RENAISSANCE®

Pathway to Proficiency: Linking Star Reading® and Star Math® to Arizona's Measurement of Educational Readiness to Inform Teaching (AzM2 formerly AzMERIT)

Contents

Figures

Tables

Initial publication June 29, 2018

Introduction

At Renaissance, we know that as an educator, chief among your responsibilities is making decisions about how to allocate limited resources to best serve diverse student needs. A good assessment system supports your efforts, by providing timely, relevant information to help address key questions about which students are on track to meet important standards and who may need additional assistance.

Assessments that identify early any students at risk of missing academic standards are especially useful, as they inform instructional decisions to improve student performance and reduce gaps in achievement. Assessments that do this while taking little time away from instruction are particularly valuable. *Interim assessments*, one of three broad categories of educational assessment,[1] indicate which students are on track to meet later expectations (Perie, Marion, Gong, & Wurtzel, 2007).

This linking study applied results from two interim assessments, Renaissance Star Reading® and Renaissance Star Math®, to help you predict whether individual students are on track or need more assistance to succeed on the year-end summative Arizona's Measurement of Educational Readiness to Inform Teaching (AzMERIT) tests in English Language Arts (ELA) and mathematics in grades 3 through 8.[2]

> Assessments that identify early any students at risk of missing academic standards are especially useful.

Main Findings

Results from the linking analysis revealed that Star Reading and Star Math are accurate predictors of the AzMERIT tests, meaning as an educator you can use Star scores to:

1. Identify early in the year students likely to miss reading and math yearly progress goals in time to make meaningful adjustments to instruction well before the year-end test.

2. Forecast the percent of students at each AzMERIT performance level to serve as an early warning system for building and district administrators and allow redirection of resources as needed.

Study

To determine if Star Reading and Star Math can predict student achievement on the end-of-year AzMERIT tests in ELA and mathematics, we began by linking the score scales for each assessment.

---

1 **Formative assessments** are short and frequent processes, embedded in instruction, that support learning and provide specific feedback on what students know and can do versus where gaps in knowledge exist. **Summative assessments** evaluate whether students have met a set of standards, and serve most commonly as year-end state-mandated tests. **Interim assessments** represent the middle ground, in terms of duration and frequency and can serve purposes including informing instruction, evaluating curriculum and student responsiveness to intervention, and forecasting performance on high-stakes summative year-end tests.

2 Technical manuals are available for Star Reading and Star Math by request to research@renaissance.com.

School-Level Data collection

To find a sample of students who were assessed by both the AzMERIT and Star Assessments, we began by gathering all Star Reading and Star Math test records from 2015 – 2016 and 2016 – 2017 for Arizona. Then, each school's Star Reading and Star Math data were aggregated by grade and subject area. The next step was to match Star data with the AzMERIT data from the same school year by district and school name. To do this, performance level distribution data from the AzMERIT was obtained from the public data provided by the Arizona Department of Education. The file included the number of students tested in each grade and the percentage of students who were at each AzMERIT performance level.[3]

Sample characteristics

Once we determined how many students in each grade at a school were tested on the AzMERIT and took a Star assessment, we calculated the percentage of students assessed on both tests. In each subject and grade at each school, if between 95% and 105% of the students who tested on the AzMERIT had taken a Star assessment, that grade was included in the sample. This method of sample selection ensured that our sample consisted of cases in which all or nearly all the enrolled students who took the AzMERIT also took a Star test within the specified window of time. If a total of approximately 1,000 or more students per grade met the sample criteria, that grade's sample was considered sufficiently large for analysis.

The reading sample included 85,039 Star Reading students from 267 schools. The math sample included 23,984 Star Math students from 92 schools. Table 1 displays by-grade test summaries for the reading and math samples. It also includes percentages of students in the *Minimally Proficient, Partially Proficient, Proficient,* and *Highly Proficient* performance levels, both for the sample and statewide.

Table 1. Performance characteristics of reading and math samples

Star Reading® sample performance										
Grade	Star Reading® students	AzMERIT ELA students	Minimally Proficient		Partially Proficient		Proficient		Highly Proficient	
			Sample	State	Sample	State	Sample	State	Sample	State
3	15,849	15,612	31%	45%	13%	13%	36%	30%	20%	12%
4	14,187	13,895	24%	38%	13%	15%	43%	35%	20%	12%
5	17,389	17,079	20%	32%	21%	24%	39%	32%	20%	12%
6	15,225	14,989	24%	37%	22%	23%	45%	34%	9%	6%
7	11,084	10,859	28%	37%	19%	20%	41%	35%	12%	8%
8	11,305	11,062	37%	44%	23%	22%	29%	25%	11%	9%
Star Math® sample performance										
Grade	Star Math® students	AzMERIT Math students	Minimally Proficient		Partially Proficient		Proficient		Highly Proficient	
			Sample	State	Sample	State	Sample	State	Sample	State
3	4,258	4,184	15%	24%	24%	29%	30%	29%	31%	18%
4	4,067	3,981	16%	27%	24%	27%	39%	34%	21%	12%
5	4,904	4,805	16%	26%	24%	27%	36%	32%	24%	15%
6	4,983	4,880	29%	36%	24%	23%	27%	25%	20%	16%
7	3,556	3,510	38%	46%	22%	20%	24%	20%	16%	14%
8	2,216	2,187	54%	50%	26%	23%	14%	18%	6%	9%

