Istation ISIP ER

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)	No evidence of split-half reliability or coefficient alpha present. Page 41 Test is adaptive – students do not respond to a common set of items(excluding Coefficient alpha as a measure) Test-retest reliability – 0.927 to 0.970 (n=416) See table 7, p. 42 Test-retest completed across seven sections from Oct. to Feb. Internal Consistency high and stable over time(up to 5 months apart)	Does not meet- Partially meets-1 Meets or exceeds- 3
	Standard error of measurement or standard estimate of error is reported Evidence includes: • SEM estimates are reported for score	DOES NOT MEET-evidence was not provided for this criteria or information does not	Reviewers are not able to locate this information and the technical manual was downloaded from	

SEM report range for eachers	es and cut-scores. estimates are rted for score es and cut-scores ach assessment e-level, form, est).	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)	Not reported No standard of error of measure included due to adaptability of assessment	Does not meet-4 Partially meets- Meets or exceeds-
have been sample us inter-rated represent administrated representation r	ators.	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	Computer administered assessment Computer administered assessment Page42 States there was inter-rater reliability but does not provide data to correlate at grade level or specific coefficient. States IRR is within two points of all administrators. Due to adaptive, non-identical assessments, interrater reliability is difficult to determine.	Does not meet-1 Partially meets- Meets or exceeds-3

		evidence. (2)		
	Studies have been conducted to establish reliability with all subcategories of students who will take the assessment. Evidence Includes: Studies that demonstrate reliability has been established from scoring samples of students that include: Non-ELLs with and without reading deficiencies and ELLs with and without reading deficiencies.	DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0) PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1) MEETS OR EXCEEDS —most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)	Reliability was reported but not broken up by demographics groups. Page 38 All subcategories of students are represented in the reliability and validity report Noted that studies were based off of one school district in North Texas.	Does not meet-1 Partially meets- Meets or exceeds-3
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.	MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0)	Evidence of parallel forms, but not given about how often the test can be given. Computer based has bank of	Does not meet- Partially meets-2
	Technical reviews indicate all forms for each grade level have demonstrated evidence of comparability and content specifications.	PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided	questions instead of alternate forms. This assessment is based on IRT not classical test theory, therefore it does not need	Meets or exceeds-2

	• Sufficient forms are provided to allow for progress monitoring between interim assessments. • Split-half reliability. • Coefficient alpha reliability.	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)	Since no student receives the same assessment, multiple forms are inherent in the program. Page 2 –states ISIP ER assess students automatically each month or on demands, as desired to measure progress throughout the school year Coefficient alpha is noted as difficult to measure due to adaptability of test from student to student. Data shows comparability at each grade level to assessments	
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.	Rating DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0) PARTIALLY MEETS-partial evidence was	Statistics given about content validity of test. Page 1-web-based assessment of critical early reading skills that are predictive of later reading success Page 13 -Shows subtest	Does not meet- Partially meets- Meets or exceeds-4

• 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.	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)	administration for teach grade level Page 14- Description of each subtest contains purpose and intended use of each.	
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 correlations	Partial evidence showing they are reliable, does not show how they were made. Not evident in the technical manual or technical report The RFI and technical manual report that readability has been established using a "commonly accepted readability formula" but it does not say what this formula is. It also does not address whether these passages were field tested and which populations they were tested with.	Does not meet-2 Partially meets-2 Does not meet-

	demonstrate ranges of .7 or higher. (2)	There is also no mention of a content review by a field of experts related specifically to the passages. Page18- To publisher, each of these passages was carefully written to conform to specific word level features follow linear story grammar structure and have readability according to a commonly accepted	
		readability according to a commonly accepted readability formula for the end grade level in each grade.	
		Demographics mirror school/district	
		Levels of assessment passages are not reported	
		Page 35- Field expert and	
If appropriate, findings from alignment studies to demonstrate alignment with	DOES NOT MEET-evidence was not	Alignment to both CAS are provided on the Istation	Does not meet-2
Colorado Academic Standards for Language Arts and resolution for any	provided for this criteria or information	web page No evidence of	Partially meets-
and resolution for any	does not	alignment to	Meets or

resulting concerns	domonstrata	Colorada	ovenede 3
resulting concerns.	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)	Colorado Academic Standards.	exceeds-2
There are studies of construct validity, such as convergent and discriminant analysis, demonstrating correlations of .7 or above.	DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0) PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1) MEETS OR EXCEEDS —most information for the criterion is provided. Information and data provided	Items previously tested for constructed validity were used to create. Author failed to demonstrate that this new test maintains construct validity. Not all correlations are at or about .70 Page 42 Studies listed under validity section. However under multiple external measures, a correlation of .7 or higher is not always evident.	Does not meet- Partially meets- 3 Meets or exceeds-1

