

Heineman Benchmark Assessment System – Spanish

**Spanish Assessment not Considered Because English Version Did not Meet 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><b>Evidence includes:</b> 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"> <li>• Split-half reliability</li> <li>• Coefficient alpha</li> <li>• Test-retest reliability</li> <li>• Classification consistency</li> </ul>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS</b> –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>	Lacking detailed evidence	<p><b>Does Not Meet–</b></p> <p><b>Partially Meets – I</b></p> <p>Meets or Exceeds -</p>
	<p>Standard error of measurement or standard estimate of error is reported</p> <p><b>Evidence includes:</b></p>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information</p>	Lacking detailed results	<p>Does not meet –</p> <p>Partially Meets – I</p>

	<ul style="list-style-type: none"> <li>SEM estimates are reported for score ranges and cut-scores.</li> <li>SEM estimates are reported for score ranges and cut-scores for each assessment (grade-level, form, subtest).</li> </ul>	<p>does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS --</b> Information and data provided suggests acceptable or strong evidence. (2)</p>		<p>Meets or Exceeds</p>
	<p>Inter-rater reliability studies have been conducted. Study sample used to establish inter-rater reliability represents test administrators.</p> <p><b>Evidence includes:</b></p> <ul style="list-style-type: none"> <li>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.</li> <li>Inter-rater reliability coefficients exceed .7.</li> </ul>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS --</b>most information for the criterion is provided. Information and data provided suggests acceptable or</p>	<p>Evidence not provided</p>	<p>Does not meet – I</p> <p>Partially Meets –</p> <p>Meets or Exceeds -</p>

		strong evidence. (2)		
	<p>Studies have been conducted to establish reliability with all subcategories of students who will take the assessment.</p> <p><b>Evidence Includes:</b> 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><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	Evidence found online in Full Report of Field Study for Reliability and Validity	<p>Does Not Meet –</p> <p>Partially Meets –</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"> <li>Technical reviews indicate all forms for each grade level have demonstrated evidence of comparability and</li> </ul>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion</p>		<p>Does Not Meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - I</p>

	<p>content specifications.</p> <p><b>Evidence includes:</b></p> <ul style="list-style-type: none"> <li>• Sufficient forms are provided to allow for progress monitoring between interim assessments.</li> <li>• Split-half reliability.</li> <li>• Coefficient alpha reliability.</li> </ul>	<p>and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS</b> –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>Evidence of content and construct validity</p>	<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><b>Evidence includes:</b></p> <ul style="list-style-type: none"> <li>• A clear description is provided that demonstrates the purpose of the assessment is to screen students for reading concerns.</li> <li>• Content specifications for each grade-level, including a complete description of the test content, purpose(s), and</li> </ul>	<p><b>Rating</b></p> <p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS</b> –most information for the criterion is</p>		<p>Does Not Meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - I</p>

	<p>intended use(s), and assessment blueprint as appropriate, is provided.</p>	<p>provided. Information and data provided suggests acceptable or strong evidence. (2)</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><b>Evidence includes:</b></p> <ul style="list-style-type: none"> <li>• Field testing populations should be clear and should mirror the school/district demographics.</li> <li>• Statistics used to establish the reading levels are reported with both ELL and Non-ELL populations.</li> <li>• Findings from a content review by field experts, including teachers in tested grade levels.</li> </ul>	<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 information for the criterion is provided. Information and data provided suggests acceptable or</p>		<p>Does Not meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - I</p>

		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><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	Insufficient evidence for a 3-standards are addressed but not detailed	<p>Does Not Meet –</p> <p>Partially Meets – I</p> <p>Meets or Exceeds -</p>
	There are studies of construct validity, such as convergent and discriminant analysis, demonstrating correlations of .7 or above.	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided</p>		<p>Does Not Meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - I</p>

