Predicting Proficiency on the State of Texas Assessments of Academic Readiness (STAAR) Spanish Reading based on NWEA MAP Growth Scores

February 2022

NWEA Psychometric Solutions

nwea

Linking Study Updates

Date	Description
2022-02-07	Initial study conducted for STAAR Spanish in grades 3–5 reading using Spring 2019 data.

Acknowledgements: This report benefited from the project management of Debbie Bauman, the statistical analysis of Sarah-Truclinh Tran and Rudraxi Majumdar, the editorial assistance of Kelly Rivard, and the psychometric leadership of Ann Hu. A special thank you to Shudong Wang for his expertise and support regarding the Spanish user norms. We appreciate all our colleagues at NWEA who assisted in partner recruitment and are grateful to all our partners who provided data for the study.

Table of Contents

Executive Summary 4
1. Introduction 7
    1.1. Purpose of the Study 7
    1.2. Assessment Overview 7
2. Methods 8
    2.1. Data Collection 8
    2.2. Post-Stratification Weighting 8
    2.3. MAP Growth Cut Scores 8
    2.4. Classification Accuracy 9
    2.5. Proficiency Projections 10
3. Results 11
    3.1. Study Sample 11
    3.2. Descriptive Statistics 11
    3.3. MAP Growth Cut Scores 12
    3.4. Classification Accuracy 13
    3.5. Proficiency Projections 13
4. References 16

List of Tables

Table 2.1. Description of Classification Accuracy Summary Statistics 9
Table 3.1. Linking Study Sample Demographics 11
Table 3.2. Descriptive Statistics of Test Scores 12
Table 3.3. MAP Growth Cut Scores 12
Table 3.4. Classification Accuracy Results 13
Table 3.5. Proficiency Projections based on RIT Scores 14

Executive Summary

To predict student achievement on the State of Texas Assessments of Academic Readiness (STAAR) Spanish assessment in grades 3–5 reading, NWEA® conducted a linking study using Spring 2019 data to derive Rasch Unit (RIT) cut scores on the Spanish MAP® Growth™ Reading assessments that correspond to the STAAR Spanish performance levels. Educators can use this information to identify students at risk of not meeting state proficiency standards early in the year and provide tailored educational interventions.[1]

E.1. Proficiency Cut Scores

Table E.1 presents the STAAR *Meets Grade Level* performance level cut scores and the corresponding MAP Growth RIT cut scores that allow teachers to identify students who are on track for proficiency on the state summative test and those who are not. For example, the *Meets Grade Level* cut score on the third-grade STAAR Spanish Reading test is 1444. A third-grader with a Spanish MAP Growth Reading RIT score of 193 in the fall is likely to meet proficiency on the STAAR Spanish Reading test in the spring, whereas a third-grader with a RIT score lower than 193 in the fall is in jeopardy of not meeting proficiency.

Table E.1. MAP Growth Cut Scores for STAAR Spanish Reading Proficiency

Assessment		*Meets Grade Level* Cut Scores by Grade		
		3	4	5
STAAR Spanish Reading, Spring		1444	1539	1582
Spanish MAP Growth Reading	Fall	193	203	203
	Winter	198	207	207
	Spring	199	208	209

E.2. Assessment Overview

The STAAR program includes Texas' state summative assessments aligned to the Texas Essential Knowledge and Skills (TEKS) curriculum. The STAAR Spanish assessment is aligned to the TEKS in Spanish and is administered to eligible students for whom a Spanish version of STAAR is the most appropriate measure of their academic progress. Based on their test scores, students are placed into one of four performance levels: *Did Not Meet Grade Level*, *Approaches Grade Level*, *Meets Grade Level*, and *Masters Grade Level*. The *Meets Grade Level* cut score demarks the minimum level of achievement considered to be proficient for accountability purposes.

The Spanish MAP Growth Reading tests were developed to help partners better understand what Spanish-speaking students know and are ready to learn. Like their English language counterparts, they are adaptive interim assessments aligned to state-specific content standards and administered in the fall, winter, and spring. Scores are reported on the RIT vertical scale with a range of 100 to 350.