Evidence of criterion/predictive validity accurately identifying students with "significant reading deficiency" Determination of cutscores based upon	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.	suggests acceptable or strong evidence. (2) DOES NOT MEET-evidence was not provided for this criteria or information does not demonstrate evidence. (0) PARTIALLY MEETS-partial evidence was provided related to the criterion and/ or data provided demonstrates weak evidence. (1) MEETS OR EXCEEDS —most information for the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)	Not all correlations are at or about .70 Clear definition of criterion used to establish concurrent validity and the ranges of percentiles for the SRD are provided. However predictive validity is only included as a range of 0.59-0.82 on "a variety of reading test" with no explicit information provided. Criterion definitions use to establish validity for each measure compared page 40-41. Validity correlations of .7 or above are shown for significant number of external measures	Does not meet- Partially meets-2 Meets or exceed-2
well-designed pilot study	decision making about students' "significant reading deficiency" using adequate demographics	was not provided for this criteria or information does not	Three tiers of normative grouping based on percentile rank.	Partially meets-

/:		- 11 1	
representing (i.e., 10% ELL	demonstrate	Full description of	
and 25% F/R lunch),	evidence. (0)	norming sample	Meets or
appropriate criterion	PARTIALLY	which includes	exceeds-4
assessment, adequate	MEETS-partial	students from 24	
sample size, and appropriate	evidence was	states.	
statistics.	provided		
	related to the	Know	
Evidence indicates:	criterion and/or	demographics	
 Includes a description of 	data provided	information is	
the process used to	demonstrates	detailed.	
establish the cut points.	weak evidence.		
A full description of the	(1)	Large	
norming sample.		representative	
The norming sample is a	MEETS OR	sample.	
large representative	EXCEEDS –most		
• .	information for		
national sample of	the criterion is		
students at the same	provided.		
grade level and is	Information and		
representative of the	2data provided		
testing population	suggests		
according to gender, ELL	acceptable or		
status, special needs	strong		
status and F/R lunch	evidence. (2)		
status.	evidence. (2)		
Studies of classification	DOES NOT	Correlations, .7 .9	Does not
accuracy analysis provide	MEET-evidence	·	meet-2
evidence that the measure	was not	Not clearly	
appropriately identifies	provided for	evident.	
students as indicated in the	this criteria or		Partially
description of purpose of the	information	No studies of	meets-1
	does not	classification	
assessment, demonstrating	demonstrate	accuracy are	
values that exceed .8 or	evidence. (0) PARTIALLY	provided.	Meets or
higher.	MEETS-partial	F	exceeds-1
	evidence was	Contains	
	provided	information on	
	related to the	measures including	
	criterion and/or	criterion-	
	data provided	referenced points,	
	demonstrates	norm-referenced	
	weak evidence.	points and lexile	
	(1)	scores to create an	
	NATITE OF	accurate analysis.	
	MEETS OR	accurate analysis.	
	EXCEEDS –most	Values are not	
	information for	Values are not	
	the criterion is	shared to	
		demonstrate	ı

	provided. Information and data provided suggests acceptable or strong evidence. (2)	values that exceed .8	
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)	Uses percentiles. Acceptable procedures for setting tiered cut scores based on percentile rank and use of standards deviations	Does not meet- Partially meets- Meets or exceed-4
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)	Could not find. Standard error or measurement estimates are not evident to reviewers. No evidence of	Does not meet-4 Partially meets-

		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)	SEM estimates.	exceeds-
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	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	Child frequently in game format, no evidence of bias reviews. Norming samples included a large percentage of culturally diverse students. However, statistics were not disaggregated by group to indicate that it is equally reliable and valid for all subgroups. Does not identify normed score based on disaggregated groups. Proof of bias control is not clear.	Partially meets-3 Meets or exceeds-1

