BRIEF

MAP Growth with enhanced item-selection algorithm: Updates on score comparability

Karyn Lewis and Megan Kuhfeld

June 2024

nwea Research

Introduction

NWEA® is dedicated to providing partners with assessments that deliver test data with the highest level of validity and reliability. To fulfill this commitment, NWEA developed an enhanced item-selection algorithm for the MAP® Growth™ assessment. This update more closely aligns the assessment with grade-level content to enhance its content validity. The enhanced item-selection algorithm (referred to as EISA throughout this report) prioritizes grade-level content while still adapting to off-grade items where necessary to provide items of appropriate difficulty for students.[1] Teachers, students, and other stakeholders will benefit from a tighter coupling between MAP Growth and the subject matter students have an opportunity to learn in school.

Prior to a broader implementation of MAP with EISA during the 2023–24 school year, NWEA conducted a pilot study in a sample of volunteer school districts. The aim was to gather evidence of the improved content validity and high score reliability of the new algorithm. In addition, the pilot study examined the comparability of scores derived from MAP with EISA and the traditional MAP Growth assessment using a matched sample of students. Students taking MAP with EISA demonstrated higher average math scores compared to those taking traditional MAP Growth. The differences varied by grade, ranging from less than 1 RIT point to over 6 RIT points, with an average increase of about 3 RIT points. Students' reading scores did not appreciably differ between the two assessments.

The purpose of this report is to expand our understanding of score comparability between MAP with EISA and traditional MAP Growth with data from the entire 2023–24 school year. With the benefit of a complete academic year of data in the 19 states that implemented MAP with EISA in the 2023–24 school year, we can now examine score comparability in fall, winter, and spring terms, as well examine the comparability of growth patterns across seasons.[2]

Score comparability in math

To investigate the comparability of test scores and growth patterns under the updated assessment, we analyze achievement and growth trends for the 2023–24 school year and compare these to the preceding year, 2022–23. Two district partner groups comprise our sample: those who consistently used traditional MAP Growth in both years (Stable Users) and those who transitioned to MAP with EISA during 2023–24 (Transition Users). Although these groups were not matched samples, examining trends for Stable Users and Transition Users separately allows us to disentangle the effects of the algorithm from other potential influences on achievement and growth trends.

Figure 1 illustrates average fall, winter, and spring math RIT scores for the 2023–24 and 2022–23 school years for **Stable Users** (left panel) and **Transition Users** (right panel).[3] The points represent average achievement levels, while the lines connecting data points indicate gains between the test seasons. Solid lines represent 2022–23, while dashed lines represent 2023–24. The figure also captures the magnitude of the year-to-year difference in a standardized effect-size metric. Positive values indicate 2023–24 scores are higher than 2022–23 scores. For example, eighth-grade scores for Transition Users were 0.14 standard deviations (SDs) higher in winter and 0.22 SDs higher in spring during 2023–2024 compared to 2022–23. For Stable Users, there was no difference in eighth-graders' winter scores and only a 0.02 SD increase in spring.

Consistent with the pilot study, partners that transitioned to MAP with EISA in 2023–24 experienced changes in math scores. Specifically, math scores in most grades were slightly lower than in the fall compared to the preceding year. By winter, math test scores in all grades were higher than the previous winter, and this trend

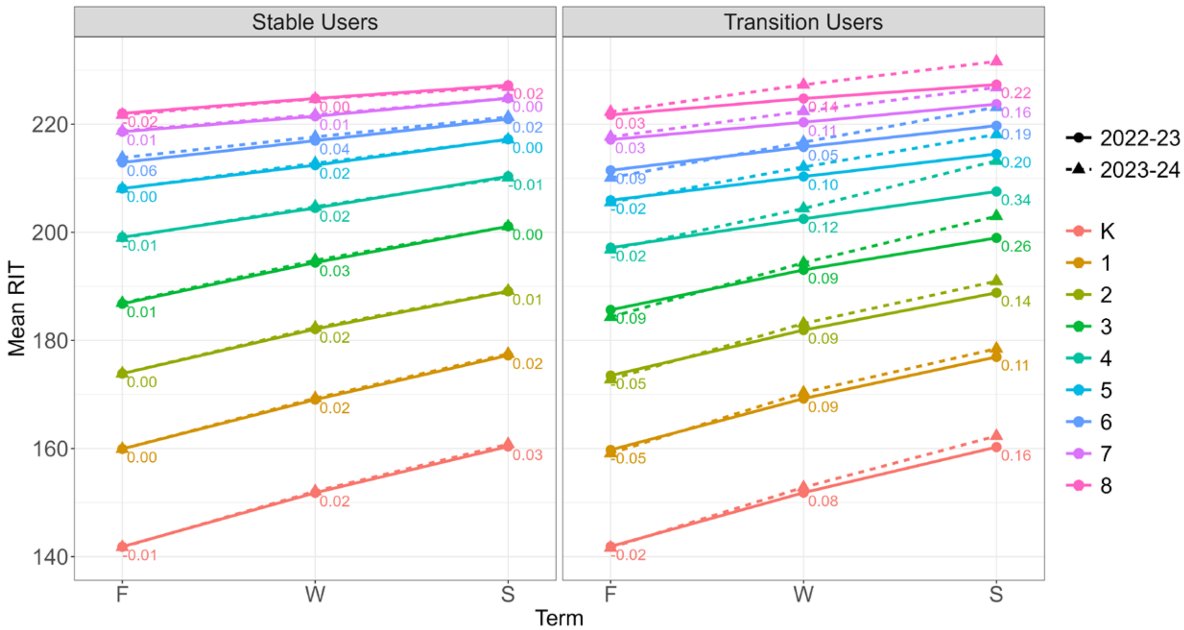
1 See the pilot study for technical details about the item-selection algorithm as well as evidence supporting the reliability and improved content validity of MAP Growth with EISA.

2 See this resource for details on the states that transitioned to MAP Growth with EISA in 2023–24.

3 This analysis is restricted to scores derived from the English version of the math test. The Spanish math test has not yet transitioned to the enhanced item-selection algorithm.

continued in the second half of the school year, with even larger differences observed in spring. An initial dip followed by greater than expected winter and spring test scores resulted in increased within-year growth compared to the previous year. In contrast, Stable Users experienced more consistent achievement and growth trends across the two years.

Figure 1. Math test scores by term in 2023–24 compared to 2022–23 separately for Stable Users (left panel) and Transition Users (right panel)



Note. Test-score patterns are denoted with solid lines for 2022–23 and dashed lines for 2023–24. Standardized mean differences between the two years are displayed for each test season, with positive values indicating that test scores were higher in 2023–24 relative to the preceding year.

To supplement and provide context for Figure 1, Table 1 shows test-score differences between 2023–24 and 2022–23 for Transition Users in unstandardized RIT points. This table shows the magnitude of differences that partners can anticipate seeing in their own data. Fall test-scores differences are minor, averaging nearly -0.5 RIT points. Winter math test-score differences show more consistent, modest increases across grades, averaging about a 1.5 RIT point increase. Spring test-scores differences are larger, averaging a 3.3 RIT point increase.

Table 1. Transition Users' math score differences between 2023–24 and 2022–23 as unstandardized RIT points and standardized effect sizes

