

Pathway to Proficiency: Linking Star Reading® and Star Math® to the Wyoming Test of Proficiency and Progress (WY-TOPP)



Contents

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

Figures

| | |
|---|---|
| 6 | Figure 1. Star Reading and Star Math highly correlate with WY-TOPP scores |
|---|---|

Tables

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

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 Wyoming Test of Proficiency and Progress (WY-TOPP) tests in English Language Arts (ELA) for grades 3–8, and Mathematics grades 3–7.²

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 WY-TOPP 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 WY-TOPP 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 WY-TOPP 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. The WY-TOPP assessment is used in grades 3-10. However, adequate sample sizes were only available for the grades listed. Renaissance will hopefully expand the linking to include additional grades in future years as additional data becomes available.

School-Level Data collection

To find a sample of students who were assessed by both the WY-TOPP and Star Assessments, we began by gathering all Star Reading and Star Math test records from 2017–2018 and 2018–2019 for Wyoming. 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 WY-TOPP data from the same school year by district and school name. To do this, performance level distribution data from the WY-TOPP was obtained from the public data provided by the Wyoming Department of Education. The file included the number of students tested in each grade and the percentage of students who were *Below Basic*, *Basic*, *Proficient*, and *Advanced*.

Sample characteristics

Once we determined how many students in each grade at a school were tested on the WY-TOPP 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 WY-TOPP 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 WY-TOPP 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 ELA sample included 11,384 Star Reading students from 71 schools. The math sample included 9,489 Star Math students from 71 schools. Table 1 displays by-grade test summaries for the reading and math samples. It also includes percentages of students in each performance level, both for the sample and statewide.

Table 1. Performance characteristics of reading and math samples

| Star Reading® sample performance | | | | | | | | | | |
|----------------------------------|------------------------|-----------------------|-------------|-------|--------|-------|------------|-------|----------|-------|
| Grade | Star Reading® students | WY-TOPP ELA students | Below Basic | | Basic | | Proficient | | Advanced | |
| | | | Sample | State | Sample | State | Sample | State | Sample | State |
| 3 | 2,323 | 2,224 | 23% | 22% | 25% | 25% | 35% | 36% | 17% | 17% |
| 4 | 2,369 | 2,278 | 23% | 23% | 27% | 27% | 30% | 30% | 20% | 19% |
| 5 | 2,568 | 2,455 | 22% | 22% | 22% | 21% | 37% | 37% | 19% | 20% |
| 6 | 2,548 | 2,464 | 19% | 22% | 19% | 19% | 43% | 42% | 19% | 17% |
| 7 | 823 | 810 | 19% | 24% | 17% | 20% | 41% | 38% | 23% | 19% |
| 8 | 753 | 753 | 19% | 22% | 18% | 18% | 39% | 40% | 24% | 19% |
| Star Math® sample performance | | | | | | | | | | |
| Grade | Star Math® students | WY-TOPP Math students | Below Basic | | Basic | | Proficient | | Advanced | |
| | | | Sample | State | Sample | State | Sample | State | Sample | State |
| 3 | 2,214 | 2,121 | 26% | 23% | 26% | 25% | 28% | 29% | 20% | 23% |
| 4 | 2,050 | 1,940 | 26% | 23% | 28% | 25% | 23% | 26% | 23% | 26% |
| 5 | 2,292 | 2,186 | 26% | 23% | 24% | 23% | 32% | 34% | 18% | 20% |
| 6 | 2,364 | 2,275 | 20% | 23% | 22% | 23% | 33% | 30% | 25% | 23% |
| 7 | 569 | 561 | 22% | 25% | 25% | 24% | 27% | 27% | 26% | 24% |

Results

Scale linkage

Renaissance linked the Star test scale to the WY-TOPP by applying equipercentile linking analysis (Kolen & Brennan, 2004). First, we aggregated the sample of schools to calculate the percentage of students performing at the *Below Basic*, *Basic*, *Proficient*, and *Advanced* performance levels 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 WY-TOPP level. For example, as shown in Table 1, 23% of students in our third-grade reading sample were classified as *Below Basic*, 25% *Basic*, 35% *Proficient*, and 17% *Advanced*. Therefore, the cutscores for proficiency levels in the third grade are at the 23rd percentile for *Basic*, the 48th percentile for *Proficient*, and the 83rd percentile for *Advanced*.

WY-TOPP cut scores and corresponding Star score equivalents

WY-TOPP results are reported in scaled scores that are split into four achievement levels: *Below Basic*, *Basic*, *Proficient*, and *Advanced*. The main purpose in linking Star Reading and Star Math to the WY-TOPP was to identify Star scores at the time of the state test that are approximately equivalent to the cut-off scores that separate the WY-TOPP levels. Table 2 displays these equivalent Star scores at the time of the state test for grades 3–8 for reading and grades 3–7 for math.³ The corresponding WY-TOPP cut scores can be found in Appendix B.