	Г		Γ = .	
	criteria.	strong	Disaggregated	
	Culturally diverse	evidence. (2)	subgroups include	
	students were included		race, gender, and	
	throughout the entire		placement.	
	process of test			
	development. For		Sample includes	
	example in the samples		students from 24	
	of pilot students, in		states with	
	cognitive interviews,		appropriate	
	etc.		representation;	
	etc.		this qualifies as an	
			adequately diverse	
	The content of the		population.	
	reading materials does			
	not favor mainstream			
	culture.			
Third party evaluation	Evidence reported to	DOES NOT	Reviewers did not	Does not
conducted	demonstrate that an	MEET-evidence	find evidence of	meet-2
	independent, qualified third	was not	this proposal.	
	party has provided a	provided for		
	thorough and unbiased	this criteria or	A panel of experts	Partially
	evaluation of the quality of	information	was used to review	meets-
	the assessment.	does not demonstrate	items for bias,	
	the assessment.	evidence. (0)	however the	
		eviderice. (0)	overall assessment	Meets or
		PARTIALLY	and study was not	exceeds-2
		MEETS-partial	reviewed by an	execus 2
		evidence was	outside party.	
		provided	outside party.	
		related to the	ISIP has	
		criterion and/ or	documented	
		data provided	independent	
		demonstrates		
		weak evidence.	reviews and	
		(1)	studies from	
		MEETS OR	school districts	
		EXCEEDS –most	comparing ISIP to	
		information for	state assessment	
		the criterion is	and other widely	
		provided.	accepted	
		Information and	assessments.	
		data provided		
		suggests		
		acceptable or		
		strong		
		evidence. (2)		
		271401100. (2)		

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)	Computer administered. Script provided in online teacher manual. Guidelines are provided for materials to be created.	Does not meet- Partially meets- Meets or exceeds-4
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	Time to administer is around 30 minutes. Notice time limit for answering questions. Progress monitoring time shortened if only monitoring specific subtest	Does not meet- Partially meets- Meets or exceeds-4

	T			
		weak evidence.		
		(1)		
		MEETS OR		
		EXCEEDS –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		Immediate scoring	Does not
	to score the assessment is		as student takes	meet-
	reasonable and balanced to	DOES NOT	assessment on	Partially
	the information provided;	MEET -evidence	computer.	meets-
	computer-assisted scoring is	was not		
		provided for	Efficient, storable	
	available; procedures for	this criteria or	computer-assisted.	Meets or
	calculating scores are clear;	information	computer-assisted.	exceed-4
	scores can be stored and	does not		exceeu-4
	reported electronically.	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 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.	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)	Accommodation are addressed, however they are minimal at best and only provided for two populations of students. Supplemental materials contains adequate accommodations and modification for students with disabilities.	Does not meet- Partially meets- Meets or exceed-4
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	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	No accommodations listed for students other than a Spanish version of the test. Students speaking languages other than Spanish are not addressed. Specific guidelines for administering assessment with accommodations not located in technical manual or teacher guide.	Does not meet-1 Partially Meets-1 Meets or exceeds-2

	accommodations are research or evidence-	EXCEEDS –most information for	ELL accommodations	
	based.	the criterion is provided. Information and data provided suggests acceptable or strong evidence. (2)	given for lesson but not assessments.	
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.	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		Does not meet- Partially Meets- Meets or exceeds-2
		suggests acceptable or strong evidence. (2)		
Cost effective: Materials, administration costs including personnel,	Materials are provided or easily accessible; time away from instruction is minimal; no additional personnel	DOES NOT MEET-evidence was not provided for	\$5.50 per student	Does not meet-
scoring, and training	required; all costs inclusive including any additional data	this criteria or information does not		Partially meets-

	platform or storage costs; minimal data entry is required.	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)		Meets or exceeds-3
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 realtime. Reports can be exported to data-base formats. Reports available in languages other than	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.	Real time report, identify level of risk, shows growth, group in tiers are useful, results for classroom and by student. No districts reports provided	Partially meets- Meets or exceeds-3
	•	the criterion is		

available provided for users.	suggests acceptable or	
000.01	strong	
	evidence. (2)	

Strengths:

- 1) <u>Cost effective, can test whole class at once, screen for levels of risk. English version SRD is clear, reports are clear.</u>
- 2) Computer based assessment and scoring streamlines administration and scoring procedures
- 3) An alignment with CAS and CCSS is provided on Istation web page.
- 4) Computer administration allows for standard administration and easy scoring.
- 5) Provides instructional materials to assist teachers in meeting student needs.

Weaknesses:

- 1) Not effective for monitoring.
- 2) Could not find SEM.
- 3) Statistical measurements such as SEM and construct validity are not readily evident.
- 4) Third party evaluation is not evident.
- 5) Not a stand-alone interim assessment. It is built to be used as an intervention curriculum with an assessment built in.
- 6) Requires school to have computers available for all students for testing purposes.

Recommend: XXX Not recommend: XX