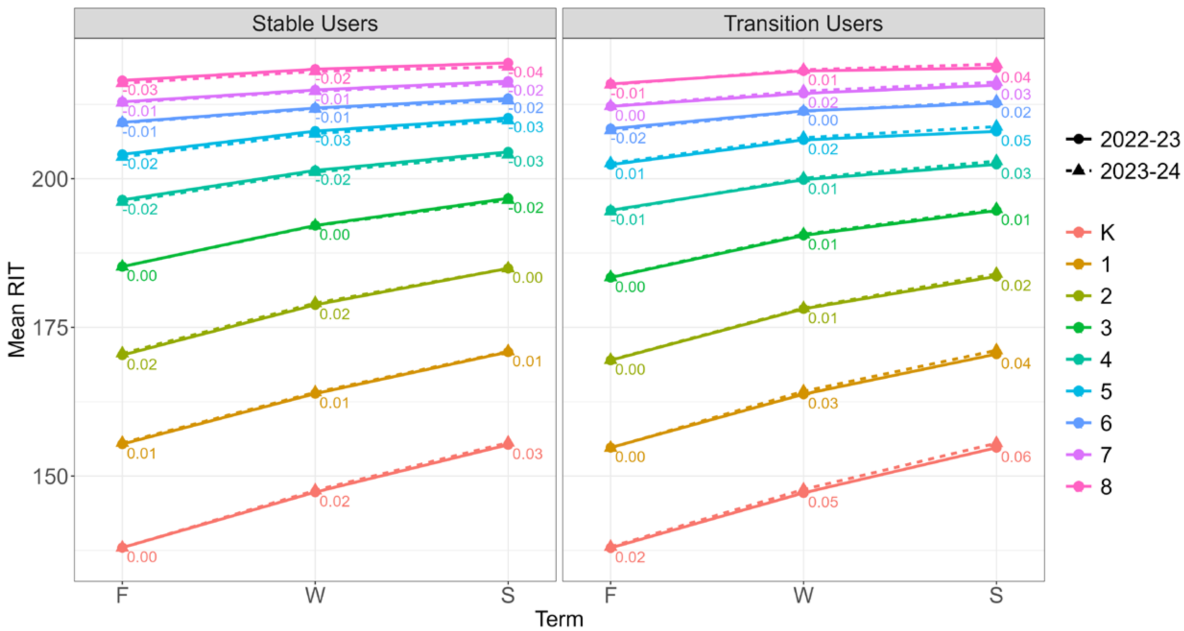
	FALL		WINTER		SPRING	
Grade	**RIT points**	**Effect size**	**RIT points**	**Effect size**	**RIT points**	**Effect size**
K	-0.22	-0.02	0.99	0.08	2.05	0.16
1	-0.64	-0.05	1.12	0.09	1.52	0.11
2	-0.72	-0.05	1.23	0.09	2.14	0.14
3	-1.31	-0.09	1.31	0.09	4.01	0.26
4	-0.37	-0.02	1.91	0.12	5.67	0.34
5	-0.38	-0.02	1.77	0.10	3.61	0.20
6	-1.38	-0.09	0.89	0.05	3.43	0.19
7	0.51	0.03	2.02	0.11	3.10	0.16
8	0.57	0.03	2.55	0.14	4.28	0.22
Avg	**-0.44**	**-0.03**	**1.53**	**0.10**	**3.31**	**0.20**

Comparing the impacts of MAP with EISA between fall, winter, and spring provides insights into its seasonal influence on test scores. We see that the impact of MAP with EISA becomes more pronounced as the school year progresses, likely due to students' increased exposure to grade-level content and improved performance as they become more familiar with grade-level material.

Score comparability in other subjects

We also assessed the comparability of test scores in the other subjects assessed by MAP Growth. Consistent with the pilot study, we found no significant changes in reading scores for partners that transitioned to MAP with EISA in 2023–24 (see Figure 2).[4] Similarly, we found no discernible impact of MAP with EISA on achievement and growth patterns in either science or language usage (not shown but figures available upon request).

Figure 2. Reading test scores by term in 2023–24 compared to 2022–23 separately for Stable Users (left panel) and Transition Users (right panel)



Note. Test-score patterns are denoted with solid lines for 2022–23 and dashed lines for 2023–24. Standardized mean differences between the two years are displayed for each test season, with positive values indicating that test scores were higher in 2023–24 relative to the preceding year.