[1] This study provides MAP Growth cut scores that predict proficiency on the STAAR Spanish Reading assessment in grades 3–5 only. They represent a higher level of achievement than universal screening cut scores designed to identify students with the most severe learning difficulties who may need intensive intervention. The Spanish MAP Growth Reading universal screening cut scores for grades K–8 are available in a separate report (He & Meyer, 2021).

The 2020 Spanish reading user norms used in this study provide basic contextual information about student performance in the fall, winter, and spring and growth between terms on the Spanish MAP Growth Reading assessments (NWEA, 2021). Given that the user norms were drawn from a limited pool of test events and are not nationally representative like the general MAP Growth norms, caution needs to be taken when the Spanish norms are used by partners who did not participate in the 2020 Spanish norms study.

E.3. Linking Methods

The equipercentile linking method was used to identify the spring MAP Growth scores that correspond to the spring STAAR Spanish performance level cut scores. MAP Growth fall and winter cut scores that predict proficiency on the spring STAAR Spanish test were then projected using the 2020 Spanish MAP Growth Reading growth norms that provide expected score gains across test administrations.

E.4. Student Sample

Table E.2 presents the number of Texas students from six districts and 70 schools who were included in the linking study sample. Only students who took both the Spanish MAP Growth Reading and STAAR Spanish Reading assessments in Spring 2019 were included. Despite having a sample size being below the usual threshold of 1,000 students, fifth grade was included in the study sample because of the robust correlation and classification accuracy statistics observed between the two assessments for this grade cohort.

Table E.2. Linking Study Sample

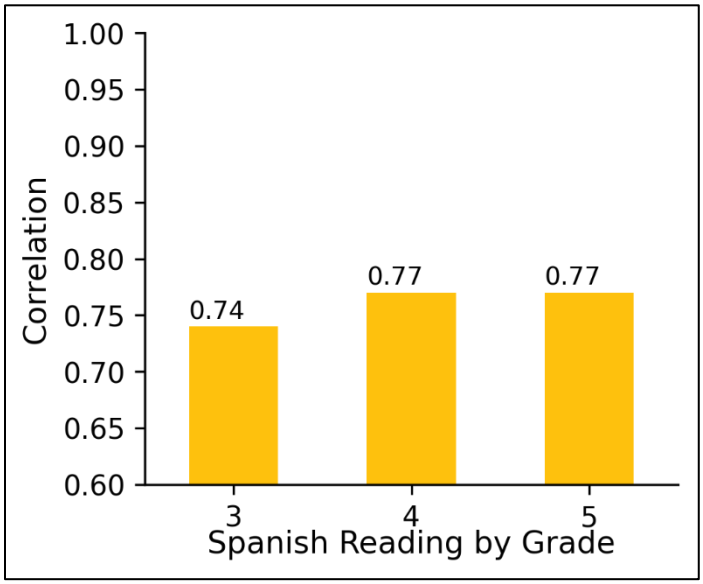
Grade	#Students
3	1,148
4	1,015
5	763

The linking study sample is voluntary and can only include student scores from partners who share their data. Also, not all students in a state take MAP Growth. The sample may therefore be different from the general student population in important characteristics. To ensure that the linking study sample represents the state student population in terms of race, sex, and performance level distributions, post-stratification weighting was applied to statistically adjust the sample so it reflects the target population on these variables. As a result, the RIT cuts derived from the study sample can be generalized to any student from the target population. All analyses in this study were conducted based on the weighted sample.

E.5. Test Score Relationships

Correlations between MAP Growth RIT scores and STAAR Spanish Reading scores range from 0.74 to 0.77, as shown in Figure E.1. These values indicate a high positive correlation among the scores, which is important validity evidence for the claim that MAP Growth scores are good predictors of performance on the STAAR Spanish Reading assessments.

