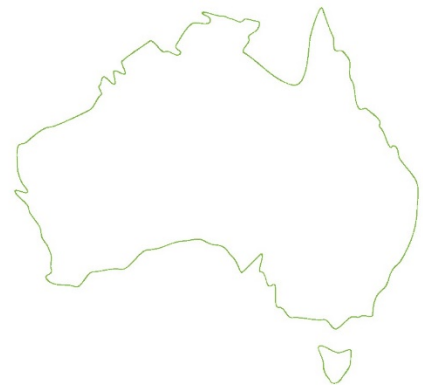


RENAISSANCE

TECHNICAL PAPER | 2021 School Year

Pathway to Proficiency: Linking Star Reading to the National Assessment Program – Literacy and Numeracy (NAPLAN) Reading Test



Contents

3	Introduction
3	Main Findings
3	Study
4	Results
8	Appendix A: About Star Reading
8	Appendix B: NAPLAN achievement levels
9	References

Figures

6	Figure 1. Star Reading highly correlates with NAPLAN scores
---	---

Tables

4	Table 1. Performance characteristics of reading sample
5	Table 2. Star Reading score equivalents for each NAPLAN achievement level range
6	Table 3. Using Star Reading scores to classify as at/above the national minimum standard for NAPLAN Reading yields accurate results
7	Table 4. Using Star Reading scores to classify as within or below the top two bands for NAPLAN Reading yields accurate results
8	Table B1. NAPLAN achievement level score ranges

Initial publication March 11, 2021

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 an interim assessment, Renaissance Star Reading®, to help determine whether individual students were on track or needed more assistance to succeed in the National Assessment Program – Literacy and Numeracy (NAPLAN) test in Reading for Years 3, 5, 7, and 9.

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

Study

To determine if Star Reading can predict student achievement on the NAPLAN reading test, 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 NAPLAN and Star Assessments, we began by gathering all Star Reading test records from 2016–2017, 2017–2018, and 2018–2019 for Australia. Then, each school's Star Reading data were aggregated by year and subject area. The next step was to match Star data with the NAPLAN data from the same school year by school name. To do this, performance level distribution data from the NAPLAN was obtained from a request submitted to the Australian Curriculum, Assessment and Reporting Authority (ACARA)'s Data Access Program. The resulting file included the number of students tested in each year and the percentage of students who scored at each band on the NAPLAN.

Sample characteristics

Once we determined how many students in each year at a school were tested in the NAPLAN area of Reading and took a Star Reading assessment, we calculated the percentage of students assessed on both tests. In each year at each school, if between 95% and 105% of the students who tested on the NAPLAN Reading test had taken a Star assessment, that year level 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 NAPLAN Reading test also took

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.

a Star Reading test within the specified window of time. Additionally, we examined the proximity of Star Reading testing in relation to NAPLAN Reading testing and included schools in which a majority of students (at least 90% for Years 3, 5, and 7 and at least 75% for Year 9²) had taken Star Reading tests within 30 days of the NAPLAN Reading test. This aspect of the sample selection ensured that a majority of the sample consisted of students with concurrent Star Reading and NAPLAN reading tests, while also ensuring a sufficient sample size for analysis.

The sample included 7,595 Star Reading students from 82 unique schools. Table 1 displays by-year test summaries for the reading sample. It also includes percentages of students in each performance level, both for the concurrent sample and nationwide for schools that use Star Reading.

Table 1. Performance characteristics of reading sample

Star Reading sample performance										
Year	Star Reading students	NAPLAN Reading students	Below the National Minimum Standard		At the National Minimum Standard		Above the National Minimum Standard		Performance in Top Two NAPLAN Bands	
			Sample	National Star Users	Sample	National Star Users	Sample	National Star Users	Sample	National Star Users
3	1,908	1,869	3%	4%	5%	5%	92%	91%	68%	60%
5	3,827	3,745	2%	4%	6%	9%	92%	88%	57%	47%
7	1,295	1,267	3%	5%	7%	12%	90%	83%	31%	26%
9	565	563	5%	8%	13%	16%	82%	75%	23%	19%

Results

Scale linkage

Renaissance linked the Star test scale to the NAPLAN by applying equipercentile linking analysis (Kolen & Brennan, 2004). First, we aggregated the sample of schools to calculate the percentage of students performing *Below the National Minimum Standard*, *At the National Minimum Standard*, and *Above the National Minimum Standard* performance levels for each subject and year level. As some performance targets focus on high levels of achievement, such as performance in the top two NAPLAN bands, we additionally aggregated the sample of schools to calculate the percentage of students in the top NAPLAN bands for reading. Then we analysed the distribution of Star scores to determine the scaled score corresponding to the same percentile as each specific NAPLAN level. For example, as shown in table 1, 2% of students in our Year 3 reading sample were classified as *Below the National Minimum Standard*, 4% *At the National Minimum Standard*, and 94% *Above the National Minimum Standard*. Therefore, the cutscores for Year 3 performance levels are at the 2nd percentile for *At the National Minimum Standard* and the 6th percentile for *Above the National Minimum Standard*. For Year 3, 32% of students performed in the first four NAPLAN bands and 68% of students performed in the top two NAPLAN bands. As a result, the cutscore for Year 3 performance in the top two NAPLAN bands is at the 32nd percentile.

² The percentage of Star records required to be in close proximity to the NAPLAN was relaxed for Year 9 to allow for a larger sample size sufficient for analysis.

NAPLAN cut scores and corresponding Star score equivalents

NAPLAN results are reported in scaled scores that are split into 10 bands associated with three achievement levels: *Below the National Minimum Standard*, *At the National Minimum Standard*, and *Above the National Minimum Standard*. Additionally, some schools track performance in the top two NAPLAN bands. The main purpose in linking Star Reading to the NAPLAN reading test was to identify Star scores at the time of the NAPLAN reading test that are approximately equivalent to the cut-off scores that separate the NAPLAN performance levels. Table 2 displays these equivalent Star scores at the time of the NAPLAN test for Years 3, 5, 7, and 9.³ The corresponding NAPLAN bands associated with each achievement level can be found in Appendix B.

Table 2. Star Reading score equivalents for NAPLAN reading achievement levels

Star Reading cut-score equivalents				
Year	Below the National Minimum Standard	At the National Minimum Standard	Above the National Minimum Standard	Top Two Bands
3	< 85	85 – 137	≥ 138	≥ 354
5	< 203	203 – 346	≥ 347	≥ 590
7	< 313	313 – 443	≥ 444	≥ 903
9	< 361	361 – 587	≥ 588	≥ 1179

Accuracy of scale linkage confirmed

Six Australian schools shared student level NAPLAN scores from 2017 and 2019 to explore the accuracy of using Star Reading for determining NAPLAN performance. The Star Reading sample consisted of 803 Star Reading records from students in Years 3, 5, and 7.⁴ who took Star Reading within 30 days of the NAPLAN Reading test.

Classification diagnostics were derived from counts of correct and incorrect classifications when using Star Reading scores to determine whether a student would 1) meet or exceed the national minimum standard on the NAPLAN Reading test and 2) receive a score falling within or below the top two NAPLAN bands on the NAPLAN Reading test. The results indicate that Star Reading Assessments provide an effective means of estimating achievement on the NAPLAN Reading test.

Concurrent Star Reading scores correlate highly with actual NAPLAN Reading scores

To summarise the relationship between Star Reading and NAPLAN Reading, we calculated correlations between observed NAPLAN Reading scale scores and concurrent Star Reading scale scores.⁵ As seen in figure 1, the correlations were strong and averaged .78 across the year levels.

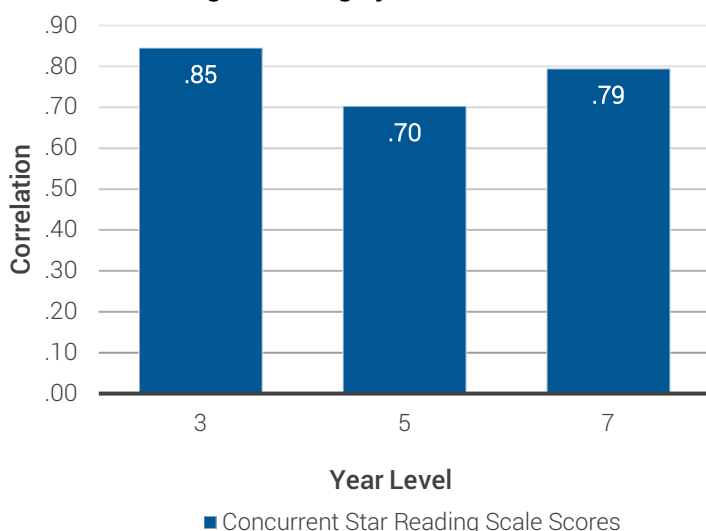
Star Reading scores have a strong relationship with NAPLAN Reading scores.