The significance of grade-level alignment for assessing math achievement

A grade-level focus is important for all subjects, but the impact of this emphasis on student scores varies depending on the content and standards of each subject. Math instruction is characterized by a more structured and sequential skill development process relative to other subjects. According to grade-level standards in math, entire concepts may not be introduced until later grades, such as learning about negative numbers in sixth grade. In contrast, reading instruction is less likely to have comparable restriction of access to "above-grade" content, which may allow for a more fluid development process. As a result, the grade-level content focus of MAP Growth with EISA likely enhances its sensitivity to detecting the impact of instruction on a student's math achievement compared to other subjects. While this hypothesis warrants further investigation, initial observations suggest that the structured nature of math instruction may lead to more discernible effects on student scores under the enhanced algorithm.

4 While we saw no significant differences for reading in fall and winter, spring reading test scores show small differences, particularly for the upper grades. Transition Users' scores are slightly higher in 2023–24 than the prior year, while Stable Users' are slightly lower in 2023–24. However, these minor differences do not translate to noticeable changes in either spring percentiles or fall-to-spring conditional growth percentiles. Therefore, a spring concordance table for reading is deemed unnecessary.

Implications for the application of norms

We observe small to moderate changes in math scores within a test season. However, even small test-score changes result in significant normative shifts. For our partners, this likely means notable shifts in student and school achievement and growth percentiles.

Table 2 summarizes achievement and growth percentiles for Transition Users before and after the transition to MAP with EISA. Shading conveys the magnitude of the differences in percentiles across years. Consistent with Figure 1, fall achievement percentiles in 2023–24 show a slight decrease across most grades relative to 2022–23. Conversely, 2023–24 winter and spring achievement percentiles show increases relative to 2022–23. The winter differences are relatively small, ranging from 0 to 7 percentile points, while the spring percentile differences are more substantial, ranging from 3 to 13 percentile points.

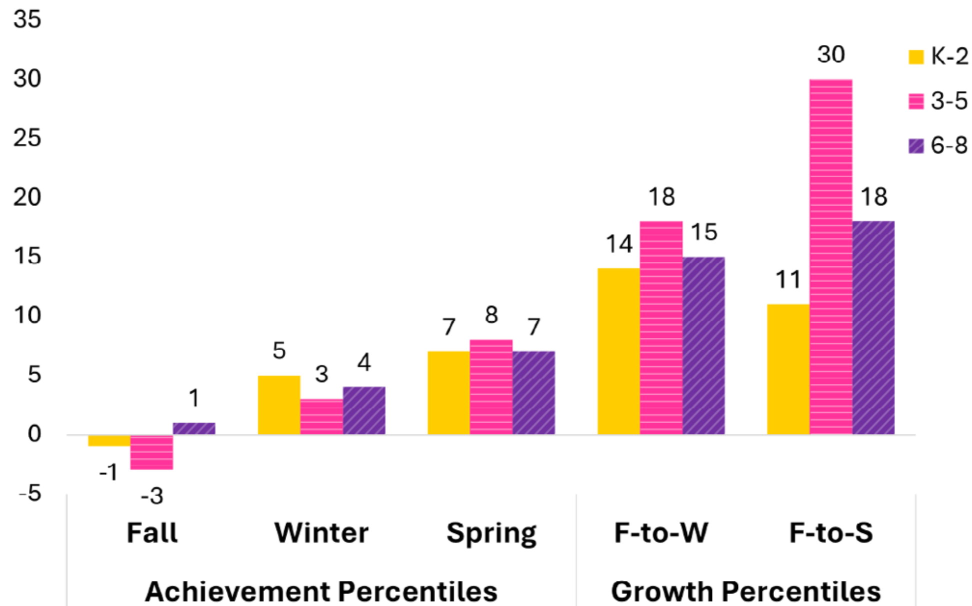
Given the combination of lower fall percentiles and higher winter and spring percentiles, conditional growth percentiles (CGPs) indicate a more pronounced shift. Increases in fall-to-winter CGPs range from 10 to 21 points, while increases in fall-to-spring CGPs range from 11 to 36 points.