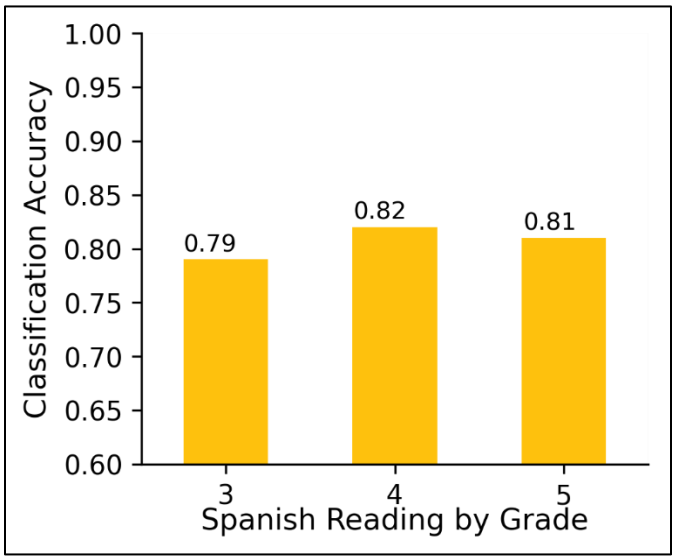
Figure E.1. Correlations between Spanish MAP Growth Reading and STAAR Spanish Reading Test Scores



E.6. Accuracy of MAP Growth Classifications

Figure E.2 presents the classification accuracy statistics that show the proportion of students correctly classified by their RIT scores as proficient or not proficient on the STAAR Spanish Reading tests. For example, the third-grade Spanish MAP Growth Reading *Meets Grade Level* cut score has a 0.79 accuracy rate, meaning it accurately classified student achievement on the state test for 79% of the sample. The results range from 0.79 to 0.82, indicating that RIT scores are good at identifying student proficiency on the STAAR Spanish Reading tests.

Figure E.2. Accuracy of MAP Growth Classifications



1. Introduction

1.1. Purpose of the Study

NWEA® is committed to providing partners with useful tools to help make inferences about student learning from MAP® Growth™ test scores. One important use of MAP Growth results is to predict a student's performance on the state summative assessment at different times throughout the year. This allows educators and parents to determine if a student is on track in their learning to meet state standards by the end of the year or, given a student's learning profile, is on track to obtain rigorous, realistic growth in their content knowledge and skills.

This document presents results from a linking study conducted by NWEA to statistically connect the scores of the State of Texas Assessments of Academic Readiness (STAAR) Spanish assessment in grades 3–5 reading with Rasch Unit (RIT) scores from the Spanish MAP Growth Reading assessments taken during the Spring 2019 term. Specifically, this report presents the following results:

1. Student sample demographics
2. Descriptive statistics of test scores
3. MAP Growth cut scores from fall, winter, and spring that correspond to the performance levels on the spring STAAR Spanish Reading assessment
4. Classification accuracy statistics to determine the degree to which MAP Growth accurately predicts student proficiency status on the STAAR Spanish Reading tests
5. The probability of achieving grade-level proficiency on the STAAR Spanish Reading assessment based on MAP Growth RIT scores from fall, winter, and spring

1.2. Assessment Overview

The STAAR Spanish assessment is aligned to the Texas Essential Knowledge and Skills (TEKS) curriculum in Spanish and is administered to eligible students for whom a Spanish version of STAAR is the most appropriate measure of their academic progress. A cut score is the minimum score a student must get on a test to be placed in a certain performance level. The STAAR Spanish Reading assessment has three cut scores that distinguish between the following performance levels: *Did Not Meet Grade Level*, *Approaches Grade Level*, *Meets Grade Level*, and *Masters Grade Level*. The *Meets Grade Level* cut score demarks the minimum level of performance considered to be proficient for accountability purposes.

The Spanish MAP Growth Reading assessments were developed to help partners better understand what Spanish-speaking students know and are ready to learn. Scores are reported on the RIT vertical scale with a range of 100 to 350. To aid the interpretation of scores, NWEA conducts norming studies of student and school performance on MAP Growth. Achievement status norms show how well a student performed on MAP Growth compared to students in the norming group. It does this by associating the student's RIT score with a percentile ranking. Growth norms provide expected score gains across test administrations (e.g., the relative evaluation of a student's growth from fall to spring).

