Research Brief

Relationship between the Independent Reading Level Assessment (IRLA) and State ELA Tests: Concurrent and Predictive Validity Evidence

American Reading Company's Independent Reading Level Assessment® (IRLA) is a standards-based formative assessment framework that is used on a regular basis throughout the year to measure the extent to which students independently demonstrate reading proficiency. Researchers have conducted a series of validity studies across the United States to examine the relationship between the IRLA and summative state ELA assessments. Findings from 14 studies provide strong evidence of the IRLA's concurrent and predictive validity. Strong positive correlations between students' scores on the state ELA test and scores on the IRLA during the state testing window indicate that the IRLA measures the same construct as the state test (concurrent validity). Strong positive correlations between students' IRLA scores from the beginning of the school year and their scores on the state test (administered in the spring) indicate that student performance on the IRLA is a good indicator of how students are likely to perform on the state test (predictive validity).

This research brief includes an overview of the IRLA, an explanation of why it is important to examine the validity of educational assessments, a description of the study methods, and findings from the validity studies.

Independent Reading Level Assessment (IRLA)

The IRLA outlines a research-based, transparent progression of skills mapped to national and state standards. Designed to work for every student at every reading level, the IRLA delivers specific and actionable data that tell the teacher where a student is, why, and the sequence of skills and behaviors needed to learn next to accelerate reading growth.

IRLA scores show students' relative placement along a continuum of grade-level proficiency. A risk status is used to identify the intensity of student need. Students who have demonstrated reading proficiency at or above their grade level are considered "proficient" and are not likely to be at risk for academic difficulties. Students who need to make more than a year of growth in one year's time are assigned an "at risk" designation that alerts teachers that the student may need additional supports to make sufficient accelerated progress. Students reading significantly below grade level are assigned "emergency" status. These students need multiple years of growth per year to gain grade-level proficiency and require the most intensive supports to make accelerated progress.

Validity

Validity is the most fundamental consideration in evaluating an assessment. Validity is the degree to which evidence supports interpretations of test scores for a given purpose.[1] The process of validation involves accumulating relevant evidence over time to provide a sound basis for the proposed score interpretations and is the responsibility of the test developer.[2]

The IRLA is used to monitor student reading progress and identify students who have not yet achieved grade-level reading proficiency and are at risk for academic difficulties. Thus, one particularly relevant form of validity evidence is the extent to which performance on the IRLA correlates with performance on other reading assessments, which are called criterion measures.

Correlation coefficients can range from –1.0 to +1.0, with values close to ±1.0 indicating a strong relationship. Positive correlations indicate that when students score high on one assessment, they also tend to score high on the other, and similarly, when students score low on one assessment, they also tend to score low on the other. In education research, correlation coefficients of .70 or greater are considered strong; coefficients ranging from .50 to .69 are considered moderate, and coefficients less than .50 are

[1] American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

[2] Ibid.

considered weak.[3] When an assessment is strongly correlated with several different measures of the same construct, there is greater confidence that results can be generalized to other measures of student proficiency.

Methods

The 14 studies described in this brief utilized annual summative state ELA assessment[4] data provided by partner school districts to examine the statistical relationship[5] between students' scores on the IRLA and the state test. In all studies, both the IRLA and state assessments were administered independently by school district personnel using standard protocols.

This brief presents two types of criterion-related validity evidence for the IRLA: concurrent and predictive. Concurrent validity evidence is used to show that the IRLA measures what it is designed to measure (reading proficiency). For this analysis, students' scores on the state ELA test were correlated with their scores on the IRLA during the state testing window. A strong positive correlation between concurrent scores on the tests indicates that the IRLA is measuring the same construct as the state test.

Predictive validity evidence is used to show that student performance on the IRLA is a good indicator of how students are likely to perform on the state test. For this analysis, students' IRLA scores from the beginning of the school year were correlated with their scores on the state test, administered in the spring. A strong positive correlation indicates that the students who score high on the IRLA early in the year are likely to score high on the state test at the end of the year, and students who score low on the IRLA early in the year are likely to score low on the state test at the end of the year.

Findings

Connecticut: Smarter Balanced Assessment (SBA)

Connecticut Study: This study was conducted in a midsize Connecticut school district that began using ARC Core in the 2018–2019 school year. District enrollment includes 31% English Learners, 20% Students with Disabilities, and 73% Eligible for Free/Reduced-Price Meals. In 2021 and 2022, the concurrent and predictive correlations between students' IRLA and SBA ELA scores were strong and statistically significant (see Table 1).

[3] Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioral sciences* (5th ed.). Houghton Mifflin.
[4] Each state administers an annual summative English Language Arts (ELA) assessment to students in Grades 3–8 and once in high school under the provisions of the Every Student Succeeds Act (ESSA), 20 U.S.C. § 6301. (2015).
[5] The Pearson Product Moment Correlation was used.

