MAP Growth Linking Studies: Intended Uses, Methodology, and Recent Studies

November 2021

Ann Hu, NWEA Psychometric Solutions

nwea

Table of Contents

List of Tables

Acknowledgements

It has been a great team effort to update the linking studies for up to 39 states in 2020–2021. I would like to express my gratitude to the NWEA partners who participated in the studies by sharing their test results with us. I would also like to thank the Legal team for preparing the data sharing agreements and the Partner Accounts team for helping us plan the study, recruit participants, collect state information, and provide input for key decisions. While planning for the linking studies, the Content Solutions team also provided valuable information about the content coverage between MAP Growth and state summative assessments, especially for the high school course-specific tests, which helped us to make decisions about the study. Their prompt inputs are greatly appreciated. I am also grateful to many of those in the Marketing, Project Management, Product Operations, and Research teams for their assistance to this project.

Finally, I would like to extend my sincere thanks to my fellow Psychometric Solutions team members who have been involved in the linking studies. Thank you, Debbie Bauman, for managing this project and communicating with the partners and internal teams; Sarah Tran and Rudraxi Majumdar for preparing the data files and conducting the analyses; Kelly Rivard for writing the reports; and Patrick Meyer for designing the report and supervising this project. Without your hard work and collaborations, it would be impossible to complete all these studies in a timely manner.

1. Introduction

MAP® Growth™ linking studies from NWEA® allow partners to use students' MAP Growth Rasch Unit (RIT) scores from fall, winter, and spring to predict their performance on the spring state summative assessment. This document presents the intended uses and methodology of the MAP Growth linking studies, a description of the results provided in the linking study reports, and a summary of the recent linking studies conducted by NWEA to incorporate the new 2020 norms (Thum & Kuhfeld, 2020).

1.1. MAP Growth Overview

MAP Growth assessments are interim adaptive tests that measure a student's academic achievement and growth in mathematics, reading, language usage, and science. It is administered in the fall, winter, and spring, with a fourth optional administration in the summer. MAP Growth has a large item bank containing dichotomously scored items aligned to various content standards. For each state, NWEA content specialists unpack the standards, select items from the MAP Growth item bank that directly align to the standard statements, and develop additional items as needed to ensure coverage of the standards. MAP Growth scores are reported on the RIT vertical scale. RIT scores have a mean of 200 and a standard deviation of 10. Scores typically range from 100 to 350. For a full description of MAP Growth assessments, please refer to the technical report (NWEA, 2019).

1.2. Purpose and Intended Uses of a Linking Study

Given that the MAP Growth item pool for a specific state is carefully assembled to measure a wide range of content standards adopted by the state, it is expected that MAP Growth RIT scores are highly correlated with scores on the state summative assessments that measure the same content standards. Therefore, a link between RIT scores and performance levels on the state summative tests can be established through linking studies. While linking studies are most often conducted in Grades 3–8 in English language arts (ELA) and mathematics, studies are also conducted for high school when applicable and for science (e.g., in Grades 5 and 8) if enough data are available. For ELA and mathematics, the RIT scores for Grade 2 students are linked to the performance levels on the Grade 3 state summative assessment given that state tests are not provided in Grade 2.

Linking study results allow educators to use students' RIT scores from fall, winter, and early spring to predict students' performance levels on the state summative tests, which are often administered in the spring after the MAP Growth testing window has opened. As a result, educators can identify students who are at risk of not meeting state proficiency standards early in the year and provide tailored educational interventions or place students in the appropriate learning groups based on their entry levels.

Many states also have a Read by Grade 3 program that identifies students in lower grades who are behind in reading and request that schools provide additional support to help these students achieve reading success by the end of Grade 3. The linking study results for Grade 2 ELA can be used to help educators identify these students. Sometimes, a state may establish a unique cut score for the Grade 3 ELA assessment to make decisions such as grade retention. In this case, a special linking study can be conducted to derive the corresponding RIT cut to either identify students at risk, or to provide alternative evidence of not placing students in the retention program.

1.3. Conditions for a New Linking Study

Linking studies must be updated periodically to sustain the connection to the state summative assessment. Changes to the state summative and MAP Growth assessments can lead to a new linking study. Common conditions that require a new linking study to be conducted are summarized below. These conditions are closely monitored, and the feasibility of conducting a new linking study is evaluated whenever a change occurs.