This study uses the 2020 Spanish reading user norms that are on the Spanish reading scale. These user norms are drawn from a limited pool of test events and are not nationally representative like the general MAP Growth norms. Nevertheless, they provide useful contextual information about student performance in the fall, winter, and spring and growth between two terms on the Spanish MAP Growth Reading assessments (NWEA, 2021).

2. Methods

2.1. Data Collection

This linking study is based on data from the Spring 2019 administrations of the Spanish MAP Growth Reading and STAAR Spanish Reading assessments. NWEA recruited Texas districts to participate in the study by sharing their student and score data for the target term. Districts also gave NWEA permission to use their students' MAP Growth scores from the NWEA in-house database. Once state score information was received by NWEA, each student's state testing record was matched to their MAP Growth score based on the student's first and last names, date of birth, student ID, and other available identifying information. Only students who took both the MAP Growth and STAAR Spanish Reading assessments in Spring 2019 were included in the study sample.

2.2. Post-Stratification Weighting

Post-stratification weights were applied to the calculations to ensure that the linking study sample represented the state's test-taking student population in terms of race, sex, and performance level. These variables were selected because they are known to be correlated with students' academic achievement and are often available in state summative assessment reports. The weighted sample will match the target population as closely as possible on the key demographics and performance characteristics as defined by the state.

A raking procedure was used to calculate the post-stratification weights that either compensate for the underrepresentation of certain groups or attenuate the overrepresentation of certain groups. Raking uses iterative procedures to obtain weights that match sample marginal distributions to known population margins. The following steps were taken during this process:

1. Calculate marginal distributions of race, sex, and performance level for the sample and population.
2. Calculate post-stratification weights with the rake function from the survey package in R (Lumley, 2019).
3. Trim the weights that are outside the range of 0.3 to 3.0.
4. Apply the weights to the sample before conducting the linking study analyses.

2.3. MAP Growth Cut Scores

Spanish MAP Growth Reading cut scores that predict student achievement on the STAAR Spanish Reading assessment are reported for grades 3–5. Percentile ranks based on the 2020 Spanish user norms are also provided. These are useful for understanding how students' scores compare to peers nationwide and the relative rigor of a state's performance level designations for its summative assessment.

The equipercentile linking method (Kolen & Brennan, 2004) was used to identify the spring Spanish MAP Growth Reading RIT scores that correspond to the spring STAAR Spanish Reading performance level cut scores. The equipercentile linking procedure matches scores on the two scales that have the same percentile rank (i.e., the proportion of tests at or below each score). For example, let x represent a score on Test X (e.g., STAAR Spanish). Its equipercentile equivalent score on Test Y (e.g., MAP Growth), $e_y(x)$, can be obtained through a cumulative-distribution-based linking function defined in Equation 1:

$$e_y(x) = G^{-1}[P(x)] \quad (1)$$

where $e_y(x)$ is the equipercentile equivalent of score x on STAAR Spanish on the scale of MAP Growth, $P(x)$ is the percentile rank of a given score on STAAR Spanish, and G^{-1} is the inverse of the percentile rank function for MAP Growth that indicates the score on MAP Growth corresponding to a given percentile. Polynomial loglinear pre-smoothing was applied to reduce irregularities of the score distributions and equipercentile linking curve.

The Spanish MAP Growth Reading conditional growth norms provide students' expected score gains across terms within the same school year, such as growth from fall to spring within the same grade. This information was used to calculate the fall and winter cut scores. Equation 2 was used to determine the previous term's MAP Growth score needed to reach the spring cut score, considering the expected growth associated with the previous RIT score:

$$RIT_{PredSpring} = RIT_{previous} + g \quad (2)$$

where:

- $RIT_{PredSpring}$ is the predicted MAP Growth spring score.
- $RIT_{previous}$ is the previous term's RIT score.
- g is the expected growth from the previous RIT (e.g., fall or winter) to the spring RIT score.

2.4. Classification Accuracy

The degree to which MAP Growth predicts student proficiency status on the STAAR Spanish Reading tests can be described using classification accuracy statistics based on the MAP Growth spring RIT cut scores. The results show the proportion of students correctly classified by their RIT scores as proficient or not proficient on the STAAR Spanish Reading test. Table 2.1 describes the classification accuracy statistics provided in this report (Pommerich et al., 2004).

