

RENAISSANCE®

TECHNICAL PAPER

Pathway to Proficiency: Linking Star Reading® and Star Math® to the New York State Testing Program (NYSTP)



Contents

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

Figures

6	Figure 1. Star Reading and Star Math highly correlate with NYSTP scores
---	---

Tables

4	Table 1. Performance characteristics of reading and math samples
5	Table 2. Star Reading and Star Math score equivalents for each NYSTP achievement level range
7	Table 3. Proficiency forecasting using Star Reading and Star Math scores yields accurate results
8	Table B1. NYSTP achievement level score ranges

Initial publication July 24, 2019

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 NYSTP 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 NYSTP, 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 NYSTP 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 NYSTP tests in ELA and mathematics, we began by linking the score scales for each assessment.

School-Level Data collection

To find a sample of students who were assessed by both the NYSTP and Star Assessments, we began by gathering all Star Reading and Star Math test records from 2017–2018 for New York. Then, each school's

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.

Star Reading and Star Math data were aggregated by grade and subject area. The next step was to match Star data with the NYSTP data from the same school year by district and school name. To do this, performance level distribution data from the NYSTP was obtained from the public data provided by the New York Department of Education. The file included the number of students tested in each grade and the percentage of students who were *Level 1*, *Level 2*, *Level 3*, or *Level 4*.

Sample characteristics

Once we determined how many students in each grade at a school were tested on the NYSTP ELA and took a Star Reading assessment, we calculated the percentage of students assessed on both tests. Then we repeated this exercise for the math assessments. In each grade at each school, if between 95% and 105% of the students who tested on the NYSTP 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 NYSTP 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 18,982 Star Reading students from 148 schools. The math sample included 16,513 Star Math students from 128 schools. Table 1 displays by-grade test summaries for the reading and math samples. It also includes percentages of students in the *Level 1*, *Level 2*, *Level 3*, and *Level 4* 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	NYSTP ELA students	Level 1		Level 2		Level 3		Level 4	
			Sample	State	Sample	State	Sample	State	Sample	State
3	3,560	3,483	18%	21%	32%	29%	43%	43%	7%	7%
4	3,650	3,569	19%	21%	33%	33%	29%	29%	18%	17%
5	3,626	3,530	33%	35%	30%	28%	22%	22%	14%	15%
6	2,529	2,485	28%	30%	23%	21%	22%	23%	27%	26%
7	3,301	3,249	29%	32%	31%	31%	28%	28%	12%	9%
8	2,316	2,271	19%	23%	33%	37%	27%	27%	21%	13%
Star Math® sample performance										
Grade	Star Math® students	NYSTP Math students	Level 1		Level 2		Level 3		Level 4	
			Sample	State	Sample	State	Sample	State	Sample	State
3	4,222	4,115	24%	30%	22%	19%	31%	28%	23%	23%
4	3,514	3,455	26%	34%	26%	24%	23%	20%	25%	22%
5	3,737	3,669	33%	35%	24%	20%	23%	23%	21%	22%
6	2,088	2,054	31%	41%	25%	22%	23%	20%	22%	17%
7	2,038	2,035	33%	42%	25%	22%	23%	19%	18%	17%
8	914	893	39%	44%	31%	28%	18%	15%	12%	13%

Results

Scale linkage

Renaissance linked the Star test scale to the NYSTP by applying equipercentile linking analysis (Kolen & Brennan, 2004). First, we aggregated the sample of schools to calculate the percentage of students categorized as *Level 1*, *Level 2*, *Level 3*, and *Level 4* 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 NYSTP level. For example, as shown in Table 1, 18% of students in the third-grade reading sample were classified as *Level 1*, 32% were classified as *Level 2*, 43% were classified as *Level 3*, and 7% were classified as *Level 4*. Therefore, the cut score was at the 18th percentile for *Level 2*, the 50th percentile for *Level 3*, and the 93rd percentile for *Level 4*.

NYSTP cut scores and corresponding Star score equivalents

NYSTP results are reported in scaled scores that are split into four achievement levels: *Level 1*, *Level 2*, *Level 3*, and *Level 4*. The main purpose in linking Star Reading and Star Math to the NYSTP was to identify Star scores at the time of the state test that are approximately equivalent to the cut-off scores that separate the NYSTP levels. Table 2 displays these equivalent Star scores at the time of the state test for grades 3-8.³ The corresponding NYSTP cut scores can be found in Appendix B.