3 The 2015−2016 AzMERIT public data file did not contain student testing counts at the school and grade level. For this school year, enrollment information was used as an estimate of student testing counts.

Results

Scale linkage

Renaissance linked the Star test scale to the AzMERIT by applying equipercentile linking analysis (Kolen & Brennan, 2004). First, we aggregated the sample of schools to calculate the percentage of students performing *Minimally Proficient, Partially Proficient, Proficient, and Highly Proficient* for each subject and grade. Then we analyzed the distribution of Star scores to determine the scaled score corresponding to the same percentile as specific AzMERIT level. For example, as shown in Table 1, 31% of students in our third-grade reading sample were classified as *Minimally Proficient,* 13% *Partially Proficient*, 36% *Proficient*, and 20% *Highly Proficient.* Therefore, the cutscores for proficiency levels in the third grade are at the 31st percentile for *Partially Proficient*, the 44th percentile for *Proficient*, and the 80th percentile for *Highly Proficient*.

AzMERIT cut scores and corresponding Star score equivalents

AzMERIT results are reported in scaled scores that are split into four achievement levels: *Minimally Proficient, Partially Proficient, Proficient*, and *Highly Proficient*. The main purpose in linking Star Reading and Star Math to the AzMERIT was to identify Star scores at the time of the state test that are approximately equivalent to the cut-off scores that separate the AzMERIT levels. Table 2 displays these equivalent Star scores at the time of the state test for grades 3−8.[4] The corresponding AzMERIT cut scores can be found in Appendix B.

Table 2. Star Reading® and Star Math® score equivalents at time of state test for each AzMERIT achievement level range

Star Reading® cut-score equivalents				
Grade	Minimally Proficient	Partially Proficient	Proficient	Highly Proficient
3	< 387	387 − 440	441 − 581	≥ 582
4	< 450	450 − 510	511 − 725	≥ 726
5	< 481	481 − 595	596 − 852	≥ 853
6	< 585	585 − 731	732 − 1132	≥ 1133
7	< 634	634 − 785	786 − 1182	≥ 1183
8	< 774	774 − 970	971 − 1264	≥ 1265
Star Math® cut-score equivalents				
Grade	Minimally Proficient	Partially Proficient	Proficient	Highly Proficient
3	< 541	541 − 610	611 − 662	≥ 663
4	< 609	609 − 679	680 − 760	≥ 761
5	< 663	663 − 741	742 − 818	≥ 819
6	< 739	739 − 800	801 − 855	≥ 856
7	< 780	780 − 831	832 − 881	≥ 882
8	< 826	826 − 877	878 − 912	≥ 913

---

4 The Star Reading and Star Math cut-score equivalents presented in Table 2 apply only to the time of the state test. Some Renaissance reports adjust the Star Reading and Star Math cut-score equivalents based on date.

Accuracy of scale linkage confirmed

Thirteen Arizona schools shared student level AzMERIT scores to explore the accuracy of using Star Reading and Star Math for forecasting AzMERIT performance. The Star Reading sample consisted of 8,517 students and the Star Math sample consisted of 9,006 students. We took students' Star scores from tests taken within 30 days before or after the AzMERIT administration (concurrent sample) and used these scores to examine the accuracy of the linkage to the AzMERIT scale.

Classification diagnostics were derived from counts of correct and incorrect classifications when using Star scores to predict whether a student would achieve proficiency on the AzMERIT. The results indicate that Star Assessments provide an effective means of estimating end-of-year achievement on the AzMERIT.

Predictive Star scores correlate highly with actual AzMERIT scores

To summarize the predictive power of Star Reading and Star Math, we calculated correlations between observed AzMERIT scores and Star scores. As seen in figure 1, the correlations were strong, averaging .82 for both Star Reading and Star Math.

> Star scores have a strong relationship with end-of-year AzMERIT scores.