Table 2. Star Reading® and Star Math® score equivalents for each WY-TOPP achievement level range

| Star Reading® cut-score equivalents | | | | |
|-------------------------------------|-------------|-----------|------------|----------|
| Grade | Below Basic | Basic | Proficient | Advanced |
| 3 | < 356 | 356 - 461 | 462 - 621 | ≥ 622 |
| 4 | < 453 | 453 - 564 | 565 - 741 | ≥ 742 |
| 5 | < 494 | 494 - 600 | 601 - 872 | ≥ 873 |
| 6 | < 533 | 533 - 652 | 653 - 971 | ≥ 972 |
| 7 | < 606 | 606 - 738 | 739 - 1028 | ≥ 1029 |
| 8 | < 633 | 633 - 793 | 794 - 1179 | ≥ 1180 |
| Star Math® cut-score equivalents | | | | |
| Grade | Below Basic | Basic | Proficient | Advanced |
| 3 | < 591 | 591 - 642 | 643 - 690 | ≥ 691 |
| 4 | < 644 | 644 - 706 | 707 - 756 | ≥ 757 |
| 5 | < 693 | 693 - 759 | 760 - 827 | ≥ 828 |
| 6 | < 714 | 714 - 774 | 775 - 841 | ≥ 842 |
| 7 | < 750 | 750 - 815 | 816 - 869 | ≥ 870 |

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

Two Wyoming districts shared student level WY-TOPP scores to explore the accuracy of using Star Reading and Star Math for forecasting WY-TOPP performance. The Star Reading sample consisted of 1,723 students and the Star Math sample consisted of 1,442 students. We took students' Star scores and used national growth norms (Renaissance Learning, 2019a, 2019b) to project what their Star scores would be at the date of the WY-TOPP administration. For each student, we used the average of their projected Star scores to examine the accuracy of the linkage to the WY-TOPP scale.

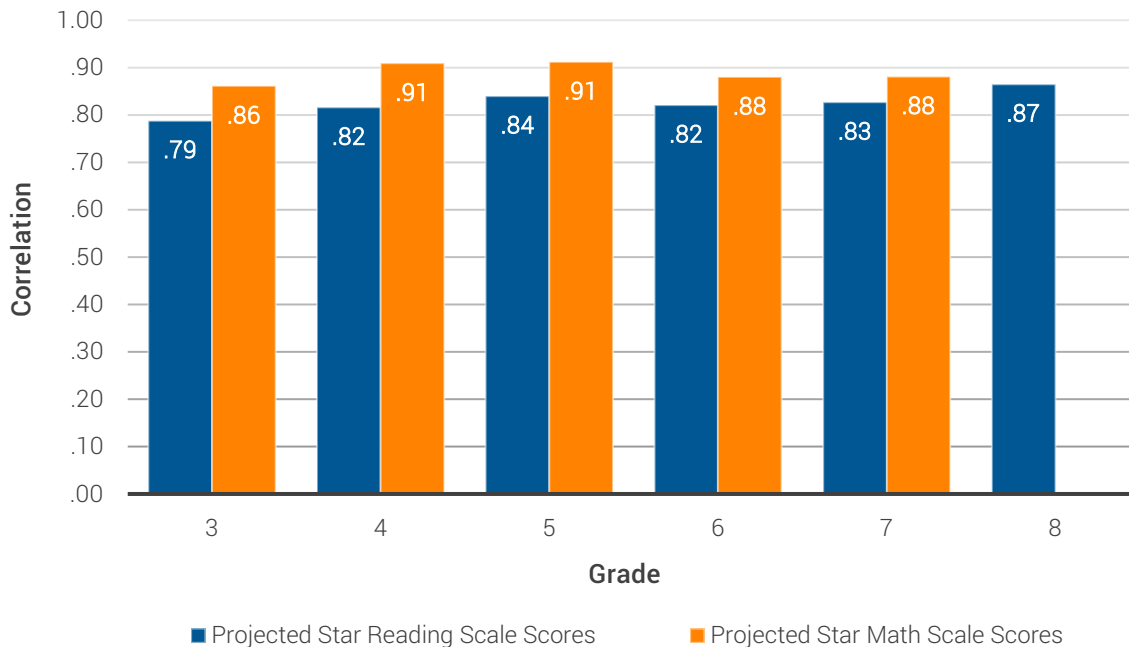
Classification diagnostics were derived from counts of correct and incorrect classifications when using Star scores to predict whether a student would achieve proficiency on the WY-TOPP. The results indicate that Star Assessments provide an effective means of estimating end-of-year achievement on the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

Projected Star scores correlate highly with actual WY-TOPP scores

To summarize the predictive power of Star Reading and Star Math, we calculated correlations between observed WY-TOPP scores and projected Star scores. As seen in figure 1, the correlations were strong, averaging .83 and .89 between WY-TOPP and Star Reading and Star Math, respectively.