Table 2. Shifts in achievement and growth percentiles for Transition Users

	ACHIEVEMENT PERCENTILES									CONDITIONAL GROWTH PERCENTILES					
	FALL			WINTER			SPRING			FALL-TO-WINTER			FALL-TO-SPRING		
GR	2022–23	2023–24	Diff	2022–23	2023–24	Diff	2022–23	2023–24	Diff	2022–23	2023–24	Diff	2022–23	2023–24	Diff
K	58	58	0	59	66	6	63	70	7	49	59	10	55	66	11
1	53	52	-1	51	56	5	55	58	3	44	58	14	50	61	11
2	50	47	-3	50	53	3	52	60	8	43	59	16	50	66	16
3	47	43	-5	47	49	2	50	58	8	46	67	21	50	80	30
4	48	46	-3	45	49	4	46	60	13	40	59	18	43	78	36
5	44	42	-2	42	45	3	43	51	8	40	58	17	40	62	23
6	44	37	-7	44	44	0	43	50	7	46	64	18	47	73	26
7	44	45	1	43	48	4	44	51	6	45	57	13	45	59	14
8	44	45	1	44	49	6	43	53	10	48	63	15	47	65	18

Figure 3 summarizes percentile shifts separately by grade bands. The increases in conditional growth percentiles are evident across all grades, but there are notable differences in the magnitude of these shifts. Specifically, grades 3–5 exhibit larger declines in fall and greater increases in spring, resulting in particularly large fall-to-spring CGP shifts. In contrast, the shifts for grades K–2 and 6–8 are smaller, though still substantial. This result suggests elementary schools that test upper elementary students more heavily are likely to see higher CGP shifts compared to middle schools.

Figure 3. Shifts in achievement and growth percentiles for Transition Users separately by grade band



Note. Bars represent the median shifts in achievement and growth percentiles within grade bands.

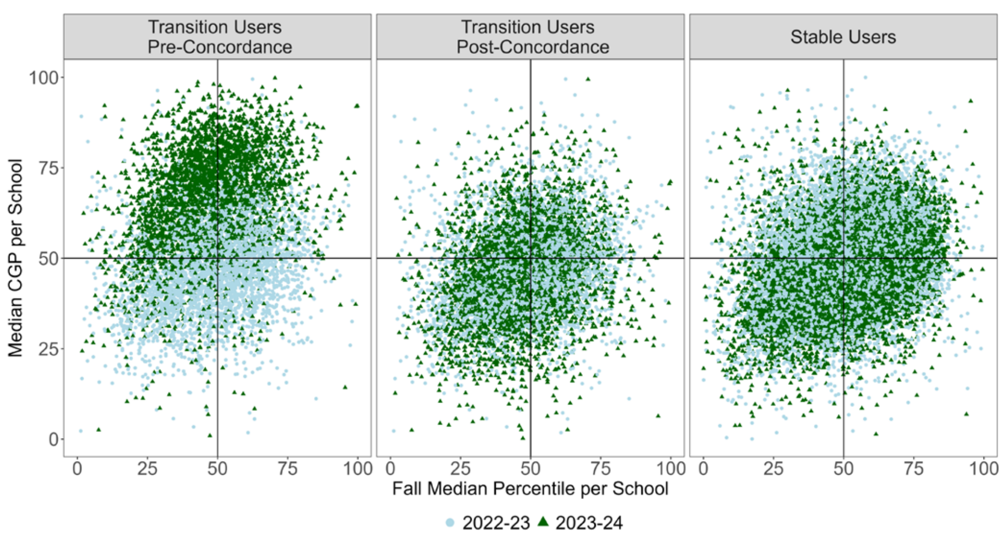
Concordance study to address math score differences

In response to the math score differences and changes in rates of growth we observe, we conducted a concordance study. The assessment industry commonly uses this established method to ensure comparability between two assessments or a test that has undergone design changes. For instance, the College Board has employed concordance studies to establish the equivalence between digital and paper-based versions of the SAT (College Board, 2024).