		<p>demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</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><b>Evidence includes:</b></p> <ul style="list-style-type: none"> <li>• A clear definition of the criterion or measure that were used to establish concurrent validity.</li> <li>• Studies with similar assessments that demonstrate the assessment measures reading ability, not other irrelevant criteria. Predictive validity correlations above .7.</li> </ul>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS</b> –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 - I</p>
<p>Determination of cut-scores based upon well-designed pilot</p>	<p>The assessment has established cut-scores for decision making about</p>	<p><b>DOES NOT MEET</b>-evidence was not</p>		<p>Does Not Meet –</p>

study	<p>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><b>Evidence indicates:</b></p> <ul style="list-style-type: none"> <li>• Includes a description of the process used to establish the cut points.</li> <li>• A full description of the norming sample.</li> <li>• 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.</li> </ul>	<p>provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and 2data provided suggests acceptable or strong evidence. (2)</p>		<p>Partially Meets –</p> <p>Meets or Exceeds - I</p>
	<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>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR</b></p>		<p>Does Not Meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - I</p>



		<p><b>EXCEEDS</b> –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>		
	<p>Acceptable, recognized procedures are followed for setting cut-scores.</p>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  <b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  <b>MEETS OR EXCEEDS</b> –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 - I</p>
	<p>SEM estimates are reported for cut-scores with guidance for score interpretation.</p>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate</p>	<p>Not enough evidence provided to understand their SEM for the cut-scores, but cut-scores are provided with</p>	<p>Does Not Meet –</p> <p>Partially Meets – I</p> <p>Meets or</p>

		<p>evidence. (0)  <b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence.(1)  <b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>guidance for score interpretation</p>	<p>Exceeds -</p>
<p>Universal Design</p>	<p>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.</p> <p><b>Evidence includes:</b></p> <ul style="list-style-type: none"> <li>• Addressed issues of equity of utility for all populations.</li> <li>• Results of bias reviews and plans that have addressed any concerns.</li> <li>• At least two to three types of classification, reliability, and validity study data have been disaggregated by</li> </ul>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence.(0)  <b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  <b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and data provided suggests</p>		<p>Does Not Meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - I</p>

	<p>subgroups and meet the criteria.</p> <ul style="list-style-type: none"> <li>• Culturally diverse students were included throughout the entire process of test development. For example in the samples of pilot students, in cognitive interviews, etc.</li> <li>• The content of the reading materials does not favor mainstream culture.</li> </ul>	<p>acceptable or strong evidence. (2)</p>		
Third party evaluation conducted	<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><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0) <b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1) <b>MEETS OR EXCEEDS</b> –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 - I</p>
Standardization of	Administration protocol is	<b>DOES NOT</b>		Does Not

<p>materials and procedures for administration</p>	<p>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>	<p><b>MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>		<p>Meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - I</p>
<p>Efficiency of administration</p>	<p>The amount of time needed to administer the assessment is reasonable and balanced to the information provided.</p>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence.</p>	<p>May be more efficient for some students than for others.</p> <p>Systems need to be in place to allow for teacher administration</p>	<p>Does Not Meet-</p> <p>Partially Meets – I</p> <p>Meets or Exceeds -</p>

		(1) <b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)		
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.		There is no computer scoring, but it is clear what you need to do	Does not Meet – Partially Meets – I Meets or Exceeds-
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. <b>Evidence includes:</b> <ul style="list-style-type: none"> <li>Any accommodations do not compromise the interpretation or purpose of the test.</li> <li>Specific administration guidelines are provided for implementing any accommodations.</li> <li>How to address accommodations is specifically addressed in the training materials or program.</li> <li>Suggested accommodations are research or evidence-based.</li> </ul>	<b>DOES NOT MEET</b> -evidence was not provided for this criteria or information does not demonstrate evidence. (0) <b>PARTIALLY MEETS</b> -partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1) <b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and	No evidence presented	Does Not Meet – I Partially Meets- Meets or Exceeds -

		data provided suggests acceptable or strong evidence. (2)		
Accommodations clearly stated and described for Second Language Learners	<p>The accommodations directly address the linguistic needs of the student.</p> <p><b>Evidence includes:</b></p> <ul style="list-style-type: none"> <li>Any accommodation does not compromise the interpretation or purpose of the test.</li> <li>Specific administration guidelines are provided for implementing any accommodations.</li> <li>How to address accommodations is specifically addressed in the training.</li> <li>Suggested accommodations are research or evidence-based.</li> </ul>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	N/A – the test would not be taken by non-Spanish speaking ELLs	<p>Does Not Meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - I</p>
Scores are easily interpreted to determine a <i>“significant reading deficiency”</i>	<p>Scores clearly specify whether a student is categorized as having a <i>“significant reading deficiency”</i>.</p> <p><b>Evidence includes:</b></p> <ul style="list-style-type: none"> <li>Score ranges or a scale is provided.</li> <li>Guides for interpretation of scores</li> </ul>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence.(0)</p> <p><b>PARTIALLY MEETS</b>-partial evidence was</p>		<p>Does Not Meet –</p> <p>Partially Meets –</p> <p>Meets or Exceeds - I</p>

