Compilation of reviewers’ comments and ratings for DRA2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 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)** | The data was grouped together for co-efficient alpha, it’s not representing individual grade levels.1-­‐Sample 112Students fortest-­‐retestmean scoreswere usedrather thanstudents’actualscores, alpha(internal consistency) not byreliability, grade levelNo grade level specified data.Coefficient alpha is reported by passage level, not grade level.Small sample size for test-retest reliability.Studies are poorly constructed (small sample size, subjective) and evidence is very weak or does not support claims.Coefficient Alpha provided for oral fluency and comprehension for DRA2 Levels 4-80; oral fluency range 0.5-0.8; comprehension range 0.5-0.8 Test-retest reliability is high, with little error associated with time sampling; sample range 90 112; ranges .93-.99 No evidence in technical report for split-half reliability or classification consistency Examiners must follow assessment guidelines | **Does Not Meet**—1,1,1**Partially meets—**1,1**MEETS OR EXCEEDS—** |
|  | 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**—most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. Correlations demonstrate ranges of .7 or higher. (2) | No evidenceThere is no evidence of standard error of measurement.Does not report.Not found | **Does Not Meet—**1,1,1,1**Partially Meets--** **Meets or Exceeds--** |
|  | 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. Correlations demonstrate ranges of .7 or higher. (2) | The inter-rater reliability co-efficient data reported was not exceeding .70-­‐inter-­‐ratersample was30students. Kappavalues of .57 forfluency, .65for comprehension,not by gradelevelCoefficient – first order is.5 – fluency.6 – comprehensionSample of educators was not representative.Evidence is based on small study sample of atypical test administrators. Coefficients do not meet threshold.Fluency. 66 Comprehension .72Raters were existing users of the DRA2Rater expert v. non-expert scrorersFluency 79%Comprehension 89% | **Does Not Meet—**1,1,1,1**Partially meets**—1**MEETS OR EXCEEDS**— |
|  | 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.  | No evidence of subcategories0Not by subgroupNo subcategories.No evidenceGender, ethnicity, free or reduced lunch, district type, grade level data sharedELL information not listed on APPENDIX A | **Does not meet**—1,1,1,1**Partially meets--**1 |
| 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)  | 0,2-­‐4 per levelThe technical manual notes differences in two cases (p.38).No research that shows alternative forms are equivalent.Evidence for passage levels, not grade levels.Multiple forms for progress monitoring exist, but evidence of reliability is lacking.2-4 passages available at each DRA2 level; fiction and nonfiction Passage equivalency-very little variability between passage variation; no significant differences between the difficulties of passages at the various levels. The passages at each level are equivalent and can be used interchangeably | **Does Not Meet**—1,1**Partially Meets**—1,1**Meets or Exceeds**--1 |
| **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.
 | **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)  | No clear description providedSpecifications per grade level are not included0-­‐constructvalidity (facevalidity) wasmeasured based on howteachers feltabout whetheror not the testmeasured whatit was supposedto measure,based onteacher perception(66teachers)The purpose for the test is not to determine students with SRDs.Multiple forms for progress monitoring exist, but evidence of reliability is lacking.Clear description provided that demonstrates the purpose of the assessment, p. 10 Description of test content, purpose and intended goals evident for DRA2 levelsAssessment is designed to measure reading growth, not screen students for significant reading deficiency.General reading ability is measured, but is not useful for designing successful remediation and interventionNo grade-level equivalencies for text levels | **Does not meet**—1, 1,1,1**Partially Meets**—1,  |
|  | 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. (2)  | No evidence found0–can’tfindNo evidence foundTeacher description of what a borderline reader should be able to do.- evaluative rather than empirical.No evidence of procedure for establishing passage levels and no evidence of passage levels being field-tested on populations that mirror demographics Based on Reading Recovery Multiple passages for each reading level Teacher ratings of the DRA2:.Measurement of Reading and Usefulness- range 81-97% agreed useful for measurement of reading ELL information is not evident11 teachers made the decision for passage levels for K-2. 9 teachers for 3-5 | **Does not meet**—1, 1,1,1**Partially meets**—1**Meets or exceeds**-- |
|  | 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) | Validity is reported at .61-­‐Teachers usedCCSS to identifyWhat studentsShould be ableTo do or not doWith given grade levelTo establishprof.benchmarksPublisher believes they are aligned to CAS and CCSSTeachers identified CCSS standards that students should be able to do at time of assessmentMinimally and/or weakly addresses alignment to foundational skills because it’s all based on word analysis No evidence provided for alignment with Colorado Academic Standards | **Does not meet**—1**Partially meets**—1,1,1**Meets or exceeds---1** |