³ The Star Reading cut-score equivalents presented in table 2 apply only to the time of the NAPLAN test.

⁴ The sample size for Year 9 was not sufficient for analysis. Year 9 results will be added as more data becomes available.

⁵ A small portion of the sample contained only NAPLAN Reading band information and NAPLAN scale scores were not provided. As a result, the Star Reading and NAPLAN Reading correlation analysis included a slightly smaller sample size (782).

Figure 1. Star Reading scores highly correlate with NAPLAN Reading scores



Star scores discriminate well between students who score at/above or below the national minimum standard

We compared actual NAPLAN Reading performance to estimated NAPLAN Reading 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 at or above the national minimum standard or not on the NAPLAN Reading test using concurrent Star scores and the Star Reading cut score equivalents for *At the National Minimum Standard* indicated in table 2. On average, students were correctly classified (i.e., overall classification accuracy) as at/above the national minimum standard or not 98% of the time. For Area Under the ROC Curve (AUC), a summary measure of diagnostic accuracy, classification of scores as at/above the national minimum standard averaged .94. 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. Using Star Reading scores to classify as at/above the national minimum standard for NAPLAN Reading yields accurate results

At/Above National Minimum Standard			
Measure	Year		
	3	5	7
Overall classification accuracy (percentage of correct classifications)	97%	98%	97%
Observed percent at/above national minimum standard	97%	98%	98%
Projected percent at/above national minimum standard	98%	99%	98%
Area Under the ROC Curve	0.94	0.91	0.96

Star scores discriminate well between students who score within or below the top two NAPLAN bands

As some schools utilise performance targets focused on high levels of achievement, such as performance in the top two NAPLAN bands, we additionally analysed whether concurrent Star Reading scores could be used to differentiate performance within or below the top two bands for NAPLAN Reading. Table 4 displays classification diagnostics about whether students were correctly or incorrectly classified as within or below the top two bands for NAPLAN Reading using concurrent Star Reading scores and the Star Reading cut score equivalents for *Top Two Bands* indicated in table 2. On average, students were correctly classified as within or below the top two NAPLAN bands 84% of the time. For Area Under the ROC Curve, classification of scores as within or below the top two NAPLAN bands averaged .90 (also displayed in table 4).

Table 4. Using Star Reading scores to classify as within or below the top two bands for NAPLAN Reading yields accurate results

Performance Within Top Two NAPLAN Bands			
Measure	Year		
	3	5	7
Overall classification accuracy (percentage of correct classifications)	88%	80%	82%
Observed percent in top two bands	68%	57%	37%
Projected percent in top two bands	72%	53%	37%
Area Under the ROC Curve	0.93	0.86	0.91

Other diagnostic accuracy measures studied:

- ✓ **Sensitivity** represents the percentage of students in the top two bands that were correctly forecasted, which averaged 83%.
- ✓ **Specificity** represents the percentage of students not in the top two bands that were correctly forecasted, which averaged 81%.
- ✓ **Positive predictive values**, which indicate that when Star scores forecasted students to be in the top two bands, they actually were in the top two bands, averaged 83%.
- ✓ **Negative predictive values**, which indicate that when Star scores forecasted students to not be in the top two bands, they actually weren't in the top two bands, averaged 82%.
- ✓ **Projection error**, the difference between actual and projected percentages in the top two bands, indicates how well scores accurately predict status in the top two bands within each year level. The projection error for the top two bands averaged 0% (negative scores indicate under-prediction while positive scores show over-prediction).

Appendix A: About Star Reading

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

RENAISSANCE
Star Reading

Star Reading is 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: NAPLAN achievement levels

Table B1. NAPLAN achievement level score ranges

NAPLAN achievement level band ranges: Reading				
Year	Band Below the National Minimum Standard	Band At the National Minimum Standard	Bands Above the National Minimum Standard	Top Two Bands
3	1	2	3 and above	5 – 6
5	3	4	5 and above	7 – 8
7	4	5	6 and above	8 – 9
9	5	6	7 and above	9 – 10

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. (2019). *Star Reading technical manual*. Available by request to research@renaissance.com

Independent technical reviews of Star Reading

- 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. (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. (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. (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>