

DRA2 Spanish

Spanish Assessment not Considered Because English Assessment did not meet the Criteria

Criterion	Specific Indicators	Rating	Feedback from Reviewers	Tally of rating
Validity, Reliability and Consistency in Scoring				
Evidence of test reliability and consistency in scoring	<p>Results of reliability studies are reported for each grade assessment</p> <p>Evidence includes: The studies are appropriate given the purpose of the measure. For each grade-level, studies provide evidence of:</p> <ul style="list-style-type: none"> • Split-half reliability • Coefficient alpha • Test-retest reliability • Classification consistency 	<p>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)</p>	<p>Internal Consistency Reliability and Test-Retest Reliability were conducted. All measures are between .5-.9, indicating acceptable evidence. They are not reported by grade, they are reported by levels.</p>	<p>Does Not Meet– Partially Meets – Meets or Exceeds - II</p>
	<p>Standard error of measurement or standard estimate of error is reported</p> <p>Evidence includes:</p>	<p>DOES NOT MEET-evidence was not provided for this criteria or information</p>	<p>Page 8 of the RFI provides standard deviation for scores on the test and re-test pilot studies separated by grade band and sub-test</p>	<p>Does not meet – I Partially Meets –</p>

	<ul style="list-style-type: none"> 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). 	<p>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)</p>	<p>Standard error of measurement not reported in technical manual for score ranges and cut-scores.</p>	<p>Meets or Exceeds - I</p>
	<p>Inter-rater reliability studies have been conducted. Study sample used to establish inter-rater reliability represents test administrators.</p> <p>Evidence includes:</p> <ul style="list-style-type: none"> 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. 	<p>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</p>	<p>p. 10 of the RFI has inter-rater reliability of above point 8 and the requirement is .7 Inter-rater reliability is evidenced however, the grade levels are combined (e.g. 2-3, 4-5) and there are not individual reliability scores per grade level.</p>	<p>Does not meet – Partially Meets – I Meets or Exceeds - I</p>

		strong evidence. (2)		
	<p>Studies have been conducted to establish reliability with all subcategories of students who will take the assessment.</p> <p>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.</p>	<p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>Score of 2 because the test was designed for a population of Spanish speaking ELLs and field testing was conducted with students from literacy squared classrooms across Colorado and other states in US, the reliability assessments were conducted with subcategories of ELL students with and without reading deficiencies. Non-ELLs are not the target population for the test and are not an appropriate subcategory for this score Info found in RFI and technical report found online. http://tinyurl.com/DRA-EDL-Technical-manual</p> <p>No detail description of studies defining non-ELLs with and without reading deficiencies and ELLs with and without reading deficiencies was done with the EDL2. They were done on the DRA2 and applied to the EDL2. The benchmark and cut-scores for the EDL2 are also taken from DRA2</p>	<p>Does Not Meet –</p> <p>Partially Meets – I</p> <p>Meets or Exceeds - I</p>
Alternative forms available for multiple assessments with demonstrated equivalence or comparability	<p>If alternative forms are provided, all forms have demonstrated evidence of equivalence or comparability such as test-retest, parallel form and internal consistency.</p> <ul style="list-style-type: none"> Technical reviews indicate all forms for each grade level have demonstrated evidence 	<p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS-partial evidence was</p>	<p>In the EDL technical report found online and the RFI p. 280 (proposal) the company states that “the assessment also includes alternative forms for multiple assessments with demonstrated comparability. In addition, I (J.B.) as a fully proficient bilingual and teacher able to determine text levels examined the assessments at various levels to determine that there were multiple forms that were comparable at different reading levels. It appears there may be two forms</p>	<p>Does Not Meet – I</p> <p>Partially Meets – I</p> <p>Meets or Exceeds -</p>