1. A new state summative test is introduced, which may include the adoption of new content standards, score scales, performance levels or cut scores, test blueprint designs, test administration mode, etc.
2. Substantial changes are made to the existing state summative and/or MAP Growth assessments, which may include the item and content standard alignment, test blueprint designs, test administration mode, rebranding of test names, etc.
3. The existing linking study has not been updated for five years.
4. MAP Growth norms are updated.
5. A request for proposal (RFP) requires a new study.

2. Methodology

2.1. Study Sample

Linking studies are based on student data from the same spring administrations of the MAP Growth and state summative assessments. A minimum sample size of 1,000 students per grade and subject is needed to conduct a linking study. NWEA requires that districts within a state provide their state summative student data and permission to use their students' MAP Growth data in the study. Once state score information is received by NWEA, each student's state testing record is matched to their MAP Growth score based on the student's first and last names, student ID, and other available identifying information. Only students who take both the MAP Growth and state summative assessments in the designated spring term are included in the study sample.

The linking study sample is voluntary and can include only 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 is applied to statistically adjust the sample so it reflects the target population on these variables. As a result, the RIT cuts derived from the linking study sample can be generalized to any student from the target population.

The variables of race, sex, and performance level are used for post-stratification weighting because they are known to be correlated with students' academic achievement and are often available in state summative assessment reports. A raking procedure is 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 are 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 that are outside the range of 0.3 to 3.0.
- Apply the weights to the sample before conducting the linking study analyses.

As a result of post-stratification weighting, the weighted sample will match the target population as closely as possible on the key demographics and performance characteristics as defined by the state.

2.2. Methods to Derive MAP Growth Cut Scores

To derive the RIT cut scores, the equipercentile linking method is used to identify the summative test score percentiles that correspond to the spring MAP Growth score percentiles. The RIT-score equivalent of the summative cut score is then identified. MAP Growth fall and winter cut scores that predict proficiency on the spring state summative assessment are then projected using the 2020 NWEA conditional growth norms that provide expected score gains across test administrations. The norms are also used to identify the Grade 2 fall, winter, and spring cut scores. Percentile ranks are provided along with the RIT cuts to show how a nationally representative sample of students in the same grade scored on MAP Growth for each administration.

Specifically, the equipercentile linking method (Kolen & Brennan, 2004) is used to identify the spring MAP Growth RIT scores for Grades 3–8 that correspond to the spring state summative 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., State Test). 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:

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

where $e_y(x)$ is the equipercentile equivalent of score x on the state summative test on the scale of MAP Growth, $P(x)$ is the percentile rank of a given score on the state summative test, 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 is 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 within the same grade or from spring of a lower grade to the spring of the adjacent higher grade. This information is used to calculate the fall and winter cut scores for Grades 3–8. Equation 2 is used to determine the previous term's or grade's MAP Growth score needed to reach the spring cut score, considering the expected growth associated with the previous RIT score:

$$RIT_{PredSpring} = RIT_{previous} + g \tag{2}$$

where:

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

The MAP Growth conditional growth norms are also used to calculate the fall, winter, and spring cuts for Grade 2. Students typically do not begin taking the state summative assessment until Grade 3. Thus, cut scores for Grade 2 are interpolated by obtaining longitudinal data for the Grade 3 cohort. For each Grade 3 student in the study sample, their MAP Growth data from the prior year when they were in Grade 2 are obtained. In this way, the data come from the same cohort of students beginning when they were in Grade 2 and continuing through Grade 3. To derive the spring cut scores for Grade 2, the growth score from spring of one year to the next is used (i.e., the growth score from spring Grade 2 to spring Grade 3). The calculation of fall and winter cuts for Grade 2 follows the same process as Grades 3–8. For example, the growth score from fall to spring in Grade 2 is used to calculate the fall cuts for Grade 2.

2.3. Classification Accuracy

The degree to which MAP Growth predicts student proficiency status on the state summative 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 state summative test. A summary of how well the interpolated Grade 2 cuts predict Grade 3 proficiency status is also reported in the classification accuracy statistics. Table 2.1 describes the classification accuracy statistics provided in the linking study reports (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
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.