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

Figure 1. Star Reading® and Star Math® scores highly correlate with WY-TOPP scores



Star scores discriminate well between students who score proficient or not

We compared actual WY-TOPP performance to students' estimated WY-TOPP 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 WY-TOPP using projected Star scores. On average, students were correctly classified (i.e., overall classification accuracy) 85% of the time by Star Reading and 87% of the time by Star Math.

For Area Under the ROC Curve (AUC), a summary measure of diagnostic accuracy, Star Reading averaged .94 and Star Math averaged .95 (also displayed in table 3). AUC values closer to 1 indicate an assessment perfectly distinguishes between students who are proficient versus those who are not, whereas values of .50 indicate prediction no better than chance. In general, an AUC of .70 to .80 is considered acceptable, .80 to .90 is excellent, and greater than .90 is outstanding (Hosmer et al., 2013).

Table 3. Proficiency forecasting using Star Reading® and Star Math® scores yields accurate results

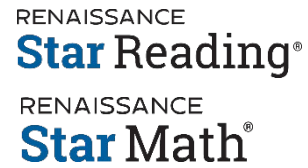
| Star Reading® | | | | | | |
|---|-------|------|------|------|------|------|
| Measure | Grade | | | | | |
| | 3 | 4 | 5 | 6 | 7 | 8 |
| Overall classification accuracy (percentage of correct classifications) | 85% | 87% | 89% | 88% | 83% | 88% |
| Area Under the ROC Curve | 0.92 | 0.94 | 0.96 | 0.95 | 0.92 | 0.95 |
| Star Math® | | | | | | |
| Measure | Grade | | | | | |
| | 3 | 4 | 5 | 6 | 7 | |
| Overall classification accuracy (percentage of correct classifications) | 86% | 88% | 88% | 85% | 88% | |
| Area Under the ROC Curve | 0.92 | 0.96 | 0.96 | 0.93 | 0.97 | |

Other diagnostic accuracy measures studied:

- ✓ **Sensitivity** represents the percentage of proficient students that were correctly forecasted, which for Star Reading averaged 92% and for Star Math also averaged 89%.
- ✓ **Specificity** represents the percentage of not-proficient students that were correctly forecasted, which for Star Reading averaged 77% and for Star Math averaged 84%.
- ✓ **Positive predictive values**, which indicate that when Star scores forecasted students to be proficient, they actually were proficient, were 87% for both Star Reading and Star Math.
- ✓ **Negative predictive values**, which indicate that when Star scores forecasted students to miss proficiency, they actually weren't proficient, were 86% for reading and 87% 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 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: WY-TOPP achievement levels

Table B1. WY-TOPP achievement level score ranges

| WY-TOPP achievement level score ranges: ELA | | | | |
|---|-------------|-----------|------------|----------|
| Grade | Below Basic | Basic | Proficient | Advanced |
| 3 | < 563 | 564 – 591 | 592 – 627 | ≥ 628 |
| 4 | < 579 | 580 – 612 | 613 – 648 | ≥ 649 |
| 5 | < 600 | 601 – 626 | 627 – 667 | ≥ 668 |
| 6 | < 613 | 614 – 639 | 640 – 688 | ≥ 689 |
| 7 | < 616 | 617 – 643 | 644 – 691 | ≥ 692 |
| 8 | < 625 | 626 – 652 | 653 – 707 | ≥ 708 |
| 9 | < 628 | 629 – 655 | 656 – 708 | ≥ 709 |
| 10 | < 630 | 631 – 666 | 667 – 712 | ≥ 713 |
| WY-TOPP achievement level score ranges: Mathematics | | | | |
| Grade | Below Basic | Basic | Proficient | Advanced |
| 3 | < 416 | 417 – 437 | 438 – 460 | ≥ 461 |
| 4 | < 441 | 442 – 465 | 466 – 490 | ≥ 491 |
| 5 | < 465 | 466 – 493 | 494 – 532 | ≥ 533 |
| 6 | < 489 | 490 – 520 | 521 – 559 | ≥ 560 |
| 7 | < 515 | 516 – 551 | 552 – 587 | ≥ 588 |
| 8 | < 542 | 543 – 584 | 585 – 626 | ≥ 627 |
| 9 | < 572 | 573 – 616 | 617 – 667 | ≥ 668 |
| 10 | < 602 | 603 – 652 | 653 – 713 | ≥ 714 |

References

- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Assessing the fit of the model. In *Applied logistic regression* (3rd ed., pp. 153–225). John Wiley & Sons.
- 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>