	<p>of comparability and content specifications.</p> <p>Evidence includes:</p> <ul style="list-style-type: none"> • Sufficient forms are provided to allow for progress monitoring between interim assessments. • Split-half reliability. • Coefficient alpha reliability. 	<p>provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>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)</p>	<p>because there are two books per level in the kit. However, no evidence of equivalence or comparability such as test-retest, parallel form and internal consistency was provided. Furthermore, there is no mention that the two books at each level indicate alternative forms, they are not labeled as alternative forms.</p>	
Content and Construct Validity				
Evidence of content and construct validity	<p>Evidence reported to demonstrate the assessment helps correctly identify students with “<i>significant reading deficiencies</i>” 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.</p> <p>Evidence includes:</p> <ul style="list-style-type: none"> • A clear description is provided that demonstrates the purpose of the assessment is to screen students for reading concerns. • Content specifications 	<p>Rating</p> <p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR</p>	<p>There is a section specifically covering “content validity” on p. 292-294 of the RFI. There is also a section covering construct validity on p. 298-306 of the RFI. These sections address the fact that this assessment correctly identifies students with an SRD and that the assessment measures reading ability. Their extensive factor analysis on page 304-305 of the RFI also provides further evidence of the construct validity of this test. A score of 1 however is given, because each grade level is not specified with respect to content validity. Evidence reported demonstrated that the assessment may help correctly identify students with SRD. The EDL2 uses the cut-offs from the DRA2, they are not created</p>	<p>Does Not Meet –</p> <p>Partially Meets – II</p> <p>Meets or Exceeds -</p>

	<p>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.</p>	<p>EXCEEDS –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>using the EDL2 in and of itself.</p> <p>Content specification for each grade-level, including a complete description of the test content, purpose(s), and assessment blueprint is provided for the most part yet, the content specifications are not specific per grade-level.</p>	
	<p>Reading levels are reported for passages and how levels were established. Reading levels of assessment passages have been field-tested or have other evidence.</p> <p>Evidence includes:</p> <ul style="list-style-type: none"> • 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. 	<p>Does Not Meet – Evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>Partially Meets – partial evidence was provided related to the criterion and/or data provided demonstrates weak evidence.</p> <p>Meets or Exceeds – most</p>	<p>Field testing was not done with the EDL2 specifically. Testing was done with the DRA2 and transferred to the EDL2. There is no given information on how the levels are created for the EDL2 and there is no evidence stating how the passages were field tested.</p> <p>Evidence provides limited information about demographics (on page 28-29). There is also no disaggregated information on how each population did.</p>	<p>Does Not meet –</p> <p>Partially Meets – I</p> <p>Meets or Exceeds - I</p>

		information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)		
	If appropriate, findings from alignment studies to demonstrate alignment with Colorado Academic Standards for Language Arts and resolution for any resulting concerns.	<p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>clearly aligned to standards, but research findings are not shown to be up to date and aligned with CAS</p> <p>There is evidence of alignment with the CCSS and the DRA2 however there is no evidence of a parallel alignment with the EDL2.</p>	<p>Does Not Meet –</p> <p>Partially Meets – II</p> <p>Meets or Exceeds -</p>
	There are studies of construct validity, such as convergent and discriminant analysis, demonstrating	<p>DOES NOT MEET-evidence was not provided for this criteria or information</p>		<p>Does Not Meet –</p> <p>Partially Meets –</p>

	<p>correlations of .7 or above.</p>	<p>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>		<p>Meets or Exceeds - II</p>
<p>Evidence of criterion/predictive validity accurately identifying students with “<i>significant reading deficiency</i>”</p>	<p>Evidence reported to demonstrate that the assessment has established criterion and/or predictive validity to correctly identify students with and without a “<i>significant reading deficiency.</i>”</p> <p>Evidence includes:</p> <ul style="list-style-type: none"> • 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 	<p>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</p>	<p>Evidence for concurrent and predictive validity are provided. On the RFI p. 294 the correlations are an average of 0.7 for one test, but range from .51-.58 for 2 others</p> <p>The criteria states the predictive validity must be above .7 and the reported predictive validity was .51 (pg. 16).</p>	<p>Does Not Meet –</p> <p>Partially Meets – II</p> <p>Meets or Exceeds -</p>

	correlations above .7.	data provided suggests acceptable or strong evidence. (2)		
Determination of cut-scores based upon well-designed pilot study	<p>The assessment has established cut-scores for decision making about students' "<i>significant reading deficiency</i>" using adequate demographics representing (i.e., 10% ELL and 25% F/R lunch), appropriate criterion assessment, adequate sample size, and appropriate statistics.</p> <p>Evidence indicates:</p> <ul style="list-style-type: none"> • 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. 	<p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS –most information for the criterion is provided. Information and 2data provided suggests acceptable or strong evidence. (2)</p>	<p>There is evidence of a correlational study between EDL2 and the CSAP (field studies). The sample sizes are small for each grade level, even as low as n=36.</p> <p>Demographics were given for the sample however, there was no mention of how the sample was a representation of national statistics at the same grade levels / age levels. The statistics were based solely on Colorado students. Furthermore, it appears that the field study was done between the DRA2 and CSAP, not the EDL2 and the CSAP.</p>	<p>Does Not Meet –</p> <p>Partially Meets – I</p> <p>Meets or Exceeds - I</p>
	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.	<p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS-partial</p>	There is no evidence of classification accuracy with evidence to say the EDL2 appropriately identifies students as indicated in the description of the purpose of the assessment with	<p>Does Not Meet – I</p> <p>Partially Meets –</p> <p>Meets or Exceeds - I</p>