Figure 1. Star Reading® and Star Math® scores highly correlate with AzMERIT scores

■ Concurrent Star Reading Scale Scores ■ Concurrent Star Math Scale Scores
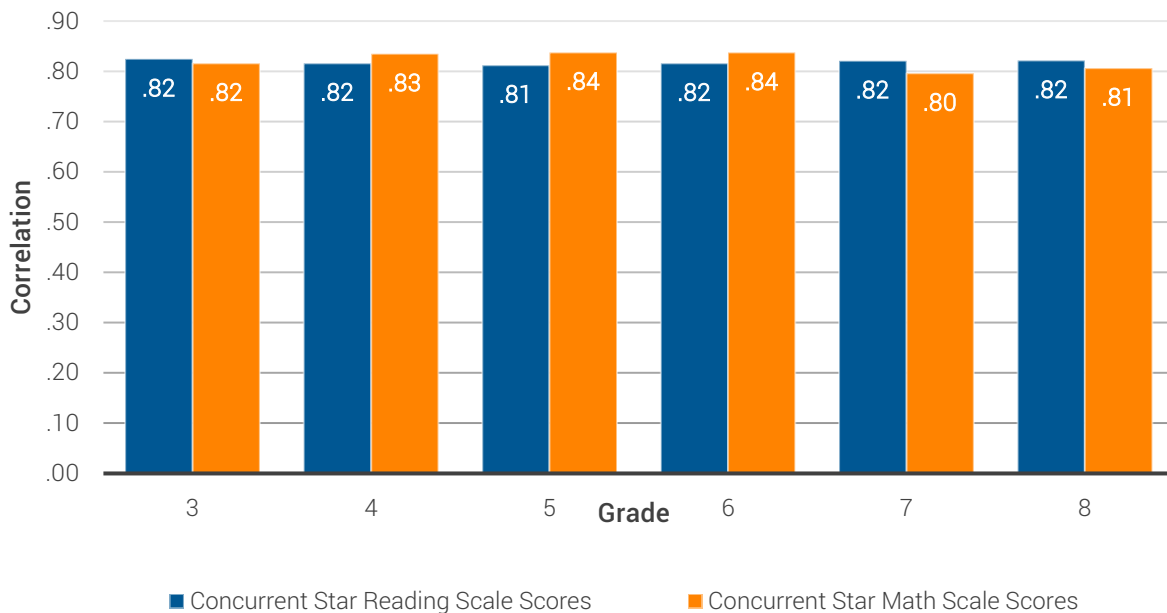
Star scores discriminate well between students who score proficient or not

We compared actual AzMERIT performance to students' estimated AzMERIT performance based on concurrent Star scores and the estimated Star cut score equivalents. Table 3 displays classification diagnostics about whether students were correctly or incorrectly classified as proficient or not on the AzMERIT using concurrent Star scores. On average, students were correctly classified (i.e., overall classification accuracy) 83% of the time by Star Reading and 84% of the time by Star Math.

For Area Under the ROC Curve (AUC), a summary measure of diagnostic accuracy, Star Reading and Star Math averaged .92 (also displayed in table 3). The AUCs met or exceeded the .85 standard set by the National Center on Response to Intervention to indicate convincing evidence that an assessment can accurately predict another assessment result or outcome.

Table 3. Proficiency forecasting using Star Reading® and Star Math® scores yields accurate results
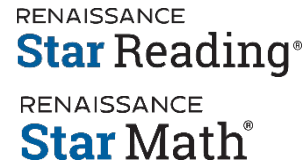
Star Reading®						
Measure	Grade					
	3	4	5	6	7	8
Overall classification accuracy (percentage of correct classifications)	83%	82%	82%	83%	84%	85%
Area Under the ROC Curve	0.91	0.91	0.91	0.91	0.93	0.92
Star Math®						
Measure	Grade					
	3	4	5	6	7	8
Overall classification accuracy (percentage of correct classifications)	81%	82%	83%	85%	85%	89%
Area Under the ROC Curve	0.91	0.92	0.92	0.92	0.93	0.94

Other diagnostic accuracy measures studied:

- ✓ **Sensitivity** represents the percentage of proficient students that were correctly forecasted, which for Star Reading averaged 80% and for Star Math averaged 85%.

- ✓ **Specificity** represents the percentage of not-proficient students that were correctly forecasted, which for Star Reading averaged 84% and for Star Math averaged 83%.

- ✓ **Positive predictive values,** which indicate that when Star scores forecasted students to be proficient, they actually were proficient, were 73% for Star Reading and 76% for Star Math.

- ✓ **Negative predictive values,** which indicate that when Star scores forecasted students to miss proficiency, they actually weren't proficient, were 89% for reading and 90% for math.

- ✓ **Proficiency status projection error**, the difference between actual and projected proficiency rates, indicates how well scores accurately predict proficiency within each grade. Star Reading averaged 4% and Star Math averaged 5% (negative scores indicate under-prediction while positive scores show over-prediction).