Table 2.1. Description of Classification Accuracy Summary Statistics

Statistic	Description*	Interpretation
Overall Classification Accuracy Rate	(TP + TN) / (total sample size)	Proportion of the study sample whose proficiency classification on the state test was correctly predicted by MAP Growth cut scores
False Negative (FN) Rate	FN / (FN + TP)	Proportion of not-proficient students identified by MAP Growth in those observed as proficient on the state test
False Positive (FP) Rate	FP / (FP + TN)	Proportion of proficient students identified by MAP Growth in those observed as not proficient on the state test
Sensitivity	TP / (TP + FN)	Proportion of proficient students identified by MAP Growth in those observed as such on the state test
Specificity	TN / (TN + FP)	Proportion of not-proficient students identified by MAP Growth in those observed as such on the state test
Precision	TP / (TP + FP)	Proportion of observed proficient students on the state test in those identified as such by the MAP Growth test

Statistic	Description*	Interpretation
Area Under the Curve (AUC)	Area under the receiver operating characteristics (ROC) curve	How well MAP Growth cut scores separate the study sample into proficiency categories that match those from the state test cut scores. An AUC at or above 0.80 is considered "good" accuracy.

*FP = false positives. FN = false negatives. TP = true positives. TN = true negatives.

2.5. Proficiency Projections

In addition to calculating the MAP Growth fall and winter cut scores, the MAP Growth conditional growth norms data were also used to calculate the probability of reaching proficiency on the STAAR Spanish Reading test based on a student's RIT scores from fall, winter, and spring. Equation 3 was used to calculate the probability of a student achieving *Meets Grade Level* performance on the STAAR Spanish Reading test based on their fall or winter RIT score:

$$Pr(Achieving\ Meets\ Grade\ Level\ in\ spring|\ starting\ RIT) = \Phi\left(\frac{RIT_{previous} + g - RIT_{SpringCut}}{SD}\right) \quad (3)$$

where:

- Φ is the standard normal cumulative distribution function.
- $RIT_{previous}$ is the student's RIT score in fall or winter.
- g is the expected growth from the previous RIT (e.g., fall or winter) to the spring RIT.
- $RIT_{SpringCut}$ is the MAP Growth *Meets Grade Level* cut score for spring.
- SD is the conditional standard deviation of the expected growth, g.

Equation 4 was used to estimate the probability of a student achieving *Meets Grade Level* performance on the STAAR Spanish Reading test based on their spring RIT score (RIT_{Spring}):

$$Pr(Achieving\ Meets\ Grade\ Level\ in\ spring\ |\ spring\ RIT) = \Phi\left(\frac{RIT_{Spring} - RIT_{SpringCut}}{SE}\right) \quad (4)$$

where SE is the standard error of measurement for MAP Growth.

3. Results

3.1. Study Sample

Only students who took both the Spanish MAP Growth Reading and STAAR Spanish Reading assessments in Spring 2019 were included in the study sample. Data used in this study were collected from six districts and 70 schools in Texas. Table 3.1 presents the demographic distributions of race, sex, and performance level in both the original unweighted and weighted study sample, as well as the distributions of the target population of students who took the Spring 2019 STAAR Spanish Reading tests. Since the original study sample is different from the target STAAR Spanish Reading population, post-stratification weights were applied to the linking study sample to improve its representativeness. The demographic distributions of the sample after weighting are almost identical to the STAAR Spanish Reading student population distributions. The analyses in this study were therefore conducted using the weighted sample.