		<p>evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>values that exceed .8 or higher.</p>	
	<p>Acceptable, recognized procedures are followed for setting cut-scores.</p>	<p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>It is stated in the narrative that reading, whether in Spanish or English, requires the same skills at the same grade level so the cut-scores for EDL2 are from the DRA2. It is unclear if these data come from the field study or otherwise.</p>	<p>Does Not Meet –</p> <p>Partially Meets – I</p> <p>Meets or Exceeds - I</p>

	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 SEM estimates are reported.	Does Not Meet – I Partially Meets – 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: <ul style="list-style-type: none"> Addressed issues of equity of utility for all populations. 	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.	p. 4 of Their technical manual –p 285 of RFI provides description of cultural validity with detail, but not with a description of bias reviews. More evidence needs to be provided to receive a 2 Minimal evidence is presented to demonstrate that the assessment has cultural validity, that fairness and bias issues have been addressed. This is most evident in the fact that many statistics are taken	Does Not Meet – Partially Meets – II Meets or Exceeds -

	<ul style="list-style-type: none"> • 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. 	<p>(1) MEETS OR EXCEEDS –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>directly from the DRA2 and not from implementation of the EDL2.</p>	
<p>Third party evaluation conducted</p>	<p>Evidence reported to demonstrate that an independent, qualified third party has provided a thorough and unbiased evaluation of the quality of the assessment.</p>	<p>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.</p>	<p>A statement in the narrative addresses this and the technical form was from PRES, their third party reviewer.</p>	<p>Does Not Meet – Partially Meets – Meets or Exceeds - II</p>

		Information and data provided suggests acceptable or strong evidence. (2)		
Administration and 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.	<p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>		<p>Does Not Meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - II</p>
Efficiency of administration	The amount of time needed to administer the assessment is reasonable and balanced to the information provided.	<p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p>	Teachers must personally administer the test, therefore it will take more time away from teaching than a computer based assessment, although teacher administration will provide first-hand information of the students' reading behaviors that could be useful for instructional planning in	<p>Does Not Meet-</p> <p>Partially Meets – II</p> <p>Meets or Exceeds -</p>

		<p>PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>addition to the usefulness of the screener</p> <p>The assessment is time intensive and the manual is not clear on how long it would take for all the tasks that are necessary. If there are 30 tasks, they take at least 2 minutes each, equaling 60 minutes total. Teachers are also required to make copies of all materials.</p>	
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.		<p>Teachers must personally score the test therefore it will take more time than a computer generated scoring system.</p> <p>Teachers are required to use a paper/pencil entry first then they are required to transfer data to a computer data management system.</p>	<p>Does not Meet –</p> <p>Partially Meets – II</p> <p>Meets or Exceeds-</p>
Accommodations clearly stated and described for students with disabilities and students with special needs (504, etc.)	<p>The differing needs of students with disabilities are specifically addressed.</p> <p>Evidence includes:</p> <ul style="list-style-type: none"> Any accommodations do not compromise the interpretation or purpose of the test. Specific administration guidelines are provided for implementing any 	<p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS-partial evidence was provided related to the criterion</p>	<p>on page 281 of the RFI the manual states that “educators should make any accommodations called for in a student’s 504 plan that are relevant.”</p> <p>The accommodations listed do not provide evidence for their use, but they are allowable.</p> <p>Accommodations are not specifically mentioned for the EDL2.</p>	<p>Does Not Meet – I</p> <p>Partially Meets- I</p> <p>Meets or Exceeds -</p>

	<p>accommodations.</p> <ul style="list-style-type: none"> • How to address accommodations is specifically addressed in the training materials or program. • Suggested accommodations are research or evidence-based. 	<p>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>		
Accommodations clearly stated and described for Second Language Learners	<p>The accommodations directly address the linguistic needs of the student.</p> <p>Evidence includes:</p> <ul style="list-style-type: none"> • 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. 	<p>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>	<p>The target population for this test is native Spanish speakers. Therefore, the fact that the test is in Spanish and is written to be inclusive of various Spanish dialects, the test itself directly addresses the linguistic needs of second language learners.</p> <p>Only given for DRA2 but not for EDL2.</p>	<p>Does Not Meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - II</p>
Utility				
Scores are easily	Scores clearly specify	DOES NOT	Yes, a chart is	Does Not

