

RENAISSANCE

TECHNICAL PAPER | 2021–2022 School Year

Relating Star Reading and Star Math to the Georgia Milestones Performance



Contents

3	Introduction
3	Main Findings
4	Study
6	Results
9	Appendix A: About Star Reading and Star Math
9	Appendix B: Milestones achievement levels
10	Appendix C: Star Reading and Star Math Enterprise score equivalents
11	References

Figures

5	Figure 1. Star Reading and Star Math scores highly correlate with Milestones scores
7	Figure 2. Projected scores from Star Reading and Star Math highly correlate with Milestones scores

Tables

6	Table 1. Star Reading and Star Math Unified score equivalents for each Milestones achievement level range
8	Table 2. Proficiency forecasting using Star Reading and Star Math scores yields accurate results
9	Table B1. Milestones achievement level score ranges
10	Table C1. Star Reading and Star Math Enterprise score equivalents for each Milestones achievement level range

Initial publication July 1, 2017

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 Milestones End-of-Grade 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 Milestones assessments, meaning as a Georgia 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 Milestones performance level to serve as an early warning system for building and district administrators and allow redirection of resources as needed.

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.

Study

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

Data collection

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

Sample characteristics

Renaissance divided the Georgia 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 Milestones administration window. This sample numbered 54,994 records in grades 3–8 with matched Milestones and Star Reading scores and 51,279 records in those grades with matched Milestones 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.

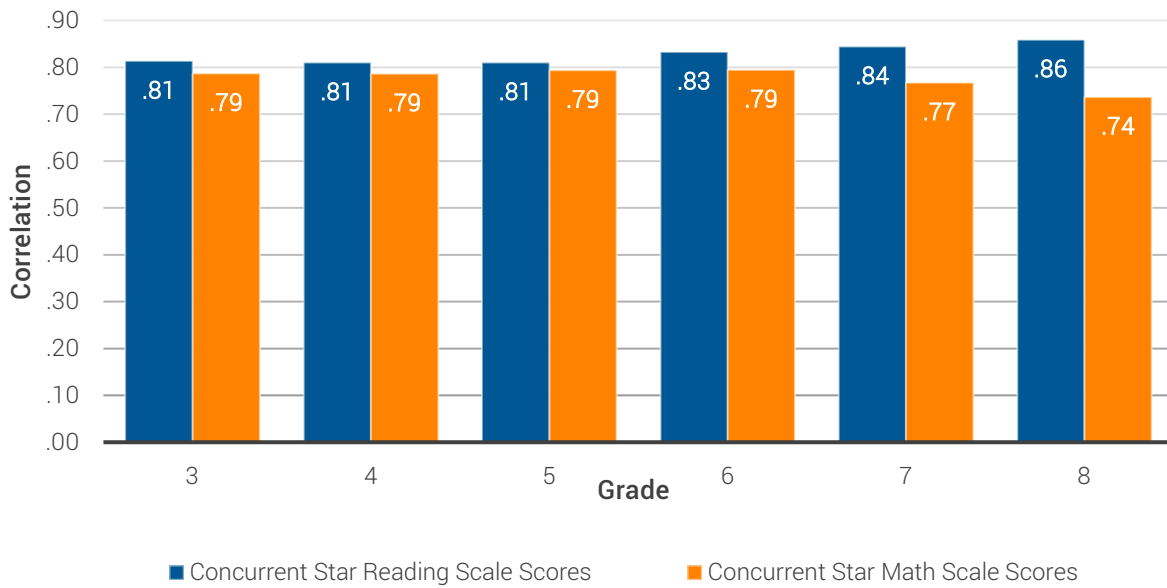
The linking analysis revealed that Star Reading and Star Math are accurate predictors of the Georgia Milestones tests.

The **predictive** sample, which included 44,436 students for reading and 44,745 students for math, included Star scores for tests taken more than 30 days before the mid-date in the Milestones testing window.

Correlations

Before linking Star tests with the Milestones, we ensured there was a strong relationship between the test scales. As seen in figure 1, the correlations were strong, averaging .83 and .78 between Milestones and Star Reading and Star Math, respectively.

Figure 1. Star Reading and Star Math scores highly correlate with Milestones scores



Scale linkage

Renaissance then linked the score scales for the Star Reading/Star Math and the Milestones in ELA and mathematics by applying equipercntile linking analysis (Kolen & Brennan, 2004) in grades 3–8. The concurrent sample (sans the holdout sample) was used in the linking (scores from all Star tests taken within 30 days before or after the Milestones testing mid-date), and the result was a table of Milestones scores for each possible Star score.

The predictive sample was then used to evaluate if the linking results could accurately predict student performance on the Milestones with Star data from earlier in the school year. To do so, we took students’ Star scores from tests taken more than 30 days prior to the Milestones testing mid-date and used national growth norms (Renaissance Learning, 2016a, 2016b) to project what their Star scores would be at the mid-date. Then the scale linkage table was used to look up the projected Star scores (or the average of the projected scores for students with multiple Star scores in the predictive sample) to see how they translated to the Milestones scale.