|  | There are studies of construct validity, such as convergent and discriminant analysis, demonstrating correlations of .7 or above. |  | Validity is reported at .6Internal validation (factor analysis) is not sufficient for construct validityDefines construct validity, but does not provide clear data or evidence to demonstrate correlation of .7 or above.Fluency & Comprehension .41, Fluency & Total .78, Comprehension & Total .89Majority of fluency and comprehension constructs compared to external assessments fell below .7 and sample size was extremely smallDefines construct validity, but does not provide clear data or evidence to demonstrate correlation of .7 or above.  | **Does not meet—1,1,1****Partially Meets--1** |
| 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) | Reported at .6 and .630-­‐123 students in 10schools(notsure aboutgrades), toDibels and Grade,31 kidsIn sample when123 totalSample is broken townTo grades 1-­‐3,No Kindergartendata.Sample size is 31.Predictive validity correlation is below .7.Very small sample (31) and below threshold on grades 1-3 comprehension and fluency predictive validity.Gray’s Oral Reading Test (GORT) DIBELS Oral Reading Fluency (DORF), Gates MacGinitie Reading Test Fall DRA2 scores predicted Spring GRADE Comprehension and DIBELS Oral Fluency scores. Correlation coefficients ranged from .51 to .89 Limited sample at grades 1-3- correlation moderate for luency and high for comprehension Correlations for grades 4-6 were large Overall predictive of other measures of reading, p. 55 | **Does not meet**—1,1,1,1**Partially Meets**--1 |
| 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 data provided suggests acceptable or strong evidence. (2) | Adequate demographic breakdown is not includedThere is not a full description of the norming sample (p.49)0-­‐based onWhat teachersBelieve andTeacher perceptionsof gradelevel readingskills forstudents. Cutpoints aredeterminedby “expertjudgments”Sample size is 31.Predictive validity correlation is below .7.Highly subjective establishment of cut-scores. Cut-scores were established by a small group of teachers who described the attributes of a borderline proficient student for each grade level. Subjective agreement was used to correlate the designated cut-points to the DRA2 national data bank. Furthermore, subjective feedback from classroom teachers whether they felt the cut-scores matched their students was gathered.Establishment of cutpoints forDRA2 book levels shared Clear cut points within tech manual,p. 50 Appendix A Scores of DRA2 and establishing SRD , p. 28 N=1676 student in grade K-8 Word Analysis- ELL populated representedThere are cut scores but there were created by a teacher work group not empirical evidence based on a norming sample not based on student data but on teacher perceptionNo disaggregated student subgroups | **Does Not Meet**—1,1**Partially Meets**—1,1,1 |
|  | 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) | No evidence0-­‐31kids, GRADE and Dibelspredictive validity, notsufficient todetermineif studentswere classified,not sufficientto determinestudents wereclassified accurately.Did they everValidate theirScale (basedOn teacherperceptions)withanother test?(.69,.65 GORTWith 66students)No evidence of studies of classification accuracy analysis.No evidence Measurement of Reading values range from .60-.92 Usefulness of DRA2 values range from .58-.76 Teachers and literacy professionals –criteria of MA degree with specialization in literacy 4 years of successful teaching under a professional license | **Does Not Meet**—1,1,1,1**Partially Meets**—1 |
|  | 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) | They confirmed benchmark scores, but not the cut scores for SRD0teacherjudgmentSubjective not empirical.Highly subjective establishment of cut-scores. Cut-scores were established by a small group of teachers who described the attributes of a borderline proficient student for each grade level. Subjective agreement was used to correlate the designated cut-points to the DRA2 national data bank. Furthermore, subjective feedback from classroom teachers whether they felt the cut-scores matched their students was gathered.Benchmark Setting Procedure evident, p.47, p. 49 (d) | **Does Not Meet**—1,1,1,1**Partially Meets**—**Meets or Exceeds**--1 |
|  | 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) | No evidence0Not foundNo evidence of SEMSEM not recorded.P. 40 contains test-retest reliability including standard deviations | **Does not meet**—1,1,1,1,1 |
| 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) | 0-­‐Samples sizesAre not sufficient, noEvidence foundTo address bias,fairness, etc.No information provided for student samples by subgroup.No bias studies provided.Small minority sample size.Small sample that does not report cultural validity or address test biasCultural validity is not addressed | **Does Not Meet**—1,1,1,1**Partially Meets**—1 |
| 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) | 0–not foundNo evidence presentNo qualified third party evident.Bias and thoroughness of evaluation by 3rd party unknown. Results not evident.Teachers and literacy professionals –criteria of MA degree with specialization in literacy Not evident if unbiased evaluators | **Does Not Meet**—1,1,1,1**Partially Meets**—1 |