Table 3.1. Linking Study Sample Demographics

		%Students by Grade								
		Linking Study Sample						Spring 2019 STAAR Spanish Reading Student Population		
		Unweighted			Weighted					
Demographic Subgroup		3	4	5	3	4	5	3	4	5
	Total N	1,148	1,015	763	1,148	1,015	763	33,060	25,566	15,979
Race	Hispanic	98.3	96.8	93.4	98.7	98.7	98.5	98.7	98.7	98.5
	Other	0.5	0.6	4.2	0.5	0.5	0.5	0.5	0.5	0.5
	White	1.1	2.6	2.4	0.7	0.9	1.0	0.7	0.9	1.0
Sex	Female	53.2	49.7	50.5	50.5	50.4	49.7	50.5	50.4	49.7
	Male	46.8	50.3	49.5	49.5	49.6	50.3	49.5	49.6	50.3
Performance Level	*Did Not Meet*	28.8	38.2	14.4	30.6	40.7	20.4	30.6	40.7	20.4
	Approaches	32.3	31.6	25.7	30.5	30.4	26.7	30.5	30.4	26.7
	Meets	18.0	19.7	34.5	18.0	16.6	31.6	18.0	16.6	31.6
	Masters	20.8	10.4	25.4	20.8	12.3	21.3	20.8	12.3	21.3

Despite having a sample size being below the usual threshold of 1,000 students, fifth grade was included in the study sample because of the robust correlation and classification accuracy statistics observed between the two assessments for this grade cohort.

3.2. Descriptive Statistics

Table 3.2 presents descriptive statistics of the Spanish MAP Growth Reading and STAAR Spanish Reading test scores from Spring 2019, including the correlation coefficient (*r*) between them. The coefficients between the scores range from 0.74 to 0.77. These values indicate a high positive correlation among the scores, which is important validity evidence for the claim that MAP Growth scores are good predictors of performance on the STAAR Spanish Reading assessments.

Table 3.2. Descriptive Statistics of Test Scores

Grade	N	*r*	STAAR Spanish Reading*				Spanish MAP Growth Reading*			
			Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
3	1,148	0.74	1409.8	159.7	1073	1978	193.4	14.4	149	234
4	1,015	0.77	1454.6	159.1	119	2056	199.9	14.8	150	240
5	763	0.77	1580.6	145.1	1154	2146	205.6	14.6	158	249

*SD = standard deviation. Min. = minimum. Max. = maximum.

3.3. MAP Growth Cut Scores

Table 3.3 presents the STAAR Spanish Reading scale score ranges and the corresponding Spanish MAP Growth Reading RIT cut scores and percentile ranges by content area and grade.[2] Bolded numbers indicate the cut scores considered to be at least proficient for accountability purposes. These tables can be used to predict a student's likely performance level on the STAAR Spanish Reading spring assessment when MAP Growth is taken in the fall, winter, or spring. For example, a third-grader who obtained a Spanish MAP Growth Reading RIT score of 193 in the fall is likely to achieve *Meets Grade Level* performance on the STAAR Spanish Reading test. A third-grader who obtained a Spanish MAP Growth Reading RIT score of 199 in the spring is also likely to achieve *Meets Grade Level* performance on the STAAR Spanish Reading assessment. The spring cut score is higher than the fall cut score because growth is expected between fall and spring as students receive more instruction during the school year.

Table 3.3. MAP Growth Cut Scores

STAAR Spanish Reading								
Grade	*Did Not Meet*		*Approaches*		*Meets*		*Masters*	
3	668–1317		1318–1443		**1444**–1531		1532–1978	
4	711–1412		1413–1538		**1539**–1635		1636–2056	
5	782–1460		1461–1581		**1582**–1700		1701–2146	
Spanish MAP Growth Reading								
	Did Not Meet		*Approaches*		*Meets*		*Masters*	
Grade	RIT	Percentile	RIT	Percentile	RIT	Percentile	RIT	Percentile
Fall								
3	100–178	1–40	179–192	41–78	**193**–200	79–91	201–350	92–99
4	100–188	1–44	189–202	45–78	**203**–212	79–92	213–350	93–99
5	100–189	1–29	190–202	30–62	**203**–214	63–87	215–350	88–99
Winter								
3	100–185	1–42	186–197	43–74	**198**–204	75–87	205–350	88–99
4	100–194	1–46	195–206	47–77	**207**–214	78–90	215–350	91–99
5	100–193	1–28	194–206	29–63	**207**–216	64–84	217–350	85–99
Spring								
3	100–186	1–42	187–198	43–74	**199**–205	75–87	206–350	88–99
4	100–196	1–46	197–207	47–75	**208**–215	76–89	216–350	90–99
5	100–197	1–31	198–208	32–61	**209**–217	62–82	218–350	83–99

[2] The fall and winter cuts are not adjusted for instructional weeks as configured by partners.

