. | Does not meet –  Partially Meets - I  Meets or Exceeds - I |
|  | Standard error of measurement or standard estimate of error is reported  **Evidence includes:**   * SEM estimates are reported for score ranges and cut-scores. * SEM estimates are reported for score ranges and cut-scores for each assessment (grade-level, form, subtest). | **DOES NOT MEET-**evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS-**partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence**.** (1)  **MEETS OR EXCEEDS --**Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet – I  Partially Meets -  Meets or Exceeds - I |
|  | Inter-rater reliability studies have been conducted. Study sample used to establish inter-rater reliability represents test administrators.  **Evidence includes:**   * Inter-rater reliability studies have been conducted for each grade level and are based on a representative sample of educators who will administer and score the assessment. * Inter-rater reliability coefficients exceed .7. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet –  Partially Meets -  Meets or Exceeds - I |
|  | Studies have been conducted to establish reliability with all subcategories of students who will take the assessment.  **Evidence Includes:**  Studies that demonstrate reliability has been established from scoring samples of students that include: Non-ELLs with and without reading deficiencies and ELLs with and without reading deficiencies. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS-**partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence**.** (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does Not Meet –  Partially Meets – I  Meets or Exceeds - I |
| Alternative forms available for multiple assessments with demonstrated equivalence or comparability | If alternative forms are provided, all forms have demonstrated evidence of equivalence or comparability such as test-retest, parallel form and internal consistency.   * Technical reviews indicate all forms for each grade level have demonstrated evidence of comparability and content specifications.   **Evidence includes:**   * Sufficient forms are provided to allow for progress monitoring between interim assessments. * Split-half reliability. * Coefficient alpha reliability. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence correlations demonstrate ranges of .7 or higher. (2) |  | Does not meet –  Partially Meets -  Meets or Exceeds - II |
| Content and Construct Validity |  |  |  |  |
| Evidence of content and construct validity | Evidence reported to demonstrate the assessment helps correctly identify students with *“significant reading deficiencies”* so that successful remediation and intervention can be provided; studies have been conducted with similar assessments to show that the assessment measures reading ability, not other irrelevant criteria.  **Evidence includes:**   * A clear description is provided that demonstrates the purpose of the assessment is to screen students for reading concerns. * Content specifications for each grade-level, including a complete description of the test content, purpose(s), and intended use(s), and assessment blueprint as appropriate, is provided. | **Rating**  **Does Not Meet –** evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **Partially Meets –** partial evidence was provided related to the criterion and/or data provided demonstrates weak evidence (1)  **Meets or Exceeds** – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence |  | Does not meet –  Partially Meets - I  Meets or Exceeds - I |
|  | Reading levels are reported for passages and how levels were established. Reading levels of assessment passages have been field-tested or have other evidence.  **Evidence includes**:   * Field testing populations should be clear and should mirror the school/district demographics. * Statistics used to establish the reading levels are reported with both ELL and Non-ELL populations. * Findings from a content review by field experts, including teachers in tested grade levels. | **Does Not Meet** – evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **Partially Meets** – partial evidence was provided related to the criterion and/or data provided demonstrates weak evidence. (1)  **Meets or Exceeds** – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence |  | Does not meet –  Partially Meets - I  Meets or Exceeds - I |
|  | If appropriate, findings from alignment studies to demonstrate alignment with Colorado Academic Standards for Language Arts and resolution for any resulting concerns. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet –  Partially Meets - I  Meets or Exceeds - I |
|  | There are studies of construct validity, such as convergent and discriminant analysis, demonstrating correlations of .7 or above. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet –  Partially Meets -  Meets or Exceeds - I |
| Evidence of criterion/predictive validity accurately identifying students with *“significant reading deficiency”* | Evidence reported to demonstrate that the assessment has established criterion and/or predictive validity to correctly identify students with and without a *“significant reading deficiency.”*  ***Evidence includes:***   * A clear definition of the criterion or measure that were used to establish concurrent validity. * Studies with similar assessments that demonstrate the assessment measures reading ability, not other irrelevant criteria. Predictive validity correlations above .7. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet –  Partially Meets - I  Meets or Exceeds - I |
| Determination of cut-scores based upon well-designed pilot study | The assessment has established cut-scores for decision making about students’ “*significant reading deficiency”* using adequate demographics representing (i.e., 10%ELL and 25% F/R lunch), appropriate criterion assessment, adequate sample size, and appropriate statistics.  **Evidence indicates**:   * Includes a description of the process used to establish the cut points. * A full description of the norming sample. * The norming sample is a large representative national sample of students at the same grade level and is representative of the testing population according to gender, ELL status, special needs status and F/R lunch status. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and 2data provided suggests acceptable or strong evidence. (2) | Would like to see more information on subgroups and include more subgroups | Does not meet –  Partially Meets - I  Meets or Exceeds - I |
|  | Studies of classification accuracy analysis provide evidence that the measure appropriately identifies students as indicated in the description of purpose of the assessment, demonstrating values that exceed .8 or higher. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet –  Partially Meets -  Meets or Exceeds - II |
|  | Acceptable, recognized procedures are followed for setting cut-scores. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet –  Partially Meets -  Meets or Exceeds - II |