2.4. Proficiency Projections

In addition to calculating the MAP Growth fall and winter cut scores (and the Grade 2 cut scores), the MAP Growth conditional growth norms data are also used to calculate the probability of reaching proficiency on the state summative test based on a student’s RIT scores from fall, winter, and spring. Equation 3 is used to calculate the probability of a student achieving proficiency on the state summative test based on their fall or winter RIT score:

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

where:

- Φ is the standard normal cumulative distribution function.
- $RIT_{previous}$ is the student’s RIT score in fall or winter (or in spring of Grade 2).
- g is the expected growth from the previous RIT (e.g., fall or winter) to the spring RIT.
- $RIT_{SpringCut}$ is the MAP Growth proficient cut score for spring. For Grade 2, this is the Grade 3 cut score for spring.
- SD is the conditional standard deviation of the expected growth, g.

Equation 4 is used to estimate the probability of a student achieving proficiency on the state summative test based on their spring RIT score (RIT_{Spring}):

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

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

3. Results

The RIT score predictions are included in the NWEA reporting system and provided on individual score reports. A linking study report is also created for each state-specific or consortium study that explains the methodology, summarizes the study sample, presents the results, and provides evidence that shows how well the RIT scores can predict state summative performance. Specifically, each linking study report presents the following information:

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 state summative assessment
4. Classification accuracy statistics to determine the degree to which MAP Growth accurately predicts student proficiency status on the state summative test
5. The probability of achieving grade-level proficiency on the state assessment based on MAP Growth RIT scores from fall, winter, and spring

Partners are cautioned that the results in the report may differ from those found in the NWEA reporting system for individual districts. The typical growth scores from fall to spring or winter to spring used in this report are based on the default instructional weeks most encountered for each term (i.e., Weeks 4, 20, and 32 for fall, winter, and spring, respectively). However, instructional weeks often vary by district, so the cut scores in the report may differ slightly from the MAP Growth score reports that reflect the specific instructional weeks set by partners. Partners are therefore encouraged to use the projected performance level in students' score reports since they reflect the specific instructional weeks set by partners.

All current MAP Growth linking study reports are posted on the NWEA website at https://www.nwea.org/resource/type/linking-studies/. For illustrative purposes, the results for the linking study between the MAP Growth and Minnesota Comprehensive Assessments-Series III (MCA-III) Mathematics assessments are presented below (NWEA, 2020).

Table 3.1 presents descriptive statistics of the MAP Growth and MCA-III Mathematics test scores from Spring 2019, including the correlation coefficient (*r*) between them. The correlation coefficients between the scores range from 0.92 to 0.93. These values indicate a strong relationship among the scores, which is important validity evidence for the claim that MAP Growth scores are good predictors of performance on the MCA-III tests.

Table 3.1. Descriptive Statistics of Test Scores

			MCA-III*				MAP Growth*			
Grade	N	*r*	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
3	4,981	0.92	354.6	16.8	315	399	205.8	15.2	136	272
4	5,036	0.92	454.8	19.2	409	499	216.4	16.9	152	298
5	5,396	0.93	549.0	14.4	515	586	225.3	18.1	137	296
6	5,621	0.93	648.1	15.3	611	688	228.5	18.6	151	284
7	4,068	0.93	748.7	12.9	718	782	233.8	19.9	146	301
8	3,355	0.92	849.8	15.8	813	888	237.2	21.7	151	305

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

Table 3.2 presents the MCA-III scale score ranges and the corresponding MAP Growth RIT cut scores and percentile ranges. The *Meets* cut score on the MCA-III state test demarks the minimum level of achievement considered to be proficient, as shown by the bolded numbers. This table can be used to predict a student's likely performance level on the spring state summative assessment when MAP Growth is taken in the fall, winter, or spring. For example, a Grade 3 student who obtained a MAP Growth Mathematics RIT score of 189 in the fall is likely to reach *Meets* proficiency on the MCA-III Mathematics test.