3.4. Classification Accuracy

Table 3.4 presents the classification accuracy summary statistics, including the overall classification accuracy rate. These results indicate how well the Spanish MAP Growth Reading spring RIT scores predict proficiency on the STAAR Spanish Reading tests, providing insight into the predictive validity of MAP Growth. The overall classification accuracy rate ranges from 0.79 to 0.82. These values suggest that the RIT cut scores are good at classifying students as proficient or not proficient on the STAAR Spanish Reading assessment.

Although the results show that MAP Growth Reading scores can be used to predict student proficiency on the STAAR Spanish Reading test with relatively high accuracy, there is a notable limitation to how these results should be used and interpreted. The STAAR Spanish Reading and MAP Growth Reading assessments are designed for different purposes and measure slightly different constructs even within the same content area. Therefore, scores on the two tests cannot be assumed to be interchangeable. MAP Growth may not be used as a substitute for the state tests and vice versa.

Table 3.4. Classification Accuracy Results

Grade	N	Spanish Reading Cut Score		Class. Accuracy*	Rate*		Sensitivity	Specificity	Precision	AUC*
		MAP Growth	STAAR		FP	FN				
3	1,148	199	1444	0.79	0.18	0.25	0.75	0.82	0.72	0.87
4	1,015	208	1539	0.82	0.15	0.25	0.75	0.85	0.67	0.89
5	763	209	1582	0.81	0.12	0.25	0.75	0.88	0.88	0.91

*Class. Accuracy = overall classification accuracy rate. FP = false positives. FN = false negatives. AUC = area under the ROC curve.

3.5. Proficiency Projections

Table 3.5 presents the estimated probability of achieving *Meets Grade Level* performance on the STAAR Spanish Reading test based on RIT scores from fall, winter, or spring. “Prob.” indicates the probability of obtaining proficiency status on the STAAR Spanish Reading test in the spring. For example, a third-grader who obtained a Spanish MAP Growth Reading score of 200 in the fall has a 78% chance of reaching *Meets Grade Level* or higher on the STAAR Spanish Reading test in the spring.

Table 3.5. Proficiency Projections based on RIT Scores

Spanish Reading											
			Fall			Winter			Spring		
Grade	Start %ile	Spring Cut	Fall RIT	Projected Proficiency		Winter RIT	Projected Proficiency		Spring RIT	Projected Proficiency	
				Meets	Prob.		*Meets*	Prob.		*Meets*	Prob.
3	5	199	159	No	<0.01	166	No	<0.01	166	No	<0.01
	10	199	164	No	<0.01	171	No	<0.01	171	No	<0.01
	15	199	168	No	<0.01	174	No	<0.01	175	No	<0.01
	20	199	170	No	<0.01	177	No	<0.01	178	No	<0.01
	25	199	173	No	0.01	179	No	<0.01	180	No	<0.01
	30	199	175	No	0.02	181	No	0.01	182	No	<0.01
	35	199	177	No	0.03	183	No	0.01	184	No	<0.01
	40	199	179	No	0.06	185	No	0.02	186	No	<0.01
	45	199	180	No	0.08	187	No	0.05	188	No	<0.01
	50	199	182	No	0.10	189	No	0.09	190	No	<0.01
	55	199	184	No	0.15	190	No	0.11	191	No	0.01
	60	199	186	No	0.22	192	No	0.18	193	No	0.03
	65	199	187	No	0.26	194	No	0.27	195	No	0.11
	70	199	189	No	0.35	196	No	0.38	197	No	0.27
	75	199	191	No	0.40	198	Yes	0.50	199	Yes	0.50
	80	199	194	Yes	0.55	200	Yes	0.62	201	Yes	0.73
	85	199	196	Yes	0.65	203	Yes	0.78	204	Yes	0.94
	90	199	200	Yes	0.78	206	Yes	0.89	208	Yes	>0.99
	95	199	205	Yes	0.90	211	Yes	0.98	213	Yes	>0.99
4	5	208	167	No	<0.01	172	No	<0.01	175	No	<0.01
	10	208	172	No	<0.01	177	No	<0.01	180	No	<0.01
	15	208	176	No	<0.01	181	No	<0.01	184	No	<0.01
	20	208	179	No	<0.01	184	No	<0.01	187	No	<0.01
	25	208	181	No	0.01	186	No	<0.01	189	No	<0.01
	30	208	183	No	0.01	188	No	<0.01	191	No	<0.01
	35	208	185	No	0.02	190	No	<0.01	193	No	<0.01
	40	208	187	No	0.04	192	No	0.01	195	No	<0.01
	45	208	189	No	0.07	194	No	0.02	196	No	<0.01
	50	208	191	No	0.09	196	No	0.04	198	No	<0.01
	55	208	193	No	0.14	198	No	0.08	200	No	0.01
	60	208	195	No	0.17	200	No	0.15	202	No	0.03
	65	208	197	No	0.25	202	No	0.19	204	No	0.11
	70	208	199	No	0.34	204	No	0.30	205	No	0.17
	75	208	201	No	0.39	206	No	0.43	208	Yes	0.50
	80	208	203	Yes	0.50	208	Yes	0.57	210	Yes	0.73
	85	208	206	Yes	0.61	211	Yes	0.76	213	Yes	0.94
	90	208	210	Yes	0.75	215	Yes	0.92	216	Yes	0.99
	95	208	215	Yes	0.89	220	Yes	0.99	221	Yes	>0.99

