Linking Study Report: Predicting Performance on the State of Texas Assessments of Academic Readiness (STAAR) End-of-Course (EOC) Algebra 1 Assessment based on NWEA MAP Growth Scores

July 2020

NWEA Psychometric Solutions

nwea

Suggested citation: NWEA. (2020). *Linking study report: Predicting performance on the State of Texas Assessments of Academic Readiness (STAAR) End-of-Course (EOC) Algebra 1 assessment based on NWEA MAP Growth scores*. Portland, OR: Author.

Table of Contents

List of Tables

Executive Summary

To predict student achievement on the State of Texas Assessments of Academic Readiness (STAAR) End-of-Course (EOC) Algebra 1 assessment, NWEA® conducted a linking study using Spring 2019 data to derive Rasch Unit (RIT) cut scores on the MAP® Growth™ Algebra 1 assessment that correspond to the STAAR performance levels. With this information, educators can identify students at risk of failing to meet state proficiency standards early in the year and provide tailored educational interventions. The linking study has been created using the new 2020 NWEA MAP Growth norms (Thum & Kuhfeld, 2020).

Table E.1 presents the STAAR *Meets* 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 EOC test and those who are not. For example, the *Meets Grade Level* cut score on the STAAR Algebra 1 test is 4000. A student with a MAP Growth Algebra 1 RIT score of 232 in the fall is likely to meet proficiency on the STAAR Algebra 1 test, whereas a student with a RIT score lower than 232 in the fall is in jeopardy of not meeting proficiency.

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

Assessment		*Meets Grade Level* Cut Scores
STAAR Algebra 1		4000
MAP Growth Algebra 1	Fall	232
	Spring	240

Please note that the results in this report may differ from those found in the NWEA reporting system for individual districts. The typical growth scores from fall to spring used in this report are based on the default instructional weeks most commonly encountered for each term (i.e., Weeks 4 and 32 for fall and spring, respectively). However, instructional weeks often vary by district, so the cut scores in this report may differ slightly from the MAP Growth score reports that reflect spring instructional weeks set by partners.

E.1. Assessment Overview

The STAAR Algebra 1 test is part of Texas' state summative assessment system aligned to the Texas Essential Knowledge and Skills (TEKS) curriculum. The STAAR EOC tests are part of Texas' graduation requirements and are designed to ensure that students are learning the specific course material. 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. MAP Growth tests 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 100–350.

E.2. Linking Methods

Based on scores from the Spring 2019 test administration, the equipercentile linking method was used to identify the spring RIT scores that correspond to the STAAR Algebra 1 performance level cut scores. MAP Growth fall cut scores that predict proficiency on the STAAR Algebra 1 test were then projected using the 2020 NWEA growth norms that provide expected score gains across test administrations. Growth norms for MAP Growth Algebra 1 are only available for fall-to-spring projections for all eligible grades combined, so only the fall RIT cut was estimated regardless of grade level.

E.3. Student Sample

Only students who took both the MAP Growth Algebra 1 and STAAR Algebra 1 assessments in Spring 2019 were included in the study sample. From the 18 districts and 83 schools who participated in this study, 7,772 students were included in the linking study sample. 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 not represent the general student population as well as it should. To ensure that the linking study sample represents the state student population in terms of race, sex, and performance level, weighting (i.e., a statistical method that matches the distributions of the variables of interest to those of the target population) was applied to the sample. 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.4. Test Score Relationships

The correlation between STAAR Algebra 1 scores and MAP Growth Algebra 1 RIT scores is 0.77. This indicates a strong relationship among the scores, which is important validity evidence for the claim that MAP Growth Algebra 1 scores are good predictors of performance on the STAAR Algebra 1 assessments.

E.5. Accuracy of MAP Growth Classifications

Classification accuracy statistics indicate the proportion of students correctly classified by their RIT scores as proficient or not proficient on the STAAR Algebra 1 test. The overall MAP Growth Algebra 1 *Meets Grade Level* cut score has a 0.82 classification accuracy rate, meaning it accurately classified student achievement on the state test for 82% of the sample. This indicates that RIT scores on the Algebra 1 test have a high accuracy rate of identifying student proficiency on the STAAR Algebra 1 test.