|  | SEM estimates are reported for cut-scores with guidance for score interpretation. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence.(1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet –  Partially Meets - I  Meets or Exceeds - I |
| Universal Design | Evidence reported to demonstrate that the assessment has cultural validity, that fairness and bias issues have been addressed; the assessment is accessible to all learners, considering minimizing language load; the format is not a barrier to student performance.  **Evidence includes:**   * Addressed issues of equity of utility for all populations**.** * Results of bias reviews and plans that have addressed any concerns. * At least two to three types of classification, reliability, and validity study data have been disaggregated by subgroups and meet the criteria. * Culturally diverse students were included throughout the entire process of test development. For example in the samples of pilot students, in cognitive interviews, etc. * The content of the reading materials does not favor mainstream culture. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence.(0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) | Not by subtype, but explanation of language used in passages satisfactory | Does not meet –  Partially Meets - II  Meets or Exceeds - |
| Third party evaluation conducted | Evidence reported to demonstrate that an independent, qualified third party has provided a thorough and unbiased evaluation of the quality of the assessment. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet –  Partially Meets -  Meets or Exceeds - II |
| **Administration & Scoring** |  |  |  |  |
| Standardization of materials and procedures for administration | Administration protocol is scripted and provides precise guidelines; administration windows are clearly identified; materials are provided or clear guidelines are provided if materials are to be created; includes both electronic and hard copy administration manual that is clear and concise. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) | Could not find the online admin manual | Does not meet –  Partially Meets - I  Meets or Exceeds - I |
| Efficiency of administration | The amount of time needed to administer the assessment is reasonable and balanced to the information provided. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet –  Partially Meets -  Meets or Exceeds - II |
| Efficiency of scoring | The amount of time needed to score the assessment is reasonable and balanced to the information provided; computer-assisted scoring is available; procedures for calculating scores are clear; scores can be stored and reported electronically. |  |  | Does not meet –  Partially Meets -  Meets or Exceeds - II |
| Accommodations clearly stated and described for students with disabilities and students with special needs (504, etc.) | The differing needs of students with disabilities are specifically addressed.  **Evidence includes:**   * Any accommodations do not compromise the interpretation or purpose of the test. * Specific administration guidelines are provided for implementing any accommodations. * How to address accommodations is specifically addressed in the training materials or program. * Suggested accommodations are research or evidence-based. | **DOES NOT MEET-**evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS-**partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS –**most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) | Not found  The only current accommodation is using the ipad for students with differing needs or disability. | Does not meet – I  Partially Meets - I  Meets or Exceeds - |
| Accommodations clearly stated and described for Second Language Learners | The accommodations directly address the linguistic needs of the student.  **Evidence includes**:   * Any accommodation does not compromise the interpretation or purpose of the test. * Specific administration guidelines are provided for implementing any accommodations. * How to address accommodations is specifically addressed in the training. * Suggested accommodations are research or evidence-based. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) | Cannot find  Does not have accommodations for ELLS; such as dialect or articulation errors | Does not meet – II  Partially Meets -  Meets or Exceeds - |
| **Utility** |  |  |  |  |
| Scores are easily interpreted to determine a *“significant reading deficiency”* | Scores clearly specify whether a student is categorized as having a *“significant reading deficiency”.*  **Evidence includes:**   * Score ranges or a scale is provided. * Guides for interpretation of scores are provided. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence.(0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) | p. 49 | Does not meet –  Partially Meets -  Meets or Exceeds - II |
| Cost effective: Materials, administration costs including personnel, scoring, and training | Materials are provided or easily accessible; time away from instruction is minimal; no additional personnel required; all costs inclusive including any additional data platform or storage costs; minimal data entry is required. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence.(0)  **PARTIALLY MEETS** -partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) |  | Does not meet –  Partially Meets -  Meets or Exceeds - II |
| Reports provide guidance for interpretation useful to educators, administrators, and parents | Information is displayed in a format and language that is understandable to educators, administrators and parents;   * Data reports are easily read and interpreted. * Clear description of how to interpret results. * Reports provide trajectory for student progress. * District, school, classroom, and student reports provided. * Reports available in real-time. * Reports can be exported to data-base formats. * Reports available in languages other than English. * Customer service is available provided for users. | **DOES NOT MEET**-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  **PARTIALLY MEETS**-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  **MEETS OR EXCEEDS** –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2) | Limited visuals accessed online and seem helpful. It is not clear if there is a district level report, if it is compatible with data storage systems, or available in other languages | Does not meet –  Partially Meets - I  Meets or Exceeds - I |

Strengths:

1. Computer adaptive, price, easy to administer
2. Cost-effective, currently is free with contributing research; estimated costs are $2 per student
3. Well-researched from the university of Minnesota

Weaknesses:

1. It would be helpful to see more information about accommodations, information about reports and compatability
2. No accommodations for dialect, articulation for ELLs
3. Does not disaggregate subgroups

Recommend: X X Not Recommended: