RESEARCH REPORT

November 2022

Predictive Validity of the IXL Real-Time Diagnostic Using the Pennsylvania System of School Assessment as Criterion

Xiaozhu An, Ph.D.

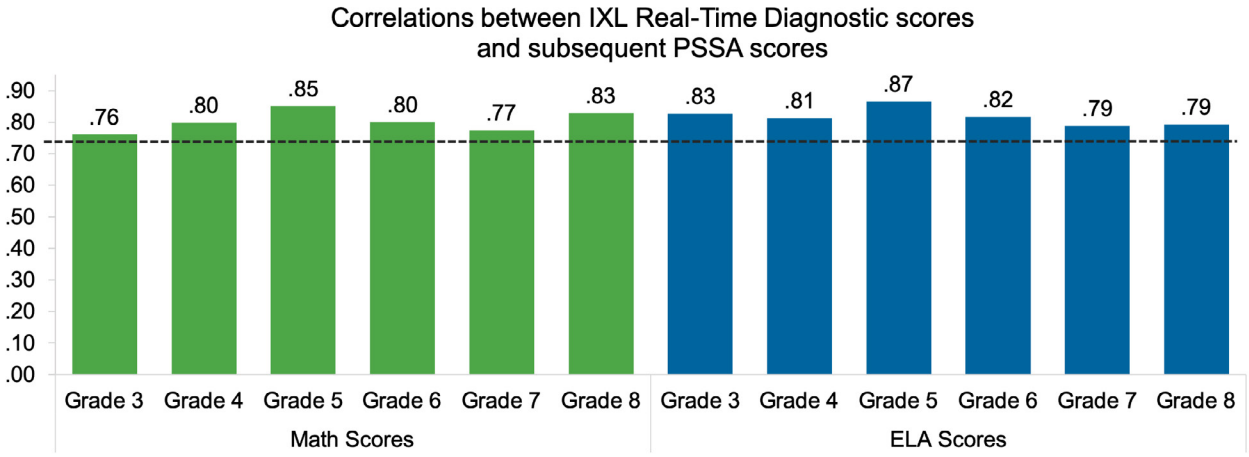
IXL LEARNING 777 Mariners Island Blvd., Suite 600, San Mateo, CA 94404
650-372-4040 | www.ixl.com

Executive Summary

IXL is an end-to-end teaching and learning solution that engages learners in grades Pre-K through 12 with a comprehensive curriculum and a first-of-its-kind assessment suite. A core component of IXL is the IXL Real-Time Diagnostic, an interim assessment that delivers up-to-date insights on students' knowledge levels in math and English language arts. Analyzing student response patterns via Item Response Theory (Lord, 1980), the diagnostic creates personalized learning plans, helping every learner achieve more.

IXL's Real-Time Diagnostic has been validated with other standardized assessments including the NWEA MAP Growth, Star, ILEARN, FSA, and SOL assessments (An, 2021; An, 2022; IXL Learning, 2020b; Schonberg, 2021a; Schonberg, 2021b; Schonberg, 2022). In this study, we examined data from 1,987 students in grades 3 to 8 from five public schools in a suburban school district in Pennsylvania to collect additional predictive validity evidence for the IXL Real-Time Diagnostic using Pennsylvania System of School Assessment (PSSA) as criterion. Key findings include:

- **The IXL Real-Time Diagnostic was a strong predictor of subsequent academic performance as measured by the 2022 PSSA.** For both math and ELA, we found strong positive correlations between IXL Real-Time Diagnostic and PSSA scores, with *r*s = [.76 - .87] across grades 3 to 8.



- **The IXL Real-Time Diagnostic and the 2022 PSSA showed a high degree of alignment in classifying students into proficiency levels.** For both math and ELA, the IXL Real-Time Diagnostic accurately predicted the grade-level proficiency for about 80% of students as determined by subsequent PSSA tests. Students with IXL Diagnostic math scores at or above grade level were 35 times more likely to reach or exceed the PSSA math proficiency standard. Students with IXL Diagnostic ELA scores at or above grade level were 18 times more likely to reach or exceed the PSSA ELA proficiency.

Predictive Validity of the IXL Real-Time Diagnostic Using the Pennsylvania System of School Assessment as Criterion

Background

IXL's Real-Time Diagnostic is an interim assessment developed by educators and subject matter experts that covers material aligned with state standards. As a rigorous assessment rooted in learning sciences research (Bashkov et al., 2021), the IXL Real-Time Diagnostic is widely used in schools and classrooms to assess student knowledge and track student growth.

We have conducted several studies examining the psychometric properties of the IXL Real-Time Diagnostic and have garnered desirable reliability and validity evidence, including coherent internal structure (IXL Learning, 2020a), multi-group measurement invariance (An et al., 2022), high reliability (IXL Learning, 2020a; Schonberg, 2021a), and strong predictive validity coefficients using multiple well-established assessments as criterion measures (An, 2021; An, 2022; IXL Learning, 2020b; Schonberg, 2021a; Schonberg, 2021b; Schonberg, 2022).

The primary goal of this study was to further validate IXL's Diagnostic with a dataset from a suburban school district in Pennsylvania using a new criterion measure: the Pennsylvania System of School Assessment (PSSA). This study also aimed to provide additional predictive validity evidence of IXL's Diagnostic across upper elementary and middle school grade levels.

Study Design and Methodology

We studied a total of 1,987 students from grades 3 to 8 who used IXL during the 2021-22 school year. The students were attending five public schools from a suburban school district in Pennsylvania, which serves more than 4,000 Pre-K through 12 students. The average size of the schools was 546 students, with an average student-teacher ratio close to 13:1. Four of the schools were Title I schools, with about 44% of the students qualifying for free- or reduced-price lunch. Among all the students, 306 (15%) were 3rd-grade students, 319 (16%) were 4th-grade students, 308 (16%) were 5th-grade students, 341 (17%) were 6th-grade students, 374 (19%) were 7th-grade students, and 339 (17%) were 8th-grade students. In terms of gender, 979 (49%) of the students were female. More than half of the students (56%) were White, 23% were Hispanic, and 9% were Black. There were 165 (8%) English language learners, 418 (21%) students with disabilities, and 75 (4%) gifted students.

DATA SOURCES

Data from two sources were used in this study: students' IXL Real-Time Diagnostic scores during the spring 2022 semester and students' end-of-year state assessment data from the Spring 2022 PSSA.

Student IXL Diagnostic Data

IXL Real-Time Diagnostic data during the spring 2022 semester were obtained from IXL's internal database. When a student completes a sufficient number of questions in IXL's Diagnostic, they receive a pinpointed score that indicates their overall grade-level proficiency. For example, a score of 550 indicates that the student has acquired about 50% of 5th-grade material, whereas a score of 600 indicates that the student is ready to learn 6th-grade material. In the analysis, we focused on the available end-of-year diagnostic scores from the date closest to students' state assessment in spring 2022.

Student Assessment Data

The district provided students' Spring 2022 state assessment data in math and ELA. Every spring, the Pennsylvania System of School Assessment (PSSA) math and ELA tests are administered to students in grades 3 to 8. Student performance on the PSSA falls into four performance levels: *Below Basic*, *Basic*, *Proficient*, and *Advanced*, with 600 as the minimum score and 1,000 as the cut score between *Basic* and *Proficient* across all grade levels (see PSSA Technical Report for more details). In this study, *Proficient* and *Advanced* were combined and used to represent student grade-level proficiency. Students' performance on the 2022 PSSA math and ELA tests served as the criterion in evaluating the predictive validity of IXL's Diagnostic. The analysis requires complete data, so students with missing IXL Diagnostic or assessment data were excluded from the analysis, leaving a sample of 1,695 for math and a sample of 1,681 for ELA.

RESEARCH QUESTIONS

This study aimed to answer two research questions, for math and ELA separately:

Research Question 1. Grade-level predictive validity: Was there a strong and positive correlation (i.e., $r >= .70$) between students' IXL Real-Time Diagnostic scores and PSSA scores in each grade?

Research Question 2. Classification alignment: What was the degree of alignment or agreement in classifying students' proficiency levels based on the IXL Real-Time Diagnostic and the PSSA across grades?

ANALYTIC APPROACH

To assess predictive validity, we calculated the correlation coefficients between IXL Real-Time Diagnostic scores and PSSA scores using the Pearson product moment correlation *r*. This correlation coefficient measures the linear relationship between two variables with values ranging from -1.00 to 1.00. An *r* value of 0 indicates no correlation, whereas an *r* value greater than .70 indicates a strong positive relationship (Ratner, 2009). High positive correlations between the IXL Real-Time Diagnostic and a subsequent PSSA test would support the validity of the IXL Real-Time Diagnostic in predicting future academic achievement. Given the nature of the PSSA scoring system with a similar score range across all grade levels, we opted to calculate the correlation between IXL Real-Time Diagnostic scores and PSSA scores for each grade level separately.

To examine classification alignment, we checked the extent to which students were classified as proficient (at or above grade level) versus not proficient (below grade level) by both the IXL Real-Time Diagnostic and the PSSA. For example, were students classified as performing at or above grade level by IXL's Diagnostic also classified as such by the PSSA? Furthermore, we conducted chi-square tests and logistic regressions to evaluate whether students with diagnostic scores at or above grade level were more likely to reach proficiency on the PSSA. We conducted these analyses separately for math and ELA across grades.

Each analysis is accompanied by a test of statistical significance and a probability (*p*) value. The *p*-value is the probability of observing the current or more extreme data, assuming the effect is zero (Cohen, 1994). As such, the smaller the *p*-value, the less likely that the result occurred at random, with .05, .01, and .001 commonly used as thresholds in research practice. Effects associated with *p*-values smaller than these thresholds are considered statistically significant at each of these significance levels.

Results

DESCRIPTIVE STATISTICS

Among the analyzed students for math (n = 1,695), the average IXL Diagnostic math score was 537.18 (SD = 183.72), with 912 (54%) students performing at or above grade level. For ELA (n = 1,681), the average IXL Diagnostic score was 597.22 (SD = 248.34), with 1,007 (60%) students performing at or above grade level.

In terms of PSSA performance, the average 2022 math PSSA scale score was 977.27 (SD = 112.53), with 648 (38%) students reaching proficiency. For ELA, the average 2022 ELA PSSA scale score was 1033.01 (SD = 108.65), with 1,050 (62%) students reaching proficiency.

CORRELATIONS

IXL Math

We found that the IXL Real-Time Diagnostic math assessment and the PSSA math test had a strong positive correlation in all grade levels from 3 to 8, supporting the predictive validity of IXL's Diagnostic math assessment. The correlations ranged from .76 to .85 with ps < .001 and are displayed via scatter plots in Figure 1.

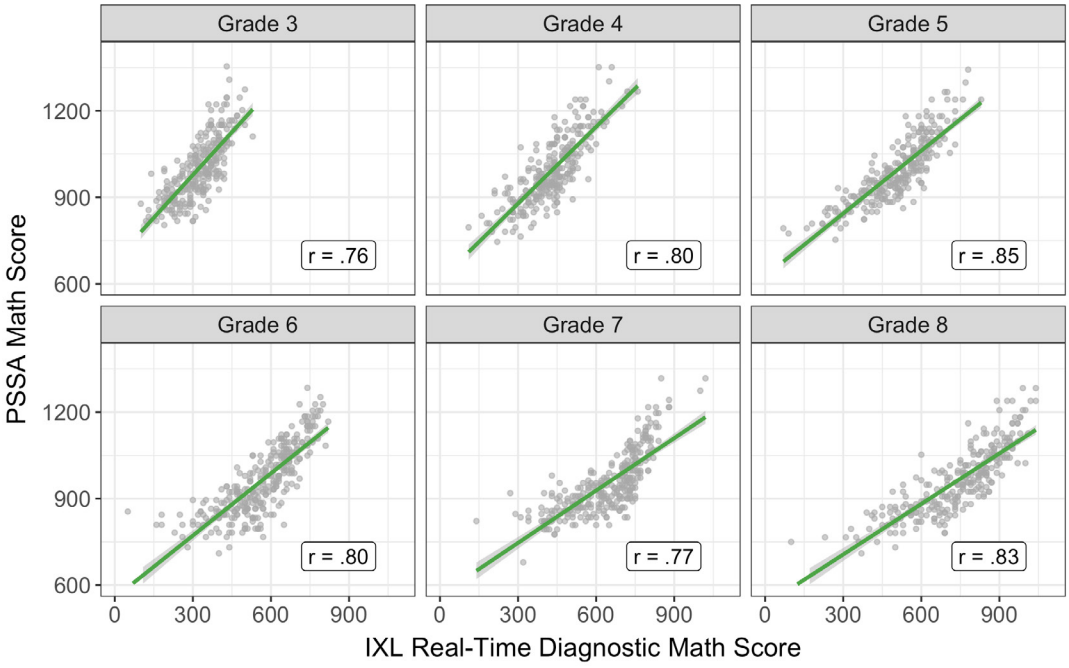


Figure 1. Correlations between IXL Real-Time Diagnostic math scores and PSSA math scores

IXL ELA

Similarly, we found that the IXL Real-Time Diagnostic ELA assessment and the PSSA ELA test had a strong positive correlation in all grade levels from 3 to 8, supporting the predictive validity of IXL's Diagnostic ELA assessment. The correlations ranged from .79 to .87 with *p*s < .001 and are displayed via scatter plots in Figure 2.

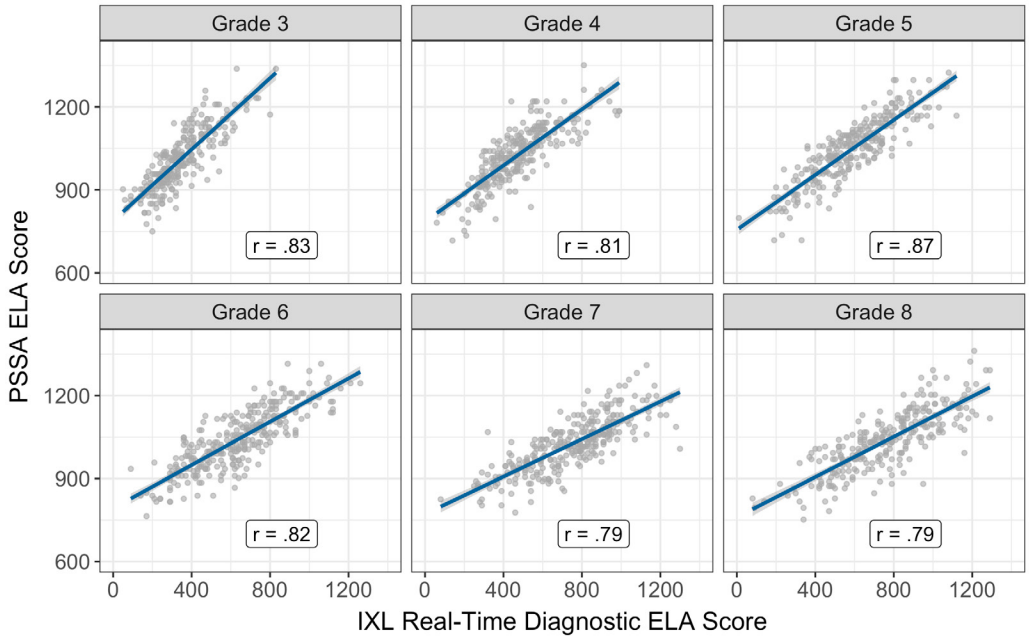


Figure 2. Correlations between IXL Real-Time Diagnostic ELA scores and PSSA ELA scores

CLASSIFICATION ALIGNMENT

IXL Math

We found a significant association between the classifications of students into proficiency levels based on the IXL Real-Time Diagnostic math assessment and the PSSA math test ($\chi^2 = 665.66$, $p < .001$; see Table A1 in Appendix A for details). Overall, the IXL Real-Time Diagnostic math assessment accurately predicted the proficiency status for 1,347 out of 1,695 (79%) students. Specifically, 66% of the students classified as proficient (i.e., performing at or above grade level) by the IXL Real-Time Diagnostic math assessment were also identified as proficient by the PSSA math assessment. Meanwhile, 95% of students classified as below grade level by the IXL Real-Time Diagnostic math assessment also did not reach proficiency on the PSSA math assessment.

To further examine the likelihood of reaching proficiency on the 2022 PSSA math assessment for students whose IXL Real-Time Diagnostic math scores were at or above grade level, we calculated odds ratios using logistic regression. Results showed that students with IXL Diagnostic math scores at or above their grade level were 35 times more likely to reach PSSA proficiency in math than students with IXL Real-Time Diagnostic math scores below grade level (see Table A1 in Appendix A for more details).

IXL ELA

Similarly, we also found a significant association between the classifications of students into proficiency levels based on the IXL Real-Time Diagnostic ELA assessment and the PSSA ELA test ($\chi^2 = 628.89$, $p < .001$; see Table A2 in Appendix A for details). Overall, the IXL Real-Time Diagnostic ELA assessment accurately predicted the proficiency status for 1,370 out of 1,681 (81%) students. Specifically, 87% of the students classified as proficient (i.e., performing at or above grade level) by the IXL Real-Time Diagnostic ELA assessment were also identified as proficient by the PSSA ELA assessment. Meanwhile, 74% of students classified as below grade level by the IXL Real-Time Diagnostic ELA assessment also did not reach proficiency on the PSSA ELA assessment.

To further examine the likelihood of reaching proficiency on the 2022 PSSA ELA assessment for students whose IXL Real-Time Diagnostic ELA scores were at or above grade level, we calculated odds ratios using logistic regression. Results showed that students with IXL Diagnostic ELA scores at or above their grade level were 18 times more likely to reach PSSA proficiency in ELA than students with IXL Real-Time Diagnostic ELA scores below grade level (see Table A2 in Appendix A for more details).