Table 1. IRLA-SBA ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2021 Gr. 3–5	621	.747*	482	.760*
2022 Gr. 3–8	1501	.760*	1009	.742*

*$p<.001$

Delaware: Smarter Balanced Assessment (SBA)

Delaware Study 1: This study was conducted over four years in a midsize district in Delaware that began using ARC Core in the 2017–2018. The school district serves a population of students that is 67% White; 17% of students are classified as Low Income and 9% are English Language Learners. The number of students in the study grew each year as the implementation expanded from Grades K–5 to Grades K–8. Concurrent and predictive correlations between students' IRLA and SBA ELA scores were strong and statistically significant in 2019, 2021, 2022 and 2023 (see Table 2). The study did not examine scores from the 2019–2020 school year because the SBA was not administered that year due to the pandemic.

Table 2. IRLA-SBA ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2019 Gr. 3–5	1099	.737*	1195	.716*
2021 Gr. 3–8	1446	.723*	1523	.714*
2022 Gr. 3–8	2273	.702*	2510	.702*
2023 Gr. 3–8	2543	.715*	2412	.697*

*$p<.001$

Delaware Study 2: This study was conducted in a midsize school district in Delaware that has been using ARC Core since the 2019–2020 school year. The district serves a population of students that is 49% Nonwhite; 26% of students are classified as Low Income and 5% are English Language Learners. In 2022 and 2023, the concurrent and predictive correlations between students' IRLA and SBA ELA scores were strong and statistically significant (see Table 3).

Table 3. IRLA-SBA ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2022 Gr. 3–5	1854	.698*	1754	.700*
2023 Gr. 3–5	1875	.707*	1796	.716*

*p<.001

Delaware Study 3: This study was conducted in a midsize school district in Delaware that has been using ARC Core since the 2020–2021 school year. The district serves a population of students that is 49% Nonwhite; 22% of students are classified as Low Income and 4% are English Language Learners. In 2021, 2022, and 2023, the concurrent and predictive correlations between students' IRLA and SBA ELA scores approached or exceeded the threshold for strong (see Table 4).

Table 4. IRLA-SBA ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2021 Gr. 3–6	1445	.723*	1145	.714*
2022 Gr. 3–6	1867	.681*	1768	.665*
2023 Gr. 3–6	1845	.685*	1805	.684*

*p<.001

Illinois: Illinois Assessment of Readiness (IAR)

Illinois Study 1: This study was conducted in a midsize school district in Illinois that began using ARC Core in the 2021–2022 school year. The district serves a population of students that is 61% Nonwhite; 36% of students are classified as Low Income and 28% are English Language Learners. In 2022, the concurrent and predictive correlations between students' IRLA and IAR ELA scores approached or exceeded the threshold for strong (see Table 5).

Table 5. IRLA-IAR ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2022				
Gr. 3	274	.693*	269	.687*
Gr. 4	275	.727*	276	.678*
Gr. 5	326	.716*	319	.665*

*p<.001

Illinois Study 2: This study was conducted in a small school district in Illinois that began using ARC Core in the 2022–2023 school year. The district serves a population of students that is 96% Nonwhite; 74% of students are classified as Low Income and 36% are English Language Learners. In 2023, both the concurrent and predictive correlations between students' IRLA and IAR ELA scores were strong and statistically significant (see Table 6).

Table 6. IRLA-IAR ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2023				
Gr. 3	196	.704*	191	.729*
Gr. 4	221	.734*	219	.698*

*p<.001

Oregon: Oregon Assessment of Knowledge and Skills (OAKS)

Oregon Study: A 2016 study conducted by researchers at the University of Portland and Northwest Evaluation Association (NWEA) and published in *The Journal of At-Risk Issues*[6] examined the relationship between scores on the IRLA and the OAKS in one Oregon school district. The district serves almost 11,000 ethnically and linguistically diverse students with nearly 75% qualifying for Free/Reduced-Price lunch. The study found strong statistically significant concurrent correlations (see Table 7).

[6] Ralston, N.C., Waggoner, J. M., Tarawasa, B., & Jackson, A. (2016). Concurrent validity of the Independent Reading Level Assessment framework and a state assessment. *Journal of At-Risk Issues, 19*(2), 1–8.

Table 7. IRLA-OAKS Correlation Coefficients

	Concurrent	
2016	n	*r*
All 3–5	2303	.766*
Gr. 3	803	.713*
Gr. 4	720	.775*
Gr. 5	780	.751*

*p<.001

Rhode Island: Rhode Island Comprehensive Assessment System (RICAS)

Rhode Island Study 1: This study was conducted in a large school district in Rhode Island that began using ARC Core in the 2021–2022 school year. The district serves a population of students that is 92% Nonwhite; 89% of students are classified as Low Income and 35% are English Language Learners. In 2022, both the concurrent and predictive correlations between students' IRLA and RICAS ELA scores were strong and statistically significant (see Table 8).