Georgia cut scores and corresponding Star score equivalents

Milestones results are reported in scaled scores that describe each student’s location on an achievement continuum ranging from approximately 180 to 830 and using four achievement levels: *Beginning Learner*, *Developing Learner*, *Proficient Learner*, and *Distinguished Learner*. A main purpose in linking Star Reading and Star Math to the Milestones was to identify Star scores at the time of the state test that are approximately equivalent to the cut-off scores that separate the Georgia achievement levels. Table 1 displays these equivalent Star Unified scores at the time of the state test for grades 3–8.³ The corresponding Milestones cut scores can be found in Appendix B.

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.

Table 1. Star Reading and Star Math Unified score equivalents at time of state test for each Milestones achievement level range

Star Reading Unified cut-score equivalents				
Grade	Beginning Learner	Developing Learner	Proficient Learner	Distinguished Learner
3	< 949	949 – 1003	1004 – 1055	≥ 1056
4	< 973	973 – 1032	1033 – 1094	≥ 1095
5	< 997	997 – 1054	1055 – 1127	≥ 1128
6	< 1018	1018 – 1073	1074 – 1147	≥ 1148
7	< 1033	1033 – 1093	1094 – 1175	≥ 1176
8	< 1034	1034 – 1108	1109 – 1181	≥ 1182
Star Math Unified cut-score equivalents				
Grade	Beginning Learner	Developing Learner	Proficient Learner	Distinguished Learner
3	< 941	941 – 1002	1003 – 1052	≥ 1053
4	< 973	973 – 1040	1041 – 1096	≥ 1097
5	< 1017	1017 – 1085	1086 – 1126	≥ 1127
6	< 1032	1032 – 1104	1105 – 1154	≥ 1155
7	< 1045	1045 – 1117	1118 – 1158	≥ 1159
8	< 1070	1070 – 1139	1140 – 1170	≥ 1171

Results

Accuracy of scale linkage confirmed

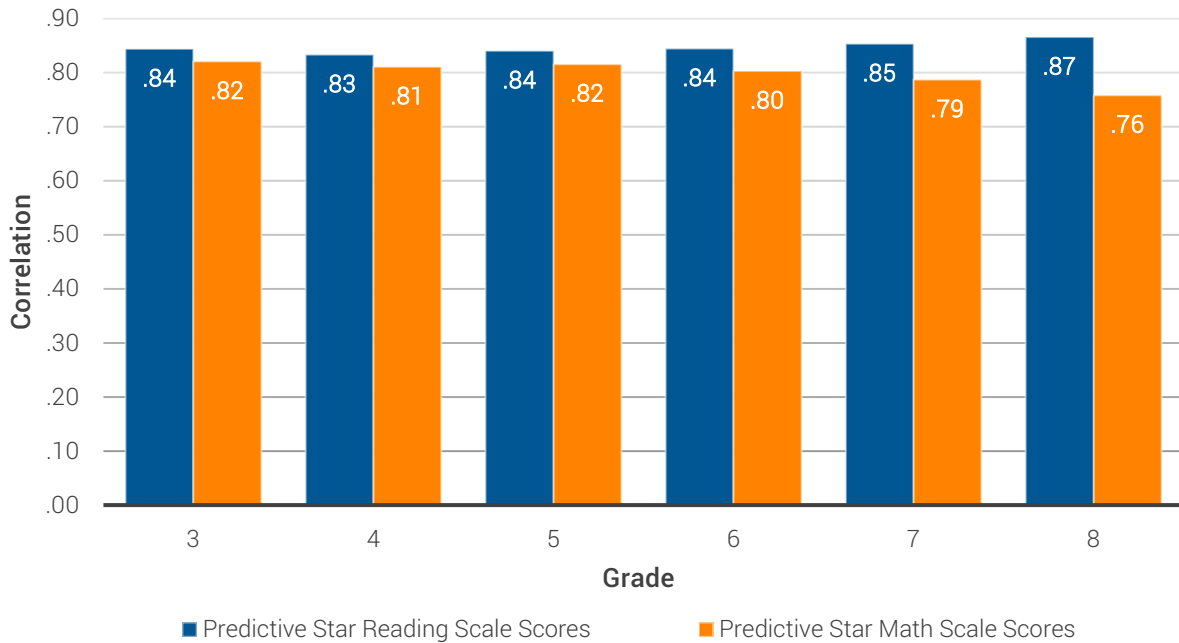
In evaluating the accuracy of the scale linkage, we used two methods to examine the differences between students' observed (actual) Milestones scores and our Star equivalents: (1) computing the RMSEL (the root mean squared errors of linking) using the scores from the linking study, and (2) applying the 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.

Predictive Star scores correlate highly with actual Milestones scores

To summarize the predictive power of Star Reading and Star Math, we calculated correlations between observed (actual) Milestones scores and projected Star scores. As figure 2 shows, the predictive correlation showed a strong relationship between the assessments (similar to the correlations from the concurrent sample, see figure 1, p. 4), indicating that earlier Star scores have a strong relationship with end-of-year Milestones scores. For reading, the correlations averaged .85 and for math, the associations were also high, averaging .80.