| 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) | 1-­‐Given 3x a year,electronicmaterials anddatabase, someconcerns regardingstandardprocedures,example, teacher is nottold when tosupply words,by astandardizedapproach (3second pause)Some evidence of scripted protocol and standardized scoring of oral reading fluency, but not provide precise or concise guidelines. No clear admin. window evident.Administration protocol is scripted Fall, Winter, Spring admin windows Materials are provided CD versions of Blackline Masters and hard copies | **Partially meets**—1,1,1**Meets or Exceeds**—1,1 |
| 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) | 0-­‐10-­‐15conferencetime,totaltime40-­‐80min/childAdministered individually can take up to 40 minutes per studentAdministration time is extensive and not balanced to information provided.Depending on level, time for administration varies No time limit; primary grades impact on instructional time would be minimal | **Does Not Meet**—1,1,1,1**Partially Meets**—1 |
| 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—**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) | 1-­‐Onlinedatabase, fluency is scoredon a ratingdetermined bya teacher.Rubrics areused. Requiresteacherfamiliarityscoring forcomprehension,same forwrittenanswers.Extensive admin. time, no electronic scoring, rubric laborious and subjectiveNeed for calibration for scoring Computer-assisted scoring not available Procedures for calculating on each blackline master for each DRA2 level Online data management system available | **Does Not Meet**—1**Partially Meets**—1,1,1,1 |
| 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) | Accommodations are addressed, but not specifically addressed in regards to research or evidence basedp. 93 –Accommodations need to align with IEP goals | **Does not meet--**1**Partially Meets**—1**Meets or Exceeds**—1,1,1 |
| 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) | P. 93- accommodation allowed is “Reading the directions in the student’s native language (if available) | **Does Not Meet**—1**Partially Meets**—**Meets or Exceeds**—1,1,1,1 |
| 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**-­‐evidenceWas not providedFor this criteriaOr informationDoes notdemonstrateevidence.(0)**PARTIALLY****MEETS**-­‐partialEvidence wasprovidedrelated to thecriterion and/ordata provideddemonstratesweak evidence.(1)**MEETS OR****EXCEEDS**–mostinformationfor the criterionis provided.Information andData providedSuggests. (2) | 0-­‐Designed toAssess astudent’sindep/readinglevel, notdesigned to show level ofriskNo evidence providedLack of predictive validity.Small sample size.Scores specify advanced, independent, instructional, intervention levels DRA2 continuum supports interpretation of scoresValidity of setting cut scores is weak so difficult to interpret scores based on weak data | **DOES NOT MEET**—1,1,1,1**PARTIALLY MEETS**—**MEETS OR EXCEEDS**--1 |
| 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**-­‐evidenceWas not providedFor this criteriaOr informationDoes notdemonstrateevidence.(0)**PARTIALLY****MEETS**-­‐partialEvidence wasprovidedrelated to thecriterion and/ordata provideddemonstratesweak evidence.(1)**MEETS****OR****EXCEEDS**–mostinformationfor thecriterion isprovided.InformationAnd dataprovidedsuggestsacceptable orstrong evidence.(2) | 2Time away from instruction is not minimalPotential personnel required to cover classrooms while assessingTime away from instruction is significant.Administration time and personnel is extensive.$320 for comprehensive package Additional $90.97 per year per classroom for access to DRA2 and EL2. Heavy data entry required | **DOES NOT MEET**—1,1**PARTIALLY MEETS**—1,1**MEETS OR EXCEEDS**--1 |
| 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**-­‐evidencewas notprovided for thiscriteria orinformation doesnot demonstrateevidence.(0)**PARTIALLY****MEETS**-­‐partialEvidence wasprovided relatedto the criterionand/or dataprovideddemonstrates weak evidence.(1)**MEETS OR****EXCEEDS**–mostInformation forthe criterion isprovided.Information anddata providedsuggests acceptable orstrong evidence.(2) | 1-­‐Reports providedat all levels, maynot providesufficient infobeyond instr.readinglevelWould need to purchase the online portion to have access to the reportsUnable to find trajectory reportsNo trajectory provided.Reports for individuals and classroom only.Data reports are do not provide specificity for practical use.Online reports available Broken down by level of mastery and component of reading Historical reports available District, school, classroom, and student reports provided Online customer serviceNot evident Reports in other languages real time exported to data-based formats and trajectory | **DOES NOT MEET**—1**PARTIALLY MEETS**—1,1,1,1**MEETS OR EXCEEDS**-- |

**STRENGTHS:\_\_** Aligned to Common Core, **\_**Accommodations were clearlydescribed**.\_** Alignment toCommon Core State Standards; Representative of typical classroom instruction; Multiple texts at multiple levels; Self reported that teachers like it

**WEAKNESSES:\_**No predictive validity, Cut scores were not determined through a reliable process**,** Cut scores were not determined by a norming sample of student performance, but by teacher perception**.\_\_** Sample sizes were small. Self-reported that teachers like it.\_ Lack of research that provides evidence of reliability, validity, and utility. \_\_ Time-consuming and costly 3. Not designed as a screening tool; Time spent administering/scoring; Ease of use questionable; highly dependent on professional development to ensure calibration

**Recommend\_\_\_\_\_\_\_ Not Recommend\_\_X\_X\_X\_X X**