Conclusion

In this study, we investigated the predictive validity of the IXL Real-Time Diagnostic with a novel set of state assessments, the Pennsylvania state assessment—PSSA. We examined test-criterion relationships between the two assessments by grade, providing empirical evidence in each upper elementary and middle school grade level (grades 3 to 8). The IXL Real-Time Diagnostic demonstrated desirable predictive validity properties: we found strong correlations between the IXL

Real-Time Diagnostic and the PSSA in each grade as well as a high degree of alignment in the two measures' classifications of student proficiency.

Coupled with prior studies of construct validity, internal consistency, and predictive validity, our findings corroborate a strong program of reliability and validity for the IXL Real-Time Diagnostic. Our findings also provide further evidence in support of using the diagnostic to identify at-risk students at an early stage, and to keep students on track in the classroom and ready for interim and state assessments.

References

An, X. (2021). *Assessing the predictive validity of the IXL Real-Time Diagnostic among marginalized students* (pp. 1-10). https://www.ixl.com/research/IXL-Real-Time-Diagnostic-Validation-Study-Marginalized-Students.pdf

An, X. (2022). *Examining the predictive validity of the IXL Real-Time Diagnostic using the Florida Standards Assessments as criterion* (pp. 1-8). https://www.ixl.com/materials/us/research/IXL_Real-Time_Diagnostic_Validation_Study_(FL_FSA).pdf

An, X., Xiong. Y., & Bashkov, B. M. (2022). *Multi-group measurement invariance of the IXL Real-Time Diagnostic math assessment* (pp. 1-11). https://www.ixl.com/materials/us/research/Multi-Group_Measurement_Invariance_of_the_IXL_Real-Time_Diagnostic_Math_Assessment.pdf

Bashkov, B. M., Mattison, K., & Hochstein, L. (2021). *IXL design principles: Core features grounded in learning science research* (pp. 1-18). https://www.ixl.com/research/IXL_Design_Principles.pdf

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 49,* 997-1003.

IXL Learning. (2020a). *Assessing the construct validity and the predictive validity of the IXL Real-Time Diagnostic* (pp. 1-15). https://www.ixl.com/research/IXL-Real-Time-Diagnostic-Validation-Study-ILEARN.pdf

IXL Learning. (2020b). *Validation study of the IXL Real-Time Diagnostic using MAP Growth Assessments* (pp. 1-11). https://www.ixl.com/research/IXL-Real-Time-Diagnostic-Validation-Study.pdf

Lord, F. (1980). *Applications of Item Response Theory to practical testing problems.* Mahwah, NJ: Lawrence Erlbaum Associates

Ratner, B. (2009). The correlation coefficient: Its values range between +1/−1, or do they? Journal of Targeting, *Measurement Analysis for Marketing*, *17*, 139-142.

Schonberg C. (2021a). *Demonstrating Grade-Level Predictive Validity of the IXL Real-Time Diagnostic Using the Virginia SOL as Criterion*. https://www.ixl.com/research/IXL_Real-Time_Diagnostic_Grade-Level_Validation_Study_VA_SOL.pdf

Schonberg C. (2021b). *The Impact of IXL on Math and ELA Achievement in a Virginia School District.* (pp. 1-14). https://www.ixl.com/research/The_Impact_of_IXL_in_a_Virginia_School_District.pdf

Schonberg C. (2022). *Assessing the predictive validity of the IXL Real-Time Diagnostic using Star Assessments as criterion* (pp. 1-11). https://www.ixl.com/materials/us/research/Assessing_the_Predictive_Validity_of_the_IXL_Real-Time_Diagnostic_(Star).pdf

Appendix A: Classification Alignment

Table A1. 2x2 Table, Chi-square Test, and Logistic Regression for Math

		PSSA math (n = 1,695)		χ^2	Odds ratio
		Below proficient	At or above proficient		
IXL Real-Time Diagnostic math	Below grade level	741 (94.64%)	42 (5.36%)	665.66 ***	34.94 ***
	At or above grade level	306 (33.55%)	606 (66.45%)		

Note. ***: significance at the .001 level.

Table A2. 2x2 Table, Chi-square Test, and Logistic Regression for ELA

		PSSA ELA (n = 1,681)		χ^2	Odds ratio
		Below proficient	At or above proficient		
IXL Real-Time Diagnostic ELA	Below grade level	497 (73.74%)	177 (26.26%)	628.89 ***	18.29 ***
	At or above grade level	134 (13.31%)	873 (86.69%)		

Note. ***: significance at the .001 level.