Table 8. IRLA-RICAS ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2022				
Gr. 3	1536	.735*	1461	.724*
Gr. 4	1477	.714*	1403	.698*
Gr. 5	1414	.735*	1335	.725*

*p<.001

Rhode Island Study 2: This study was conducted in a midsize school district in Rhode Island that began using ARC Core in the 2020–2021 school year. The district serves a population of students that is 81% Nonwhite; 98% of students are classified as Low Income and 47% are English Language Learners. In 2022, the concurrent and predictive correlations between Grade 3 students' IRLA and RICAS ELA scores were moderately strong and statistically significant while the correlations in Grades 4 and 5 were strong and statistically significant (see Table 9).

Table 9. IRLA-RICAS ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2022				
Gr. 3	139	.667*	140	.603*
Gr. 4	147	.801*	141	.765*
Gr. 5	166	.762*	156	.793*

*p<.001

Rhode Island Study 3: This study was conducted in a small school district in Rhode Island that began using ARC Core in the 2020–2021 school year. The district serves a population of students that is 6% Nonwhite; 13% of students are classified as Low Income. In 2022, the concurrent and predictive correlations between students' IRLA and RICAS ELA scores approached or exceeded the threshold for strong and were statistically significant (see Table 10).

Table 10. IRLA-RICAS ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2022				
Gr. 3	87	.685*	86	.671*
Gr. 4	85	.724*	83	.707*
Gr. 5	91	.694*	91	.727*

*p<.001

Washington: Smarter Balanced Assessment (SBA)

Washington Study 1: This study was conducted in a large school district in Washington that began implementation of ARC Core in the 2021–2022 school year. This study included students from Grades 3–5, of whom 82% are Nonwhite, 40% are English Language Learners, and 14% are classified as Students with Disabilities. In 2022, the concurrent and predictive correlations between students' IRLA and SBA ELA scores approached the threshold for strong and were statistically significant (see Table 11).

Table 11. IRLA-SBA ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2022 Gr. 3–5	3691	.693*	2487	.686*

**p*<.001

Washington Study 2: This study was conducted in a midsize district in Washington that has been using ARC Core since the 2019–2020 school year. This district's enrollment includes 49% Hispanic/Latino students and 45% White students; 23% of students are English Language Learners and 63% are classified as Low Income. The study was conducted over two years. Washington State postponed the Spring 2021 SBA until the fall of 2021[7], therefore, SBA scores examined from Fall 2021 are reflective of the grade students were in during the 2020–2021 school year. The 2021–2022 SBA scores were administered on a typical schedule in spring of 2022. In both years, the correlations between students' IRLA and SBA ELA scores were strong and statistically significant (see Table 12).

Table 12. IRLA-SBA ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2021 (Fall) Gr. 4–8	1072	.734*	-	-
2022 (Spring) Gr. 3–8	1439	.749*	1289	.735*

**p*<.001

Wyoming: Wyoming's Test of Proficiency and Progress (WY-TOPP)

Wyoming Study 1: This study was conducted in a small Wyoming school district that has been using ARC Core since the 2018–2019 school year. This district's population includes 83% White students, 9% Hispanic students, and 8% students of another race/ethnicity. District-wide, 14% are classified as Students with Disabilities and 9% are classified as Low Income. In 2022 and 2023, the concurrent and predictive correlations between students' IRLA and WY-TOPP ELA scores were strong and statistically significant (see Table 13).

[7] Washington Office of Superintendent of Public Instruction (OSPI), August 2021 Update: News from Assessment and Student Information.

Table 13. IRLA-WY-TOPP ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2022 Gr. 3–6	106	.758*	102	.720*
2023 Gr. 3–6	105	.764*	101	.753*

*p<.001

Wyoming Study 2: This study was conducted in a small Wyoming school district that began using ARC Core during the 2022–2023 school year. This district's population includes 88% White students, 8% Hispanic students, and 4% students of another race/ethnicity. District-wide, 13% are classified as Students with Disabilities and 42% are classified as Low Income. The concurrent and predictive correlations between students' IRLA and WY-TOPP ELA scores were strong and statistically significant (see Table 13).

Table 14. IRLA-WY-TOPP ELA Correlation Coefficients

	Concurrent		Predictive	
	n	*r*	n	*r*
2023 Gr. 3–5	156	.789*	152	.749*

*p<.001

Suggested Citation

Famularo, L., Jackson, A., & Boyd, A. T. (2024). Relationship between the Independent Reading Level Assessment (IRLA) and State ELA Tests: Concurrent and Predictive Validity Evidence. Blue Bell, PA: American Reading Company.

AMERICAN READING COMPANY

ARC's mission is to ensure every student reads and writes on or above grade level. As a triple bottom line company (people, planet, profit), ARC prides itself on being more than a curriculum provider. ARC strives to position itself as a partner with schools and districts, working alongside them to ensure all students are equipped to meet the robust demands of college, career, and life. ARC has a unique ability to produce, iterate, and integrate curricular materials in a way that connects the vision, mission, and values that districts already hold dear. This seamless integration sets ARC apart from other curricula.

www.americanreading.com