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

Figure 2. Projected scores from Star Reading® and Star Math® highly correlate with Milestones scores



Star scores discriminate well between students who score proficient or not

Using the sample of actual Milestones scores, we were able to compare how our projected Star scores aligned with the observed Milestones scores. Table 2 displays classification diagnostics about whether students were correctly or incorrectly classified as proficient or not on the Milestones using projected Star scores. On average, students were correctly classified (i.e., overall classification accuracy) 87% of the time for reading and 89% of the time for math.

For Area Under the ROC Curve (AUC), a summary measure of diagnostic accuracy, Star Reading averaged .94 and Star Math averaged .94 (also displayed in table 2). The AUCs far 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	Measure					
	3	4	5	6	7	8
Overall classification accuracy (percentage of correct classifications)	87%	87%	86%	87%	87%	87%
Area Under the ROC Curve	0.94	0.94	0.93	0.94	0.94	0.94
Star Math						
Measure	Measure					
	3	4	5	6	7	8
Overall classification accuracy (percentage of correct classifications)	88%	88%	88%	89%	90%	90%
Area Under the ROC Curve	0.94	0.94	0.94	0.95	0.95	0.94

Other diagnostic accuracy measures studied:

- ✓ **Sensitivity** represents the percentage of proficient students that were correctly forecasted, which for Star Reading averaged 79% and for Star Math averaged 77%.
- ✓ **Specificity** represents the percentage of not-proficient students that were correctly forecasted, which for Star Reading averaged 90% and for Star Math averaged 93%.
- ✓ **Positive predictive values** indicate that when Star scores forecasted students to be proficient, they actually were proficient 78% of the time for Star Reading and 80% of the time for Star Math.
- ✓ **Negative predictive values** indicate that when Star scores forecasted students to miss proficiency, they actually weren't proficient 91% of the time for reading and 92% 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 averaged 0% and Star Math averaged -1% (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: Milestones achievement levels

Table B1. Milestones achievement level score ranges

Milestones achievement level score ranges: English Language Arts				
Grade	Beginning Learner	Developing Learner	Proficient Learner	Distinguished Learner
3	180 – 474	475 – 524	525 – 580	581 – 830
4	210 – 474	475 – 524	525 – 573	574 – 775
5	210 – 474	475 – 524	525 – 586	587 – 760
6	140 – 474	475 – 524	525 – 598	599 – 820
7	165 – 474	475 – 524	525 – 591	592 – 785
8	225 – 474	475 – 524	525 – 580	581 – 730
Milestones achievement level score ranges: Mathematics				
Grade	Beginning Learner	Developing Learner	Proficient Learner	Distinguished Learner
3	290 – 474	475 – 524	525 – 579	580 – 705
4	270 – 474	475 – 524	525 – 584	585 – 715
5	265 – 474	475 – 524	525 – 579	580 – 725
6	285 – 474	475 – 524	525 – 579	580 – 700
7	265 – 474	475 – 524	525 – 579	580 – 740
8	275 – 474	475 – 524	525 – 578	579 – 755

Appendix C: Star Reading and Star Math Enterprise score equivalents

Since the 2017–2018 school year, Renaissance has offered educators the additional option of reporting student Star scores on an alternative scale referred to as the Unified scale. Table C1 below includes the Enterprise scaled scores for Star Reading and Star Math that correspond to achievement levels on the Milestones assessments for English language arts and Mathematics. Table 1 contains the Unified scaled scores for Star Reading and Star Math that correspond to achievement levels on the Milestones assessments for English language arts and Mathematics.

Table C1. Star Reading and Star Math Enterprise score equivalents at time of state test for each Milestones achievement level range

Star Reading Enterprise cut-score equivalents				
Grade	Beginning Learner	Developing Learner	Proficient Learner	Distinguished Learner
3	< 346	346 – 478	479 – 640	≥ 641
4	< 401	401 – 563	564 – 843	≥ 844
5	< 463	463 – 639	640 – 994	≥ 995
6	< 518	518 – 721	722 – 1164	≥ 1165
7	< 563	563 – 837	838 – 1296	≥ 1297
8	< 566	566 – 905	906 – 1312	≥ 1313
Star Math Enterprise cut-score equivalents				
Grade	Beginning Learner	Developing Learner	Proficient Learner	Distinguished Learner
3	< 544	544 – 643	644 – 725	≥ 726
4	< 595	595 – 706	707 – 796	≥ 797
5	< 668	668 – 778	779 – 846	≥ 847
6	< 692	692 – 810	811 – 891	≥ 892
7	< 713	713 – 831	832 – 897	≥ 898
8	< 753	753 – 866	867 – 916	≥ 917

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. (2016a). *Star Math technical manual*. Available by request to research@renaissance.com

Renaissance Learning. (2016b). *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>