Spanish Reading											
			Fall			Winter			Spring		
Grade	Start %ile	Spring Cut	Fall RIT	Projected Proficiency		Winter RIT	Projected Proficiency		Spring RIT	Projected Proficiency	
				Meets	Prob.		*Meets*	Prob.		*Meets*	Prob.
5	5	209	174	No	<0.01	178	No	<0.01	182	No	<0.01
	10	209	179	No	<0.01	183	No	<0.01	187	No	<0.01
	15	209	183	No	0.01	187	No	<0.01	190	No	<0.01
	20	209	185	No	0.02	190	No	<0.01	193	No	<0.01
	25	209	188	No	0.05	192	No	0.01	195	No	<0.01
	30	209	190	No	0.07	194	No	0.02	197	No	<0.01
	35	209	192	No	0.11	196	No	0.04	199	No	<0.01
	40	209	194	No	0.14	198	No	0.08	201	No	0.01
	45	209	196	No	0.21	200	No	0.14	203	No	0.03
	50	209	198	No	0.29	202	No	0.24	205	No	0.11
	55	209	200	No	0.34	204	No	0.29	206	No	0.17
	60	209	202	No	0.45	206	No	0.43	208	No	0.38
	65	209	204	Yes	0.50	207	Yes	0.50	210	Yes	0.62
	70	209	206	Yes	0.61	209	Yes	0.64	212	Yes	0.83
	75	209	208	Yes	0.71	212	Yes	0.82	214	Yes	0.94
	80	209	210	Yes	0.75	214	Yes	0.86	216	Yes	0.99
	85	209	213	Yes	0.86	217	Yes	0.95	219	Yes	>0.99
	90	209	217	Yes	0.93	220	Yes	0.98	222	Yes	>0.99
	95	209	222	Yes	0.98	226	Yes	>0.99	227	Yes	>0.99

4. References

He, W., & Meyer, J. (2021). *MAP Growth universal screening benchmarks: Establishing MAP Growth as an effective universal screener*. NWEA Research Report. https://www.nwea.org/content/uploads/2021/05/MAP-Growth-Universal-Screening-Benchmarks-2021-03-12_NWEA_report.pdf

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking*. Springer.

Lumley, T. (2019). *Survey: Analysis of complex survey samples*. R package version 3.36. https://CRAN.R-project.org/package=survey.

NWEA. (2021). *Spanish MAP Growth Reading technical report*. https://www.nwea.org/content/uploads/2021/05/Spanish-MAP-Growth-Reading-Technical-Report-2021-03-17.pdf

Pommerich, M., Hanson, B., Harris, D., & Sconing, J. (2004). Issues in conducting linkage between distinct tests. *Applied Psychological Measurement*, *28*(4), 247–273.