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 in July 2020 to statistically connect the scores of the State of Texas Assessments of Academic Readiness (STAAR) End-of-Course (EOC) Algebra 1 assessment with Rasch Unit (RIT) scores from the MAP Growth Algebra 1 assessment taken during the Spring 2019 term. The linking study has been created using the new 2020 NWEA MAP Growth norms (Thum & Kuhfeld, 2020). This report presents the following results:

1. Student sample demographics
2. Descriptive statistics of test scores
3. MAP Growth Algebra 1 cut scores that correspond to the STAAR Algebra 1 performance levels using the equipercentile linking procedure for the spring results and the 2020 norms for the fall results
4. Classification accuracy statistics to determine the degree to which MAP Growth accurately predicts student proficiency status on the STAAR Algebra 1 test
5. The probability of achieving grade-level proficiency on the STAAR Algebra 1 assessment based on MAP Growth Algebra 1 RIT scores from fall and spring using the 2020 norms

1.2. Assessment Overview

The STAAR Algebra 1 assessment is part of Texas' state summative assessment system aligned to the Texas Essential Knowledge and Skills (TEKS) curriculum. The assessment has three cut scores (i.e., the minimum score a student must get on a test to be placed in a certain performance level) 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.

MAP Growth interim assessments from NWEA are computer adaptive and aligned to state-specific content standards. Scores are reported on the RIT vertical scale with a range of 100–350. Each content area has its own scale. To aid the interpretation of scores, NWEA periodically conducts norming studies of student and school performance on MAP Growth. Achievement status norms show how well a student performed on the MAP Growth test compared to students in the norming group by associating the student's performance on the MAP Growth test, expressed as a 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). The most recent norms study was conducted in 2020 (Thum & Kuhfeld, 2020).

2. Methods

2.1. Data Collection

This linking study is based on data from the Spring 2019 administrations of the MAP Growth and STAAR Algebra 1 assessments. NWEA recruited Texas districts to participate in the study by sharing their student and score data from the STAAR Algebra 1 test taken in Spring 2019. Districts also gave NWEA permission to access students' associated MAP Growth scores from the NWEA in-house database. Once Texas state score information was received by NWEA, each student's state testing record was matched to the MAP Growth score by using the student's first and last names, date of birth, student ID, and other available identifying information. Only students who took both MAP Growth Algebra 1 and the STAAR Algebra 1 assessment 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 population in terms of race, sex, and performance level. These variables were selected because they are correlated with the student's academic achievement within this study and are often provided in the data for the state population. The weighted sample matches the target population as closely as possible on the key demographics and test score characteristics. Specifically, a raking procedure was used to calculate the post-stratification weights and improve the representativeness of the sample. Raking uses iterative procedures to obtain weights that match sample marginal distributions to known population margins. The following steps were taken during this process:

- Calculate marginal distributions of race, sex, and performance level for the sample and population.
- Calculate post-stratification weights with the rake function from the survey package in R (Lumley, 2019).
- Trim the weight if it is not in the range of 0.3 to 3.0.
- Apply the weights to the sample before conducting the linking study analyses.

2.3. MAP Growth Cut Scores

The equipercentile linking method (Kolen & Brennan, 2004) was used to identify the spring RIT scores that correspond to the STAAR Algebra 1 cut scores. Since the state Algebra 1 test is not grade-dependent (i.e., any student can take the assessment once they finish the course), the spring RIT cuts were established based on all the students in the study sample regardless of their grades. The growth norms for the MAP Growth Algebra 1 test are available for fall-to-spring projections. Therefore, only the fall RIT cut was reported using the 2020 growth norms and the spring RIT cuts.

Percentile ranks are also provided that show how a nationally representative sample of students in the same grade scored on MAP Growth for each administration, which is an important interpretation of MAP Growth test scores. This is useful for understanding (1) how student scores compare to peers nationwide and (2) the relative rigor of a state's performance level designations for its summative assessment.

The MAP Growth spring cut scores could be calculated using the equipercentile linking method because that data are directly connected to the STAAR Algebra 1 spring data used in the study. 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 Algebra 1). Its equipercentile equivalent score on Test Y (e.g., MAP Growth Algebra 1), $e_y(x)$, 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 Algebra 1 on the scale of MAP Growth, $P(x)$ is the percentile rank of a given score on STAAR, 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 MAP Growth conditional growth norms provide students' expected score gains across terms, such as growth from fall to spring within the same grade or from spring of a lower grade to the spring of the adjacent higher grade. This information can be used to calculate the fall RIT 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) to the spring RIT.

