

SPANISH ASSESSMENT: I-STATION

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
Translation and adaptation 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) 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>minimal translation has been used which preserved authenticity of language 2</p> <p>Authors of the assessment were part of a bilingual assessment company, TRI-LIN. Resumes for several authors were provided to document their expertise. This assessment is not a translation, but rather a fully Spanish assessment.</p>	<p>Does not meet – 0</p> <p>Partially Meets - 0</p> <p>Meets or Exceeds- III</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) 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</p>	<p>Though the pilot testing had a relatively large sample size, the majority of the schools and districts were from a single state (Texas- 95.6%) which limits the generalizability of the findings.</p>	<p>Does not meet – I</p> <p>Partially Meets - I</p> <p>Meets or Exceeds- I</p>

SPANISH ASSESSMENT: I-STATION

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
		evidence. (2)		
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.	<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 – 0</p> <p>Partially Meets - 0</p> <p>Meets or Exceeds- III</p>
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 provided that the reading constructs	<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</p>	A considerable amount of validity research has been conducted. However these concurrent validity studies have been conducted with assessments used in Texas. Because of this it would be difficult to ensure that this concurrent validity is the same for states that use different summative assessments. Additionally	<p>Does not meet – I</p> <p>Partially Meets - II</p> <p>Meets or Exceeds- 0</p>

SPANISH ASSESSMENT: I-STATION

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
	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.	acceptable or strong evidence. (2)	many of the correlation values between ISIP Espanol and Tejas Lee are quite low (.1-.3).	
3. Demonstrated comparability	Evidence is provided on the psychometric comparability of measures in English and measures in the target language.	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)	The validity of this assessment was only examined in relation to Texas assessments. Additionally, there was no classification analysis conducted to provide evidence that classification of students within Tiers using this assessment was accurate and comparable to other assessments.	Does not meet – 1 Partially Meets - 0 Meets or Exceeds- 11
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. For example, do teachers need to scale the score of the translated test	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	Scores are calculated by the computer program with interpretation help provided within the reports. Again, no classification analysis was conducted to examine how these scores and placement of students with this assessment compare to placement in	Does not meet – 0 Partially Meets - 11 Meets or Exceeds- 1

SPANISH ASSESSMENT: I-STATION

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
	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?	acceptable or strong evidence. (2)	other Universal Screeners.	
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)	Review conducted by a panel of Spanish language experts. Review covered colloquial terms, regionalisms, bias and ambiguity.	Does not meet – 0 Partially Meets - 0 Meets or Exceeds- III
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 Spanish - Spanish from Spain - is not privileged over Mexican or Puerto Rican dialects). Specifically, the	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)	Item review by panel of experts specifically covered this.	Does not meet – 0 Partially Meets - 0 Meets or Exceeds- III

SPANISH ASSESSMENT: I-STATION

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
	translation procedures took into account the wide variety of dialects of the language speakers in the United States.	MEETS OR EXCEEDS – most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)		
2. Appropriate cultural adaptation	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.	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 – 0 Partially Meets - 0 Meets or Exceeds- III
3. Address stereotypes	The cultural adaptation of the test is not based on stereotypes about cultures.	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 –		Does not meet – 0 Partially Meets - 0 Meets or Exceeds- III

SPANISH ASSESSMENT: I-STATION

Criterion	Specific Indicators	Ratings	Feedback from Reviewers	Tally of Rating
		most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)		

Strengths:

- (1) Significant research has been done to ensure the test items perform well
- (2) Thorough outside review conducted to ensure the test is free of bias and stereotyping

Weaknesses:

- (1) There are great concerns that 96% of the sample for the Spanish version of the test are taken from Texas.
- (2) Additionally, concurrent validity has only been measured using Texas specific summative assessments. It is unclear whether these results would replicate to assessments used in other states. This is a critical aspect of an interim assessment and must be weighed heavily in the decision to recommend this assessment in the state of Colorado. Because of the state specific (Texas) information provided, the reviewers do not feel comfortable recommending this assessment for use in Colorado even though other areas of this assessment (equity, usability) received high scores.

Not Recommended X

Recommended X X