

Relating Star Reading[®] and Star Math[®] to the Florida Standards Assessments (FSA) Performance



Contents

Introduction.....	3
Main Findings.....	3
Study.....	3
Results.....	6
Appendix A: About Star Reading and Star Math.....	8
Appendix B: FSA achievement levels.....	8
References.....	9

Figures

Figure 1. Star Reading and Star Math scores highly correlate with FSA.....	4
--	---

Tables

Table 1. Star Reading and Star Math score equivalents for each FSA achievement level range.....	5
Table 2. Proficiency forecasting using Star Reading and Star Math scores yields accurate results.....	7
Table B1. FSA achievement level score ranges.....	8

Initial publication June 15, 2020

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 et al., 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 Florida Standards Assessments (FSA) for English language arts in grades 3 through 10 and mathematics 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 FSA, 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 FSA 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 FSA in English language arts 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.

Data collection

Using a secure data-matching procedure compliant with the federal Family Educational Rights and Privacy Act (FERPA), staff from seven Florida districts provided Renaissance with state summative test scores for students who had taken Star Reading or Star Math during the 2014–2015, 2017–2018, and 2018–2019 school years. Each record included a student’s FSA scores and was matched with all Star scores for that year.

Sample characteristics

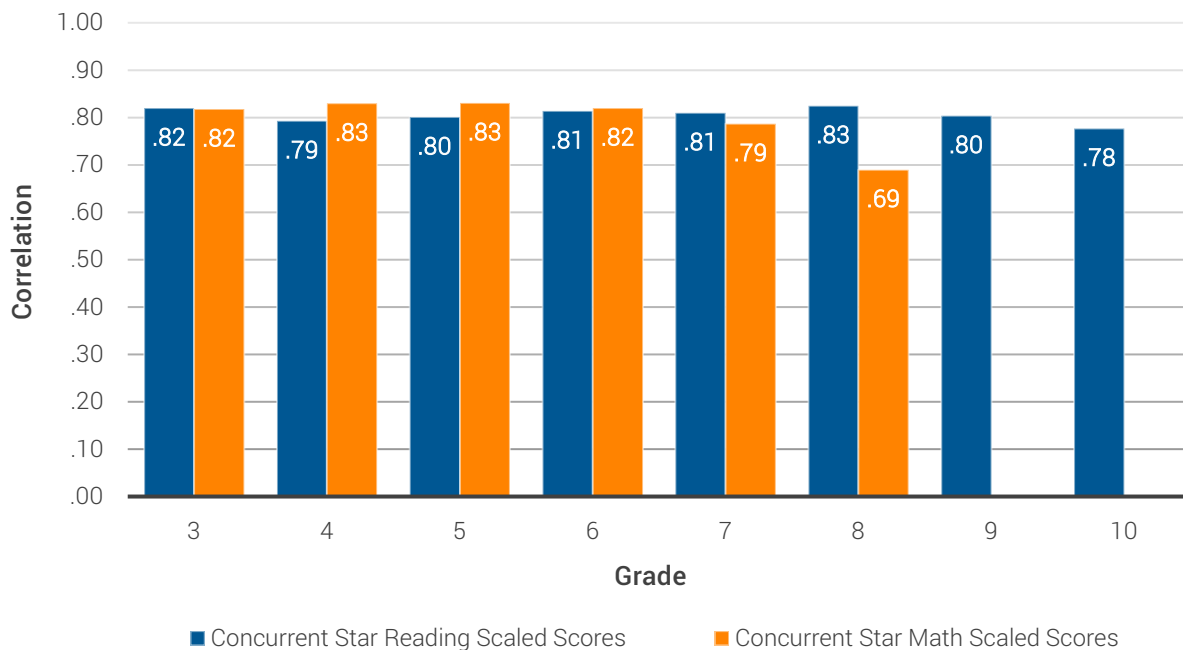
Renaissance divided the data into two samples. The **concurrent sample** included students’ scores for all Star tests taken within 30 days before or after the mid-date of the FSA administration window. This sample numbered 99,669 records in grades 3–10 with matched FSA English language arts and Star Reading scores and 73,071 records in grades 3–8 with matched FSA mathematics and Star Math scores. In each grade, we then set aside scores from a subset of these students—10%—as a **holdout sample** to use only to evaluate the scale linkage.