<p>interpreted to determine a "significant reading deficiency"</p>	<p>whether a student is categorized as having a "significant reading deficiency".</p> <p>Evidence includes:</p> <ul style="list-style-type: none"> • Score ranges or a scale is provided. • Guides for interpretation of scores are provided. 	<p>MEET-evidence was not provided for this criteria or information does not demonstrate evidence.(0)</p> <p>PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>presented to differentiate classes of students however, there are no guides for interpretation of the scores.</p>	<p>Meet –</p> <p>Partially Meets – I</p> <p>Meets or Exceeds - I</p>
<p>Cost effective: Materials, administration costs including personnel, scoring, and training</p>	<p>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.</p>	<p>DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence.(0)</p> <p>PARTIALLY MEETS -partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence (1)</p> <p>MEETS OR</p>	<p>Teachers must personally administer the test therefore it will take more time away from teaching than a computer based assessment.</p>	<p>Does Not Meet –</p> <p>Partially Meets – II</p> <p>Meets or Exceeds -</p>

		EXCEEDS –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)		
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; <ul style="list-style-type: none"> • 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)	There are extensive reports available for DRA2 starting on page 149 of the RFI and on pg 184 specifically they show that the reports are available for EDL Reports but minimal guides.	Does Not Meet – Partially Meets – I Meets or Exceeds - I

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
Translation and adaptation				

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
procedure				
<p>1. Translation has been provided by highly qualified personnel.</p>	<p>Provide documentation on the translation team used to translate and adapt the test. Include the qualifications of the individuals who translated the test. The translation team should preferably include:</p> <ul style="list-style-type: none"> • translators who are native speakers in the target language • specialists in reading in the target language • bilingual educators (not to be confused with English as a Second Language (ESL) teachers or English as a Foreign Language (EFL) teachers or teachers of Spanish as a foreign language) in the target language. 	<p>DOES NOT MEET- evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS- partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>Pgs 22 and 23 of the technical manual found on-line http://tinyurl.com/DRA-EDL-Technical-manual there is an extensive description of the translation team, their qualifications and the process. In addition, careful examination by myself (J.B.) a fully proficient bilingual educator for 16 years, shows clear evidence of quality translations of the testing materials.</p> <p>It is not clear how this assessment is translated or if it is translated. If it is not translated it is not clear where the materials come from.</p>	<p>Does not meet –</p> <p>Partially Meets - I</p> <p>Meets or Exceeds-I</p>
<p>2. Pilot test sampling appropriately considers language diversity</p>	<p>The translated test was piloted with a representative sample of speakers of the target language in the United States.</p>	<p>DOES NOT MEET- evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS- partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>age 309-310 of the RFI lists the ethnicity, free and reduced lunch status and ELL status of the sample used in the field testing of the EDL. In addition pg 23 of the technical manual (url above) shows that students were field tested in various states across the country, New York to California.</p> <p>The field tests include students from language diverse backgrounds however, this is not clearly discussed in relevance to a national norm. The sample was composed of 602</p>	<p>Does not meet –</p> <p>Partially Meets - I</p> <p>Meets or Exceeds- I</p>

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
			students in K-6 classrooms from Colorado.	
3. Consistency of appearance between the English language and the target language version of the test	Formatting should remain consistent with the English language test version. Specifically, the font size of a translated test version should not be smaller than the English version. General ideas should be consistent with the English language test version.	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
Criterion	Specific Indicators	Ratings		Notes
Psychometric and measurement considerations:				
1. Construct validity for translated test versions	Provide documentation to demonstrate that the test specifically identifies students with a <i>“significant reading deficiency”</i> in their native language. (i.e., test developers consider what constitutes a proficient reader in the target language rather than directly translating the measures of a proficient reader in English into the target language). Evidence is	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	Most evidence is provided for the DRA2 and not the EDL2.	Does not meet – Partially Meets - I Meets or Exceeds- I