Table 2. Star Reading[®] and Star Math[®] score equivalents for each NYSTP achievement level range

Star Reading [®] cut-score equivalents				
Grade	Level 1	Level 2	Level 3	Level 4
3	< 291	291 – 423	424 – 676	≥ 677
4	< 346	346 – 494	495 – 677	≥ 678
5	< 486	486 – 615	616 – 804	≥ 805
6	< 519	519 – 634	635 – 809	≥ 810
7	< 596	596 – 834	835 – 1198	≥ 1199
8	< 597	597 – 843	844 – 1199	≥ 1200
Star Math [®] cut-score equivalents				
Grade	Level 1	Level 2	Level 3	Level 4
3	< 558	558 – 606	607 – 663	≥ 664
4	< 629	629 – 691	692 – 741	≥ 742
5	< 704	704 – 764	765 – 821	≥ 822
6	< 705	705 – 775	776 – 835	≥ 836
7	< 783	783 – 838	839 – 879	≥ 880
8	< 784	784 – 850	851 – 895	≥ 896

³ 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

Four customers in New York shared student level NYSTP scores to explore the accuracy of using Star Reading and Star Math for forecasting NYSTP performance. The Star Reading sample consisted of 2,711 students and the Star Math sample consisted of 811 students. To examine the accuracy of the linkage, students' Star scores from the 2017–2018 school year were projected to the time of the NYSTP assessment using normative growth information. For students who took more than one Star Reading or Math test during the school year, projected scores were averaged.

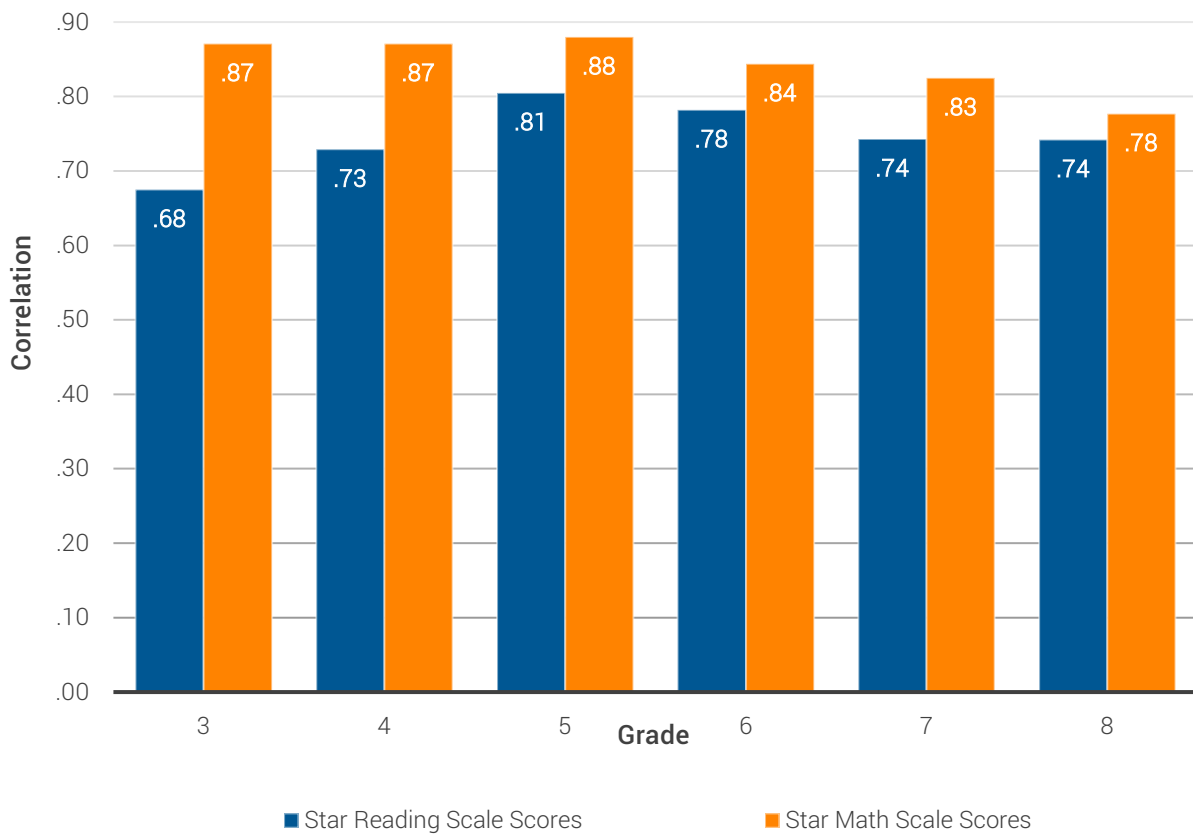
Classification diagnostics were derived from counts of correct and incorrect classifications when using projected Star scores to predict whether a student would achieve proficiency on the NYSTP. The results indicate that Star Assessments provide an effective means of estimating end-of-year achievement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