2.4. Classification Accuracy

The degree to which MAP Growth predicts student proficiency status on the STAAR Algebra 1 assessment can be described using classification accuracy statistics based on the MAP Growth spring cut scores that show the proportion of students correctly classified by their RIT scores as proficient (*Meets* or *Masters*) or not proficient (*Did Not Meet* or *Approaches*). Table 2.1 describes the classification accuracy statistics provided in this report (Pommerich, Hanson, Harris, & Sconing, 2004). The results are based on the Spring 2019 MAP Growth and STAAR Algebra 1 data for the *Meets Grade Level* cut score.

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

Statistic	Description*	Interpretation
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
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 Projection

In addition to calculating the MAP Growth fall cut scores, the MAP Growth conditional growth norms data were also used to calculate the probability of reaching proficiency on the STAAR Algebra 1 test based on a student's RIT scores from fall and spring. Equation 3 was used to calculate the probability of a student achieving *Meets Grade Level* on the STAAR Algebra 1 test based on their fall 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 a standardized normal cumulative distribution.
- $RIT_{previous}$ is the student's RIT score in fall.
- g is the expected growth from the previous RIT (e.g., fall) 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* on the STAAR Algebra 1 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 MAP Growth Algebra 1 and STAAR Algebra 1 assessments in Spring 2019 were included in the study sample. Data used in this study were collected from 18 districts and 83 schools in Texas. Table 3.1 presents the demographic distributions of race, sex, and performance level of the student population that took the Spring 2019 STAAR Algebra 1 test (TEA, 2017). It also presents the demographic distributions in the original unweighted study sample and the weighted sample. Since the unweighted data are different from the general STAAR Algebra 1 population, post-stratification weights were applied to the linking study sample to improve its representativeness. The analyses in this study were therefore conducted based on the weighted sample.

Table 3.1. Linking Study Sample Demographics

Demographic Subgroup		%Students		
		Texas Population*	Unweighted Sample	Weighted Sample
	Total N-Count	416,354	7,772	7,772
Race	Asian	4.3	4.0	4.3
	Black	13.2	16.8	13.2
	Hispanic	53.1	49.4	53.1
	Multi-Race	2.2	1.9	2.2
	Other	0.6	2.0	0.6
	White	26.6	25.8	26.6
Sex	Female	48.1	49.5	48.1
	Male	51.9	50.5	51.9
Performance Level	*Did Not Meet*	16.2	10.8	16.2
	Approaches	22.1	24.0	22.1
	Meets	22.3	25.4	22.3
	Masters	39.4	39.8	39.4

*The number of students who took the STAAR Algebra 1 assessment in Spring 2019.

3.2. Descriptive Statistics

Table 3.2 presents descriptive statistics of the MAP Growth and STAAR Algebra 1 test scores from Spring 2019, including the correlation coefficient (*r*) between them. The correlation coefficient is 0.77. This indicates a strong relationship among the scores, which is important validity evidence for the claim that MAP Growth Algebra 1 scores are good predictors of performance on the STAAR Algebra 1 assessment.

Table 3.2. Descriptive Statistics of Test Scores

Assessment*	N	*r*	Mean	SD	Min.	Max.
MAP Growth Algebra 1	7,772	0.77	4200.5	634.7	2466	6181
STAAR Algebra 1			243.5	20.8	180	307

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

3.3. MAP Growth Cut Scores

Table 3.3 presents the STAAR Algebra 1 scale score ranges and the corresponding MAP Growth RIT cut scores and percentile ranges. These tables can be used to predict a student's likely performance level on the STAAR Algebra 1 assessment when MAP Growth is taken in the fall or spring. For example, a student who obtained a MAP Growth Algebra 1 RIT score of 232 in the fall is likely to reach *Meets Grade Level* on the STAAR Algebra 1 test. A student who obtained a RIT score of 240 in the spring is also likely to reach *Meets Grade Level* performance. 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.

Within this report, the cut scores for fall are derived from the spring cuts and the typical growth scores from fall-to-spring. The typical growth scores are based on the default instructional weeks most commonly encountered for each term (Weeks 4 and 32 for fall and spring, respectively). Since instructional weeks often vary by district, the cut scores in this report may differ slightly from the MAP Growth score reports that reflect instructional weeks set by partners. If the actual instructional weeks deviate from the default ones, a student's projected performance level could be different from the generic projection presented in this document. Partners are therefore encouraged to use the projected performance level in students' profile, classroom, and grade reports in the NWEA reporting system since they reflect the specific instructional weeks set by partners.