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
	provided that the reading constructs measured by the test are relevant to the target language. As appropriate, information is reported on the procedures used to screen, select, and adapt the items of the test so that they are relevant and applicable to the target language.	and data provided suggests acceptable or strong evidence. (2)		
3. Demonstrated comparability	Evidence is provided on the psychometric comparability of measures in English and measures in the target language.	<p>DOES NOT MEET- evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS- partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	Pgs. 306-307 of the RFI show that convergent validity was investigated. The developers compared EDL2 to CSAP and DRA2 in English ranging from .51 - 0.8. There is not a requirement for the correlation statistic for this clause . Most evidence is given for the DRA2 but not the EDL2. There is no comparable data between the two assessments.	Does not meet – Partially Meets - I Meets or Exceeds- I
4. Documentation on the interpretation of scores and the scaling of scores	Scaling information is provided to ensure appropriate interpretability of scores across language versions of the test so that educators and administrative officials know how to correctly interpret the scores obtained by the students in the translated version of the test.	<p>DOES NOT MEET- evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p>PARTIALLY MEETS- partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS –</p>	N/A the test is scored for them, no scaling is necessary	Does not meet – Partially Meets - Meets or Exceeds- II

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
	For example, do teachers need to scale the score of the translated test version in order to compare it with the English language version? If so, what kind of documentation is provided to assist teachers in this scaling process?	most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)		
5. Evidence provided regarding investigation into potential item bias	Appropriate differential functioning items analyses across equivalent items have been conducted to examine bias for the same items across the two language versions. For example, for each item, is there a bias against students tested in the target language? Item bias reviews have been conducted and subsequent changes have been made based on recommendations.	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 provided regarding item bias analysis or reviews No evidence for item analysis was given.	Does not meet – II Partially Meets - Meets or Exceeds-
Criterion	Specific Indicators	Ratings		Notes
Equity and fairness considerations on the translated test version				
1. Consideration of appropriate dialect	The translation provides documentation to show that the translated test version does not privilege any dialect of the target language over others (e.g. Iberic	DOES NOT MEET- evidence was not provided for this criteria or information does not demonstrate evidence. (0) PARTIALLY MEETS- partial evidence was	It is stated that the dialect is neutral and the texts are culturally generic.	Does not meet – Partially Meets - Meets or

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
	<p>Spanish - Spanish from Spain - is not privileged over Mexican or Puerto Rican dialects). Specifically, the translation procedures took into account the wide variety of dialects of the language speakers in the United States.</p>	<p>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>		Exceeds- I
<p>2. Appropriate cultural adaptation</p>	<p>Documentation is provided to show that items have been adapted to address cultural differences inherent to language. Cultural adaptations go beyond the superficial features of the contextual information provided by the items. For example, the items do not simply mention “Juan,” instead of “John,” as characters. Instead, consider how students’ experience may influence their interpretation of the items. Provide appropriate context for items to increase students’ access to the intended interpretation of the items.</p>	<p>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>	<p>On page 285 of the RFI it states that higher levels of reading are not culturally generic at higher levels of reading and they provide stories about students in Spanish speaking cultures. In addition pgs. 23-24 of the online technical manual describe in depth the translation creation and address variations and adaptations in terms of regional language terms and cultural differences.</p> <p>They do not address this topic. Because of the stance of being culturally generic, they do not address cultural adaptation.</p>	<p>Does not meet – Partially Meets - I Meets or Exceeds- I</p>
<p>3. Address stereotypes</p>	<p>The cultural adaptation of the test is not based on stereotypes about cultures.</p>	<p>DOES NOT MEET- evidence was not provided for this criteria or information does not demonstrate evidence. (0) PARTIALLY MEETS- partial evidence was</p>	<p>As a fully proficient and bicultural bilingual, I (J.B) examined and read the testing materials and found no evidence of stereotypes or cultural bias. In addition pgs. 23-24 of the online technical manual describe in depth the translation creation and address</p>	<p>Does not meet – I Partially Meets - Meets or</p>

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
		<p>provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p>MEETS OR EXCEEDS – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>variations and adaptations in terms of regional language terms and cultural differences. This reflect a lack of stereotypes in the materials.</p> <p>This topic is minimally addressed. Because of the stance of being culturally generic, they do not address cultural adaptation.</p>	Exceeds-

Strengths:

- 1) Widely tested in bilingual classrooms with the Literacy Squared Project and others
- 2) The amount of data included supporting the reliability and validity of the test
- 3) Provides a Spanish version that is consistent with the English version
- 4) Provided adequate data analysis

Weaknesses:

- 1) There were not item bias reviews
- 2) Up to date alignment to the CCSS and CED was not given
- 3) Use the DRA2 data instead of direct use of EDL2 data
- 4) Time intensive administration.

Recommended: X

Not Recommended: X