Appendix A: About Star Reading® and Star Math®

The computer-adaptive Star Reading and Star Math assessments serve multiple purposes including screening, progress monitoring, instructional planning, forecasting proficiency, standards mastery, and measuring growth. These highly reliable, valid, and efficient standards-based measures of student performance in reading and math provide valuable information regarding the acquisition of skills along a continuum of learning expectations. The assessments can be completed in about 20 minutes, and we recommend administering them two to five times a year for most purposes and more frequently when used for progress monitoring.

RENAISSANCE
Star Reading®

RENAISSANCE
Star Math®

Star Reading and Star Math are highly rated for academic screening and academic progress monitoring by the National Center on Intensive Intervention.

National Center on
INTENSIVE INTERVENTION
at American Institutes for Research ∎

Appendix B: AzMERIT achievement levels

Table B1. AzMERIT achievement level score ranges

AzMERIT achievement level score ranges: ELA				
Grade	Minimally Proficient	Partially Proficient	Proficient	Highly Proficient
3	2395 − 2496	2497 − 2508	2509 − 2540	2541− 2605
4	2400 − 2509	2510 − 2522	2523 − 2558	2559 − 2610
5	2419 − 2519	2520 − 2542	2543 − 2577	2578 − 2629
6	2431 − 2531	2532 − 2552	2553 − 2596	2597 − 2641
7	2438 − 2542	2543 − 2560	2561 − 2599	2600 − 2648
8	2448 − 2550	2551 − 2571	2572 − 2603	2604 − 2658
AzMERIT achievement level score ranges: Mathematics				
Grade	Minimally Proficient	Partially Proficient	Proficient	Highly Proficient
3	3395 − 3494	3495 − 3530	3531 − 3572	3573− 3605
4	3435 − 3529	3530 − 3561	3562 − 3605	3606 − 3645
5	3478 − 3562	3563 − 3594	3595 − 3634	3635 − 3688
6	3512 − 3601	3602 − 3628	3629 − 3662	3663 − 3722
7	3529 − 3628	3629 − 3651	3652 − 3679	3680 − 3739
8	3566 − 3649	3650 − 3672	3673 − 3704	3705 − 3776

References

Kolen, M. J., & Brennan, R. R. (2004). *Test equating scaling and linking: Methods and practices*. New York, NY: Springer Science+Business Media.

Perie, M., Marion, S., Gong, B., & Wurtzel, J. (2007). *The role of interim assessments in a comprehensive assessment system.* Aspen, CO: Aspen Institute.

Renaissance Learning. (2018a). *Star Math: Technical manual*. Wisconsin Rapids, WI: Author. Available by request to research@renaissance.com

Renaissance Learning. (2018b). *Star Reading: Technical manual*. Wisconsin Rapids, WI: Author. Available by request to research@renaissance.com

Independent technical reviews of Star Reading® and Star Math®

U.S. Department of Education: National Center on Intensive Intervention. (2018a). *Review of academic progress monitoring tools* [Review of Star Math]. Washington, DC: Author. Retrieved from https://charts.intensiveintervention.org/chart/progress-monitoring

U.S. Department of Education: National Center on Intensive Intervention. (2018b). *Review of academic progress monitoring tools* [Review of Star Reading]. Washington, DC: Author. Retrieved from https://charts.intensiveintervention.org/chart/progress-monitoring

U.S. Department of Education: National Center on Intensive Intervention. (2018c). *Review of academic screening tools* [Review of Star Math]. Washington, DC: Author. Retrieved from https://charts.intensiveintervention.org/chart/academic-screening

U.S. Department of Education: National Center on Intensive Intervention. (2018d). *Review of academic screening tools* [Review of Star Reading]. Washington, DC: Author. Retrieved from https://charts.intensiveintervention.org/chart/academic-screening

U.S. Department of Education: National Center on Response to Intervention. (2010a). *Review of progress monitoring tools* [Review of STAR Math]. Washington, DC: Author. Retrieved from https://web.archive.org/web/20120813035500/http://www.rti4success.org/pdf/progressMonitoringGOM.pdf

U.S. Department of Education: National Center on Response to Intervention. (2010b). *Review of progress monitoring tools* [Review of STAR Reading]. Washington, DC: Author. Retrieved from https://web.archive.org/web/20120813035500/http://www.rti4success.org/pdf/progressMonitoringGOM.pdf

U.S. Department of Education: National Center on Response to Intervention. (2011a). *Review of screening tools* [Review of STAR Math]. Washington, DC: Author. Retrieved from https://web.archive.org/web/20171027185735/http://www.rti4success.org:80/resources/tools-charts/screening-tools-chart

U.S. Department of Education: National Center on Response to Intervention. (2011b). *Review of screening tools* [Review of STAR Reading]. Washington, DC: Author. Retrieved from https://web.archive.org/web/20171027185735/http://www.rti4success.org:80/resources/tools-charts/screening-tools-chart