Table 3.2. MAP Growth Cut Scores

MCA-III Mathematics				
Grade	***Does Not Meet***	***Partially Meets***	***Meets***	***Exceeds***
3	301–339	340–349	**350**–365	366–399
4	401–439	440–449	**450**–465	466–499
5	501–539	540–549	**550**–562	563–599
6	601–639	640–649	**650**–661	662–699
7	701–739	740–749	**750**–759	760–799
8	801–839	840–849	**850**–860	861–899

MAP Growth Mathematics								
	Does Not Meet		***Partially Meets***		***Meets***		***Exceeds***	
Grade	**RIT**	**Percentile**	**RIT**	**Percentile**	**RIT**	**Percentile**	**RIT**	**Percentile**
Fall								
2	100–166	1–26	167–175	27–52	**176**–191	53–89	192–350	90–99
3	100–180	1–28	181–188	29–51	**189**–202	52–85	203–350	86–99
4	100–192	1–31	193–200	32–53	**201**–214	54–85	215–350	86–99
5	100–204	1–38	205–215	39–66	**216**–231	67–92	232–350	93–99
6	100–210	1–40	211–221	41–66	**222**–236	67–91	237–350	92–99
7	100–214	1–37	215–227	38–66	**228**–243	67–90	244–350	91–99
8	100–218	1–37	219–230	38–62	**231**–246	63–87	247–350	88–99
Winter								
2	100–175	1–26	176–184	27–52	**185**–199	53–88	200–350	89–99
3	100–188	1–29	189–196	30–51	**197**–209	52–83	210–350	84–99
4	100–199	1–33	200–207	34–54	**208**–221	55–85	222–350	86–99
5	100–210	1–40	211–221	41–67	**222**–237	68–92	238–350	93–99
6	100–215	1–41	216–226	42–66	**227**–241	67–90	242–350	91–99
7	100–217	1–36	218–231	37–66	**232**–247	67–90	248–350	91–99
8	100–221	1–37	222–233	38–61	**234**–249	62–86	250–350	87–99
Spring								
2	100–181	1–28	182–189	29–51	**190**–204	52–87	205–350	88–99
3	100–193	1–30	194–201	31–52	**202**–214	53–83	215–350	84–99
4	100–203	1–33	204–211	34–53	**212**–225	54–83	226–350	84–99
5	100–214	1–40	215–225	41–66	**226**–241	67–91	242–350	92–99
6	100–218	1–40	219–229	41–65	**230**–244	66–89	245–350	90–99
7	100–220	1–37	221–234	38–66	**235**–250	67–89	251–350	90–99
8	100–223	1–37	224–235	38–60	**236**–251	61–85	252–350	86–99

Table 3.3 presents the classification accuracy summary statistics that indicate how well MAP Growth spring RIT scores predict proficiency on the MCA-III Mathematics test, providing insight into the predictive validity of MAP Growth. The overall classification accuracy rate ranges from 0.87 to 0.90, which suggests that the RIT cuts are good at classifying students as proficient or not proficient on the state test. For Grade 2, the classification accuracy rate refers to how well the MAP Growth cuts shown can predict students' proficiency status on MCA-III in Grade 3.

Table 3.3. Classification Accuracy Results

Grade	N	Cut Score		Class. Accuracy*	Rate*		Sensitivity	Specificity	Precision	AUC*
		MAP Growth	MCA-III		FP	FN				
2	4,197	190	350	0.87	0.22	0.09	0.91	0.78	0.90	0.93
3	4,981	202	350	0.90	0.17	0.07	0.93	0.83	0.91	0.96
4	5,036	212	450	0.90	0.17	0.06	0.94	0.83	0.90	0.96
5	5,396	226	550	0.89	0.15	0.08	0.92	0.85	0.87	0.96
6	5,621	230	650	0.90	0.12	0.08	0.92	0.88	0.88	0.97
7	4,068	235	750	0.90	0.12	0.08	0.92	0.88	0.89	0.97
8	3,355	236	850	0.90	0.12	0.09	0.91	0.88	0.90	0.96

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

Table 3.4 presents the estimated probability of achieving *Meets* performance on the Grade 3 MCA-III Mathematics test based on RIT scores from fall, winter, or spring. For example, a Grade 3 student who obtained a MAP Growth Mathematics score of 202 in the fall has a 97% chance of reaching *Meets* proficiency or higher on the MCA-III test. "Prob." indicates the probability of obtaining proficient status on the MCA-III test in the spring.