	are provided.	provided related to the criterion and/ or data provided demonstrates weak evidence. (1) <b>MEETS OR EXCEEDS</b> –most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)		
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.	<b>DOES NOT MEET</b> -evidence was not provided for this criteria or information does not demonstrate evidence.(0) <b>PARTIALLY MEETS</b> -partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence (1) <b>MEETS OR EXCEEDS</b> –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

<p>Reports provide guidance for interpretation useful to educators, administrators, and parents</p>	<p>Information is displayed in a format and language that is understandable to educators, administrators and parents;</p> <ul style="list-style-type: none"> <li>• Data reports are easily read and interpreted.</li> <li>• Clear description of how to interpret results.</li> <li>• Reports provide trajectory for student progress.</li> <li>• District, school, classroom, and student reports provided.</li> <li>• Reports available in real-time.</li> <li>• Reports can be exported to data-base formats.</li> <li>• Reports available in languages other than English.</li> <li>• Customer service is available provided for users.</li> </ul>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  <b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  <b>MEETS OR EXCEEDS</b> –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 – I</p> <p>Meets or Exceeds -</p>
---	--	--	--	---

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
<p><b>Translation and adaptation procedure</b></p>				
<p><b>1. Translation has been provided by highly qualified personnel.</b></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"> <li>• translators who are native speakers in the</li> </ul>	<p><b>DOES NOT MEET</b>-evidence was not provided for this criteria or information does not demonstrate evidence. (0)  <b>PARTIALLY MEETS</b>-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p>	<p>The manual states that passages were translated by native speakers. However, this team and their qualifications are never explicitly stated. There is no way to truly know the qualifications of</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
	target language <ul style="list-style-type: none"> <li>• specialists in reading in the target language</li> <li>• bilingual educators</li> </ul> (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.	<b>MEETS OR EXCEEDS</b> – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)	the translation team.	
<b>2. Pilot test sampling appropriately considers language diversity</b>	The translated test was piloted with a representative sample of speakers of the target language in the United States.	<b>DOES NOT MEET-</b> evidence was not provided for this criteria or information does not demonstrate evidence. (0) <b>PARTIALLY MEETS-</b> partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1) <b>MEETS OR EXCEEDS</b> – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)	The proposal states that the assessment was field tested in various states with Spanish speaking populations. However, because there is no sample size provided it is not possible to confidently generalize the findings to all settings. It is also unclear what is meant by “field testing.” Simply using the passages in classrooms is an inadequate pilot test for a Universal Screening Measure.	Does not meet – I  Partially Meets -  Meets or Exceeds- I
<b>3. Consistency of appearance between the English language and the target language version of the test</b>	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	<b>DOES NOT MEET-</b> evidence was not provided for this criteria or information does not demonstrate evidence. (0) <b>PARTIALLY MEETS-</b> partial evidence was provided related to the criterion and/ or data		Does not meet –  Partially Meets -  Meets or Exceeds- II

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
	consistent with the English language test version.	provided demonstrates weak evidence. (1) <b>MEETS OR EXCEEDS</b> – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)		
Criterion	Specific Indicators	Ratings		Notes
<b>Psychometric and measurement considerations:</b>				
<b>1. Construct validity for translated test versions</b>	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 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.	<b>DOES NOT MEET-</b> evidence was not provided for this criteria or information does not demonstrate evidence.(0)  <b>PARTIALLY MEETS-</b> partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  <b>MEETS OR EXCEEDS</b> – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)	It has been clearly documented in the RFI that the SEL is not a direct translation of the English version. However, no psychometric information has been provided regarding construct validity of this assessment therefore, there is no way to confidently identify students with a significant reading deficiency.	Does not meet – I  Partially Meets -  Meets or Exceeds- I