Our concordance study uses the equipercentile linking method to align math scores from traditional MAP Growth with those generated by MAP with EISA. This linking establishes a correspondence between math scores derived from the traditional and updated assessment using matched samples of students. By doing so, math scores from the updated assessment can be adjusted as needed to support accurate application of the norms and comparison to the traditional assessment.[5]

The quadrant plots in Figure 4 show the efficacy of this approach. Each panel depicts the relationship between school median fall achievement percentiles on the horizontal axis and school median fall-to-spring CGPs on the vertical axis. We use green triangles (2023–24) and blue circles (2022–23) to distinguish between school years. Data for Stable Users (rightmost panel) show the anticipated relationship: the cluster of green triangles aligns closely with the cluster of blue circles. In contrast, the data for Transition Users (leftmost panel) show the green cluster is noticeably shifted upward relative to the blue cluster. Relative to the same starting point in the fall, schools have higher CGPs in 2023–24 than in 2022–23. Our concordance of scores attenuates this difference. The middle panel shows that after applying score adjustments, the two clusters overlap, and the data are consistent with the historical pattern.

Figure 4. Quadrant plots of CGPs and fall achievement percentiles separately by year and user group



Note. Each point represents a school's median achievement percentile in fall on the horizontal axis and median fall-to-spring CGP on the vertical axis.

5 Because we observed comparability of scores under the updated assessment in other subjects, adjusting scores based on the concordance study is only necessary for math tests.

Applying the concordance table

Changes in the rate of growth in math relative to the 2020 norms mean that these norms are less applicable to math scores derived from MAP with EISA and will provide less meaningful context for student achievement and growth. Therefore, we strongly recommend that districts leverage the concordance table to support decision-making that is consistent with historical decisions based on the 2020 norms. The concordance tables should be used until new norms are released in 2025. There are two categories of use cases for applying the concordance table: adjusting practice and adjusting data.

Adjusting practice. This category of use cases involves adjusting cut scores, such as those used for program-placement decisions. Given the increases in winter and spring scores, using unadjusted cut scores results in underidentifying students for intervention services and overidentifying students for talented and gifted programs.[6] By using the concordance study to adjust these cut scores in math, partners can make decisions about student placement that are more consistent with past decisions, ensuring that students receive the appropriate level of support. Similarly, the cut scores used for determining which students are on track for proficiency, based on NWEA linking studies, should also be adjusted based on the concordance table.

Adjusting data. The second use case category involves using the concordance study to adjust individual scores. **It is important to note that the concordance table cannot be applied to aggregate data or directly to growth metrics**. Instead, it must first be applied to individual scores, which can then be used to recalculate growth metrics at the student, grade, school, or other grouping levels. This ensures those metrics are not upwardly biased and is strongly recommended when normative growth metrics are used for evaluation or decision-making. Additionally, adjusting data is necessary when scores will be used in longitudinal research that spans the transition period.

Conclusion

With the benefit of data from more time points, we now have a clearer understanding of how the enhanced item-selection algorithm impacts math achievement and growth trends. MAP with EISA prioritizes on-grade items to deliver test content that is aligned with grade-level curriculum standards while also adapting off-grade when necessary to deliver items of suitable difficulty for low- and high-performing students. This prioritization of grade-level test content appears to make the test more sensitive to instruction in math, which is a subject that changes in distinct ways from one grade to the next. While growth on the new test is genuine, the shifts in students' math test scores have resulted in steeper growth patterns than those captured in the 2020 norms. Consequently, score adjustments are necessary to interpret student growth against the national sample represented in the 2020 norms.

The credibility of our norms remains intact, and with appropriate adjustments to align scores with traditional MAP Growth, the 2020 norms can be validly applied. Our concordance study, along with accompanying tools, serves as the appropriate interim solution to address these shifts until NWEA releases new norms in 2025 that incorporate data from the updated assessment. To aid partners during this transition, we've developed a comprehensive suite of resources that can be accessed at NWEA.org/mapeisa.