Predictive Star scores correlate highly with actual NYSTP scores

To summarize the predictive power of Star Reading and Star Math, we calculated correlations between NYSTP scores and projected Star scores. As seen in figure 1, the correlations were adequate, averaging .71 and .78 between NYSTP and Star Reading and Star Math, respectively.

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

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



Star scores discriminate well between students who score proficient or not

We compared actual NYSTP performance to students' estimated NYSTP performance based on projected 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 NYSTP using Star scores. On average, students were correctly classified (i.e., overall classification accuracy) 81% of the time by Star Reading and 89% of the time by Star Math.

For Area Under the ROC Curve (AUC), a summary measure of diagnostic accuracy, Star Reading averaged .90 and Star Math averaged .94 (also displayed in table 3). All of 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

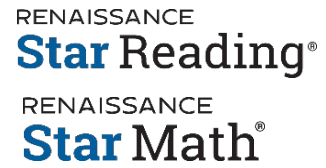
Star Reading®						
Measure	Grade					
	3	4	5	6	7	8
Overall classification accuracy (percentage of correct classifications)	82%	82%	79%	80%	85%	80%
Area Under the ROC Curve	0.88	0.92	0.91	0.90	0.92	0.89
Star Math®						
Measure	Grade					
	3	4	5	6	7	8
Overall classification accuracy (percentage of correct classifications)	88%	86%	86%	91%	96%	86%
Area Under the ROC Curve	0.95	0.92	0.93	0.98	0.99	0.88

Other diagnostic accuracy measures studied:

- ✓ **Sensitivity** represents the percentage of proficient students that were correctly forecasted, which averaged 83% for both Star Reading and Star Math.
- ✓ **Specificity** represents the percentage of not-proficient students that were correctly forecasted, which for Star Reading averaged 81% and Star Math averaged 90%.
- ✓ **Positive predictive values**, which indicate that when Star scores forecasted students to be proficient, they actually were proficient, were 62% for Star Reading and 78% for Star Math.
- ✓ **Negative predictive values**, which indicate that when Star scores forecasted students to miss proficiency, they actually weren't proficient, were 93% for reading and 91% 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 9% 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.



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: NYSTP achievement levels

Table B1. NYSTP achievement level score ranges

NYSTP achievement level score ranges: ELA				
Grade	Level 1	Level 2	Level 3	Level 4
3	530 – 582	583 – 601	602 – 628	629 – 655
4	532 – 583	584 – 602	603 – 618	619 – 654
5	509 – 593	594 – 608	609 – 621	622 – 661
6	514 – 589	590 – 601	602 – 613	614 – 657
7	511 – 590	591 – 606	607 – 622	623 – 654
8	507 – 583	584 – 602	603 – 616	617 – 651
NYSTP achievement level score ranges: Mathematics				
Grade	Level 1	Level 2	Level 3	Level 4
3	526 – 586	587 – 599	600 – 614	615 – 646
4	525 – 587	588 – 601	602 – 613	614 – 650
5	527 – 591	592 – 603	604 – 615	616 – 654
6	528 – 591	592 – 603	604 – 615	616 – 656
7	524 – 592	593 – 605	606 – 617	618 – 644
8	527 – 595	596 – 609	610 – 621	622 – 651

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>