The linking analysis revealed that Star Reading and Star Math are accurate predictors of the FSA.

Correlations

Before linking Star tests with the FSA, we ensured there was a strong relationship between the test scales. As seen in figure 1, the correlations were strong, averaging .81 and .80 between FSA and Star Reading and Star Math, respectively.

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



Scale linkage

Renaissance then linked the score scales for the Star Reading/Star Math and the FSA in English language arts and mathematics by applying equipercentile linking analysis (Kolen & Brennan, 2004). The concurrent sample (sans the holdout sample) was used in the linking (scores from Star tests taken 30 days before or after the FSA testing mid-date), and the result was a table of FSA scores for each possible Star score.

FSA cut scores and corresponding Star score equivalents

FSA results are reported in scaled scores that describe each student's location on an achievement continuum ranging from 240 to 412 for ELA and 240 to 393 for mathematics and using five achievement levels: *Level 1*, *Level 2*, *Level 3*, *Level 4*, and *Level 5*. A main purpose in linking Star Reading and Star Math to the FSA was to identify Star scores approximately equivalent to the cut-off scores that separate the FSA achievement levels. Table 1 displays these equivalent Star scores for grade 3–10 in reading and 3–8 in math.³ The corresponding FSA cut scores can be found in the Appendix B. In addition to the scores below, cut scores for high school math are also available in a separate [Florida End-of-Course linking report](#).

Table 1. Star Reading® and Star Math® score equivalents for each FSA achievement level range

Star Reading® cut-score equivalents					
Grade	Level 1	Level 2	Level 3	Level 4	Level 5
3	< 312	312 – 422	423 – 538	539 – 685	≥ 686
4	< 407	407 – 516	517 – 650	651 – 893	≥ 894
5	< 475	475 – 608	609 – 786	787 – 1055	≥ 1056
6	< 510	510 – 670	671 – 873	874 – 1182	≥ 1183
7	< 574	574 – 769	770 – 951	952 – 1221	≥ 1222
8	< 597	597 – 811	812 – 1048	1049 – 1293	≥ 1294
9	< 684	684 – 915	916 – 1173	1174 – 1321	≥ 1322
10	< 723	723 – 974	975 – 1228	1229 – 1339	≥ 1340
Star Math® cut-score equivalents					
Grade	Level 1	Level 2	Level 3	Level 4	Level 5
3	< 547	547 – 603	604 – 659	660 – 717	≥ 718
4	< 626	626 – 675	676 – 737	738 – 792	≥ 793
5	< 666	666 – 734	735 – 795	796 – 844	≥ 845
6	< 687	687 – 767	768 – 829	830 – 892	≥ 893
7	< 708	708 – 780	781 – 847	848 – 897	≥ 898
8	< 717	717 – 786	787 – 848	849 – 898	≥ 899

³ The Star Reading and Star Math cut-score equivalents presented in table 1 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.

Results

Accuracy of scale linkage confirmed

In evaluating the accuracy of the scale linkage, we used two methods to examine the differences between students' observed (actual) FSA scores and our Star equivalents: (1) computing the RMSEL (the root mean squared errors of linking) using the scores from the linking study, and (2) applying the holdout sample, consisting of the subset of concurrent scores not used in the linking, to the linking results. Results showed that our linking computation performed as intended.

Star scores discriminate well between students who score proficient or not

Using the holdout sample, we were able to compare how concurrent Star scores aligned with the observed FSA scores. Table 2 displays classification diagnostics about whether students were correctly or incorrectly classified as proficient or not on the FSA using concurrent Star scores. On average, students were correctly classified (i.e., overall classification accuracy) 83% of the time for reading and 82% of the time for math.