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
<p><b>3. Demonstrated comparability</b></p>	<p>Evidence is provided on the psychometric comparability of measures in English and measures in the target language.</p>	<p><b>DOES NOT MEET-</b> evidence was not provided for this criteria or information does not demonstrate evidence. (0)  <b>PARTIALLY MEETS-</b> partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  <b>MEETS OR EXCEEDS –</b> most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>No psychometrics were provided for either the English or Spanish version.</p>	<p>Does not meet – I  Partially Meets -  Meets or Exceeds- I</p>
<p><b>4. Documentation on the interpretation of scores and the scaling of scores</b></p>	<p>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. 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?</p>	<p><b>DOES NOT MEET-</b> evidence was not provided for this criteria or information does not demonstrate evidence. (0)  <b>PARTIALLY MEETS-</b> partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)  <b>MEETS OR EXCEEDS –</b> most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	<p>The RFI states that the leveling system used for SEL is equivalent to the English version. However, there is no research to demonstrate the passages are accurately leveled and there is no psychometric information provided to demonstrate reliability of student scores based on this assessment</p>	<p>Does not meet – I  Partially Meets -  Meets or Exceeds- I</p>
<p><b>5. Evidence</b></p>	<p>Appropriate differential</p>	<p><b>DOES NOT MEET-</b></p>	<p>Not clearly addressed in</p>	<p>Does not</p>

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
<b>provided regarding investigation into potential item bias</b>	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.	evidence was not provided for this criteria or information does not demonstrate evidence. (0) <b>PARTIALLY MEETS-</b> partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1) <b>MEETS OR EXCEEDS –</b> most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)	regards to item bias analysis  The RFI states that because this is not a translation item bias is not an issue. However, it is possible that items can still be biased even if initially written in the target language. All items, regardless of language should undergo an item review and the process and outcomes of this review should be documented. None of this evidence has been provided.	meet – I  Partially Meets - I  Meets or Exceeds-
Criterion	Specific Indicators	Ratings		Notes
<b>Equity and fairness considerations on the translated test version</b>				
<b>1. Consideration of appropriate dialect</b>	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 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	<b>DOES NOT MEET-</b> evidence was not provided for this criteria or information does not demonstrate evidence. (0) <b>PARTIALLY MEETS-</b> partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1) <b>MEETS OR EXCEEDS –</b> most information for the criterion is provided. Information and data	This test is not translated, therefore this is not applicable	Does not meet –  Partially Meets -  Meets or Exceeds- II

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
	the United States.	provided suggests acceptable or strong evidence. (2)		
<b>2. Appropriate cultural adaptation</b>	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><b>DOES NOT MEET-</b> evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS-</b> partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS –</b> most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)</p>	Because this is not a translation, it is obvious that some effort has been made to ensure that cultural adaptations have been made appropriately. However, there are no documented pilot studies that demonstrate that items are culturally appropriate for students across the United States.	Does not meet –  Partially Meets - I  Meets or Exceeds- I
<b>3. Address stereotypes</b>	The cultural adaptation of the test is not based on stereotypes about cultures.	<p><b>DOES NOT MEET-</b> evidence was not provided for this criteria or information does not demonstrate evidence. (0)</p> <p><b>PARTIALLY MEETS-</b> partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1)</p> <p><b>MEETS OR EXCEEDS –</b> most information for the criterion is provided. Information and data</p>	The authors of the test and passages are native speakers, therefore the risk of stereotypes is lower than if it was simply a direct translation. However, even with native speakers there is a risk of stereotypes. There is no documented review to	Does not meet –  Partially Meets - I  Meets or Exceeds- I

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
		provided suggests acceptable or strong evidence. (2)	demonstrate that items have been examined for cultural stereotypes.	

Strengths:

- 1) Not translated – authentic Spanish
- 2) Extensive field of Spanish version with specific evidence provided
- 3) This test is not a translation, but rather fully written by native Spanish speakers with Spanish speaking populations in mind.

Weaknesses:

- 1) Did not address inter-rater reliability
- 2) Teacher needs time to personally administer
- 3) There is absolutely no data provided to support any of the statements made in the RFI
- 4) Universal Screeners must meet a strict psychometric standard and no psychometric studies have been conducted on this assessment.

**Recommended: X**

**Not Recommended: X**