Table 3.4. Proficiency Projections based on RIT Scores

Grade	Start %ile	Spring Cut	Fall			Winter			Spring		
			Fall RIT	Projected Proficiency		Winter RIT	Projected Proficiency		Spring RIT	Projected Proficiency	
				Meets	Prob.		*Meets*	Prob.		*Meets*	Prob.
3	5	202	166	No	<0.01	174	No	<0.01	178	No	<0.01
	10	202	171	No	<0.01	179	No	<0.01	183	No	<0.01
	15	202	175	No	0.01	182	No	<0.01	186	No	<0.01
	20	202	177	No	0.03	185	No	0.01	189	No	<0.01
	25	202	179	No	0.05	187	No	0.02	192	No	<0.01
	30	202	181	No	0.10	189	No	0.04	194	No	<0.01
	35	202	183	No	0.17	191	No	0.10	196	No	0.02
	40	202	185	No	0.26	193	No	0.20	198	No	0.08
	45	202	187	No	0.37	195	No	0.33	199	No	0.15
	50	202	188	No	0.44	196	No	0.42	201	No	0.37
	55	202	190	Yes	0.56	198	Yes	0.58	203	Yes	0.63
	60	202	192	Yes	0.63	200	Yes	0.74	205	Yes	0.85
	65	202	194	Yes	0.74	201	Yes	0.80	207	Yes	0.96
	70	202	196	Yes	0.83	203	Yes	0.90	208	Yes	0.98
	75	202	198	Yes	0.90	205	Yes	0.96	211	Yes	>0.99
	80	202	200	Yes	0.95	208	Yes	0.99	213	Yes	>0.99
	85	202	202	Yes	0.97	210	Yes	>0.99	216	Yes	>0.99
	90	202	206	Yes	>0.99	214	Yes	>0.99	219	Yes	>0.99
	95	202	211	Yes	>0.99	219	Yes	>0.99	224	Yes	>0.99

4. Recent Studies

Linking studies use growth norms and the spring RIT cuts to derive the fall and winter cut scores, and the status norms (i.e., percentiles) are reported along with the RIT scores. Therefore, all existing linking studies needed to be updated after NWEA released the new norms in 2020. As such, linking studies for the following 39 states were updated using the most recent 2020 norms in 2020–2021. Of those 39 states, six of them use the collective Smarter Balanced Assessment Consortium (SBAC) linking study results (i.e., Connecticut, Delaware, Hawaii, Idaho, Montana, Vermont).

- Alaska
- Arizona
- Arkansas
- California
- Colorado
- Connecticut
- Delaware
- Florida
- Georgia
- Hawaii
- Idaho
- Illinois
- Indiana
- Iowa
- Kansas
- Kentucky
- Massachusetts
- Michigan
- Minnesota
- Mississippi
- Missouri
- Montana
- Nebraska
- Nevada
- New York
- North Carolina
- North Dakoda
- Ohio
- Oklahoma
- Oregon
- Pennsylvania
- South Carolina
- South Dakoda
- Tennessee
- Texas
- Vermont
- Virginia
- Washington
- Wisconsin

States that do not meet the minimum 1,000 study sample size or do not have enough MAP Growth partners cannot have a state-specific linking study. For these states, and for international schools, a special study was conducted to derive generic MAP Growth cut scores based on the median proficient RIT cuts from the 39 states with linking studies based on the 2020 norms. Partners can use these generic cuts as an indication of passing on MAP Growth, although state-specific studies are always preferred when possible.

All current MAP Growth linking study reports are posted on the NWEA website at https://www.nwea.org/resource/type/linking-studies/. Linking studies will continue to be updated as needed based on the close monitoring of state assessments and the presence of a condition that triggers the need for a new study.

5. References

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. (2019). *MAP Growth technical report*. https://www.nwea.org/research/publication/map-growth-technical-report/

NWEA. (2020). *Linking study report: Predicting performance on the Minnesota Comprehensive Assessments-Series III (MCA-III) based on NWEA MAP Growth scores*. https://www.nwea.org/resources/minnesota-linking-study/

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

Thum, Y. M., & Kuhfeld, M. (2020). *NWEA 2020 MAP Growth achievement status and growth norms for students and schools*. NWEA Research Report. https://teach.mapnwea.org/impl/normsResearchStudy.pdf