For Area Under the ROC Curve (AUC), a summary measure of diagnostic accuracy, Star Reading and Star Math averaged .91 and .90, respectively (also displayed in table 2). The AUCs meet or exceed 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 2. Proficiency forecasting using Star Reading® and Star Math® scores yields accurate results

Star Reading®									
Measure	Grade								
	3	4	5	6	7	8	9	10	
Overall classification accuracy (percentage of correct classifications)	83%	82%	81%	83%	82%	84%	85%	84%	
Area Under the ROC Curve	0.92	0.90	0.90	0.91	0.91	0.92	0.90	0.91	
Star Math®									
Measure	Grade								
	3	4	5	6	7	8			
Overall classification accuracy (percentage of correct classifications)	84%	83%	83%	85%	83%	78%			
Area Under the ROC Curve	0.92	0.92	0.91	0.93	0.89	0.85			

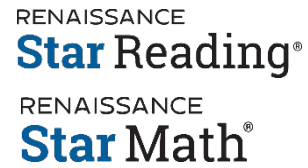
Other diagnostic accuracy measures studied:

- ✓ **Sensitivity** represents the percentage of proficient students that were correctly forecasted, which averaged 79% for both Star Reading and Star Math.
- ✓ **Specificity** represents the percentage of not-proficient students that were correctly forecasted, which for Star Reading averaged 85% and for Star Math averaged 84%.
- ✓ **Positive predictive values** indicate that when Star scores forecasted students to be proficient, they actually were proficient 79% for both Star Reading and Star Math.

- ✓ **Negative predictive values** indicate that when Star scores forecasted students to miss proficiency, they actually weren't proficient 85% of the time for reading and 84% of the time 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 average 0% and Star Math also averaged 0% (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.



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: FSA achievement levels

Table B1. FSA achievement level score ranges

FSA achievement level score ranges: English language arts					
Grade	Level 1	Level 2	Level 3	Level 4	Level 5
3	240 – 284	285 – 299	300 – 314	315 – 329	330 – 360
4	251 – 296	297 – 310	311 – 324	325 – 339	340 – 372
5	257 – 303	304 – 320	321 – 335	336 – 351	352 – 385
6	259 – 308	309 – 325	326 – 338	339 – 355	356 – 391
7	267 – 317	318 – 332	333 – 345	346 – 359	360 – 397
8	274 – 321	322 – 336	337 – 351	352 – 365	366 – 403
9	276 – 327	328 – 342	343 – 354	355 – 369	370 – 407
10	284 – 333	334 – 349	350 – 361	362 – 377	378 – 412
FSA achievement level score ranges: Mathematics					
Grade	Level 1	Level 2	Level 3	Level 4	Level 5
3	240 – 284	285 – 296	297 – 310	311 – 326	327 – 360
4	251 – 298	299 – 309	310 – 324	325 – 339	340 – 376
5	256 – 305	306 – 319	320 – 333	334 – 349	350 – 388
6	260 – 309	310 – 324	325 – 338	339 – 355	356 – 390
7	269 – 315	316 – 329	330 – 345	346 – 359	360 – 391
8	273 – 321	322 – 336	337 – 352	353 – 364	365 – 393

References

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

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

Renaissance Learning. (2019a). *Star Math technical manual*. Available by request to research@renaissance.com

Renaissance Learning. (2019b). *Star Reading technical manual*. 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]. <https://charts.intensiveintervention.org/aprogressmonitoring>

U.S. Department of Education: National Center on Intensive Intervention. (2018b). *Review of academic progress monitoring tools* [Review of Star Reading]. <https://charts.intensiveintervention.org/aprogressmonitoring>

U.S. Department of Education: National Center on Intensive Intervention. (2019a). *Review of academic screening tools* [Review of Star Math]. <https://charts.intensiveintervention.org/ascreening>

U.S. Department of Education: National Center on Intensive Intervention. (2019b). *Review of academic screening tools* [Review of Star Reading]. <https://charts.intensiveintervention.org/ascreening>

U.S. Department of Education: National Center on Response to Intervention. (2010a). *Review of progress monitoring tools* [Review of STAR Math]. <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]. <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]. <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]. <https://web.archive.org/web/20171027185735/http://www.rti4success.org:80/resources/tools-charts/screening-tools-chart>