6 The reverse would be true in fall given scores are lower in that season.

Key terms

Percentile: A statistical measure, expressed as a percentage, that indicates the relative performance of a student or group of students compared to a national reference group of students derived from the NWEA 2020 norms. *Achievement percentiles* indicate normative performance for a test-score in a specific term. *Conditional growth percentiles* indicate normative performance for test score growth between different terms. Achievement and growth are compared to same-grade students in the norming sample at a comparable stage of the school year.

Concordance study: A statistical method for linking the scores of two different assessments or different versions of the same assessment to ensure comparability between scores (see Dorans, 2004). This process adjusts scores from one assessment to make them comparable to scores from another assessment.

MAP Growth with the enhanced item-selection algorithm (EISA): The updated version of the MAP Growth assessment that uses an improved algorithm to select test items for a student that are more closely aligned with the student's grade level.

Transition Users: A term used in this study to denote partners who used traditional MAP Growth in 2022–23 and transitioned to MAP Growth with EISA in 2023–24.

Pilot study: An investigation conducted in spring of the 2021–22 school year with a small sample of volunteer partners to evaluate MAP Growth with EISA during live testing, study the comparability of scores with traditional MAP Growth assessments, and produce evidence of test-content validity and score reliability. Available for download here: https://www.nwea.org/uploads/Content-Proximity-Project-and-Pilot-Study-Spring-2022-Research-Report.pdf

RIT score: A student's overall scaled MAP Growth score for a given subject (see the MAP Growth Technical Report for more details).

Stable Users: A term used in this study to denote partners who consistently used traditional MAP Growth in both 2023–24 and 2022–23.

Standardized effect size: A statistical measure that quantifies the magnitude of the difference between two groups by expressing it in standard deviation units. Standardized effect sizes allow for the comparison of effect magnitudes.

Appendix Table 1. Unique number of students and schools in math sample by user group

	TRANSITION USERS		STABLE USERS	
Grade	Student count	School count	Student count	School count
K	82,042	1,578	268,264	4,294
1	114,153	2,040	393,993	5,760
2	130,372	2,225	445,672	6,247
3	122,733	2,177	433,805	6,126
4	124,546	2,167	437,458	6,091
5	124,014	2,084	435,225	5,821
6	114,596	1,426	394,242	3,383
7	111,234	1,200	382,300	2,914
8	105,034	1,171	325,267	2,793
9	39,401	509	85,118	877
10	36,025	503	66,023	832
11	22,735	441	35,276	719
12	9,962	250	13,237	427

ABOUT THE AUTHORS

Dr. Karyn Lewis is Director of Research and Policy Partnerships at NWEA, where she leads a team of researchers who operate at the intersection of K–12 education research, practice, and policy. Her research interests focus on the interplay between students' academic achievement and growth, their social-emotional development and well-being, and how they experience their school's climate. Prior to joining NWEA, she was a senior researcher at Education Northwest/REL Northwest, where she led a diverse portfolio of applied research, technical assistance, and evaluation projects centered around socialemotional learning. Dr. Lewis is a former data fellow with the Strategic Data Project at the Harvard Center for Education Policy Research. She completed a National Science Foundation funded postdoctoral fellowship at the University of Colorado Boulder and earned a PhD from the University of Oregon in social psychology.

Dr. Megan Kuhfeld is Director of Growth Modeling and Data Analytics at NWEA. Her research seeks to understand students' trajectories of academic and socialemotional learning and the school and neighborhood influences that promote optimal growth. Dr. Kuhfeld completed a doctorate in quantitative methods in education and a master's degree in statistics from the University of California, Los Angeles (UCLA).

About NWEA

For more than 40 years, NWEA has been a pioneer in educational research and assessment methodology with a focus on improving learning outcomes for every student. NWEA continues this discovery through dedicated research that explores foundational issues in education, practical challenges in today's schools, and the evolving role of technology in the lives of students. As a mission-based educational research organization, NWEA's research agenda reflects our commitment to attacking big challenges in education and measurement and empowering education stakeholders with actionable insights.

JUN24 | WELTSK7503