Table 3.3. MAP Growth Cut Scores

STAAR Algebra 1								
EOC Test	***Did Not Meet***		***Approaches***		***Meets***		***Masters***	
Algebra 1	0–3549		3550–3999		**4000**–4332		4333–6373	
MAP Growth Algebra 1*								
	Did Not Meet		***Approaches***		***Meets***		***Masters***	
Grade	**RIT**	**Percentile**	**RIT**	**Percentile**	**RIT**	**Percentile**	**RIT**	**Percentile**
Fall	100–212	1–12	213–231	13–52	**232**–242	53–76	243–350	77–99
Spring	100–219	1–15	220–239	16–51	**240**–251	52–75	252–350	76–99

*Cut scores for fall are derived from the spring cuts and growth norms based on the typical instructional weeks. Bolded numbers indicate the cut scores considered to be at least proficient for accountability purposes.

3.4. Classification Accuracy

Table 3.4 presents the classification accuracy summary statistics, including the overall classification accuracy rate. These results indicate how well MAP Growth spring RIT scores predict proficiency on the STAAR Algebra 1 test, providing insight into the predictive validity of MAP Growth. The overall classification accuracy rate is 0.82, which suggests that the RIT cut scores are good at classifying students as proficient or not proficient on the STAAR Algebra 1 assessment.

Although the results show that MAP Growth Algebra 1 scores can be used to accurately classify students as likely to be proficient on the STAAR Algebra 1 test, there is a notable limitation to how these results should be used and interpreted. STAAR and MAP Growth 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

N	Cut Score		Class. Accuracy*	Rate*		Sensitivity	Specificity	Precision	AUC*
	MAP Growth	STAAR Algebra 1		FP	FN				
7,772	240	4000	0.82	0.19	0.18	0.82	0.81	0.87	0.90

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

3.5. Proficiency Projection

Table 3.5 presents the estimated probability of achieving *Meets Grade Level* performance on the STAAR Algebra 1 test based on RIT scores from fall or spring. “Prob.” indicates the probability of obtaining proficient status on the STAAR test in the spring. For example, a student who obtained a MAP Growth Algebra 1 score of 242 in the fall has an 87% chance of reaching *Meets Grade Level* or higher on the STAAR Algebra 1 test.

Table 3.5. Proficiency Projection based on RIT Scores

Algebra 1							
		Fall			Spring		
Start %ile	Spring Cut	Fall RIT	Projected Proficiency		Spring RIT	Projected Proficiency	
			Meets	Prob.		*Meets*	Prob.
5	240	205	No	<0.01	207	No	<0.01
10	240	210	No	0.01	214	No	<0.01
15	240	214	No	0.02	219	No	<0.01
20	240	217	No	0.04	223	No	<0.01
25	240	220	No	0.07	226	No	<0.01
30	240	223	No	0.13	229	No	<0.01
35	240	225	No	0.22	231	No	<0.01
40	240	227	No	0.29	234	No	0.03
45	240	229	No	0.37	236	No	0.11
50	240	231	No	0.46	239	No	0.38
55	240	233	Yes	0.54	241	Yes	0.62
60	240	235	Yes	0.63	244	Yes	0.89
65	240	237	Yes	0.71	246	Yes	0.97
70	240	239	Yes	0.78	249	Yes	>0.99
75	240	242	Yes	0.87	252	Yes	>0.99
80	240	244	Yes	0.93	255	Yes	>0.99
85	240	248	Yes	0.97	259	Yes	>0.99
90	240	251	Yes	0.99	263	Yes	>0.99
95	240	257	Yes	>0.99	270	Yes	>0.99

References

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

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

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

Texas Education Agency (TEA). (2017). *STAAR statewide summary reports 2016–2017*. Retrieved from https://tea.texas.gov/student-assessment/testing/staar/staar-statewide-summary-reports-2016-2017.

Thum, Y. M., & Kuhfeld, M. (2020). *NWEA 2020 MAP Growth achievement status and growth norms for students and schools*. NWEA Research Report. Portland, OR: NWEA.