MAP Growth theory of action

March 2022

J. Patrick Meyer & Michael Dahlin

The MAP® Growth™ theory of action describes key features of MAP Growth and its position in a comprehensive assessment system. The basic premise of the theory of action is that all students learn when MAP Growth is situated in a comprehensive assessment system and used for its intended purposes to yield information about student learning and enable educators to make data-informed decisions about curriculum and instruction. The infographic illustrates the components of the MAP Growth theory of action, shedding light on the connectedness of the parts and the claims that are central to the validity argument. Evidence for the claims is described in the MAP Growth technical report (NWEA®, 2019) and other research reports.

Comprehensive assessment system

A comprehensive assessment system involves multiple forms of assessment that each emphasize a different purpose and use (Perie et al., 2009). Summative, interim, and formative assessments complement each other and provide the appropriate data for decision-making at every level of an education agency. MAP Growth is a component of a comprehensive assessment system.

A key outcome in education is for students to demonstrate proficiency with the curriculum established for the students' grades. Summative assessment is a measure of this outcome. It involves a sample of tasks from a broadly defined curriculum and a determination of a student's level of proficiency. Summative assessments are tightly aligned to standards, and test administration follows highly standardized procedures. They are typically given at the end of the school year, and results are usually available after instruction has completed. Consequently, it has very limited utility for measuring student progress during the year or informing instructional decisions. MAP Growth is not a summative assessment, but it may be used to predict performance on a summative assessment, allowing educators to identify students at risk of not meeting state proficiency standards by the end of the year.

MAP Growth is a vertically scaled interim assessment that provides information about student achievement during a particular academic term and changes in student achievement (i.e., growth) over multiple terms. The comparability of scores over multiple administrations is a strength of MAP Growth. Comparability is achieved through a test design that incorporates grade-level standards and cross-grade adaptability to meet students where they are in their learning and precisely measure their achievement and growth. MAP Growth is administered up to four times per year (fall, winter, spring, and summer). The frequency of the assessment and comparability of scores is useful for measuring growth, setting goals for students, planning instruction for a term, and program evaluation. It is less useful for providing more frequent feedback or a daily tailoring of instruction to student needs.

Formative assessments are frequent instructional practices aligned to specific lesson plans or instructional units that facilitate the tailoring of instruction on a regular basis, perhaps even daily, and adapting instruction to student needs. MAP Growth is not a formative assessment, but NWEA supports formative instructional practices through professional learning in responsive teaching and learning (Nordengren, 2021). In our professional learning, teachers learn to create and use high-quality formative assessments to support their day-to-day instruction. Because formative assessment is not designed or intended to measure student growth over time, teachers also learn to combine the formative assessment information with MAP Growth scores to make data-informed decisions about instruction and student growth.

MAP Growth theory of action

MAP Growth measures achievement and growth in grades K–12 in reading and mathematics and grades 2–12 in language usage and science. The research-backed assessment provides teachers with accurate and actionable evidence to help target instruction for students working on, above, or below grade level.

The MAP Growth theory of action illustrates the how and why of MAP Growth—from test development to leveraging results. With every test, for every use, our goal is clear: work alongside educators to help all students learn.

PROFESSIONAL LEARNING:
Responsive teaching and learning

Test users

Standards
Curriculum and instruction
Assessment

Aligned system of instruction and assessment to help monitor student growth and progress in grades K–12 in reading and mathematics and grades 2–12 in language usage and science

map GROWTH

NWEA

Input: standards and research
Focus: fair and equitable assessment
Tools: adaptivity, machine scoring

Adaptive assessments aligned to state-specific content standards that efficiently provide data on a vertical scale (within and across grades) for informing instruction and program evaluation

PROFESSIONAL LEARNING:
MAP Growth basics

Students take the grade-appropriate MAP Growth assessment with immediate results

Results

Estimates of student achievement (RIT scores)

Measures of growth (change in observed achievement across time)

Achievement status and growth norms

Predicted proficiency on the state summative assessment

Conversions of scores to other well-known metrics (e.g., Lexile®, Quantile)

College and career readiness predictors (ACT®, SAT®)

Test engagement metrics

Instructional resources

MAP® Accelerator™

College Explorer tools

Instructional content providers

PROFESSIONAL LEARNING:
Applying reports

Intended use of results

Student & Classroom (instructional planning)	School & District (school improvement)	Education & Policy Research	Special Use Cases
Differentiated instruction	Identifying schools on the basis of growth and achievement characteristics of students (e.g., premium reports)	NWEA internal research	Universal screening
Instructional planning for a term		External research	Read by Grade 3
Customized growth goals for students	Monitoring the testing process via operational reports (e.g., how much time students are spending on test events)		
	Design and evaluation of curriculum and instruction		
	Monitoring for equitable impact of education for all students		

USES THAT REQUIRE DISTRICT RESEARCH AND EVIDENCE FOR EFFICACY AND INCLUSION

Gifted and talented identification

Course placement

IEP/504 evaluation and planning

PROFESSIONAL LEARNING:
Informing instruction, focusing on growth

All students learn.

Learn more about MAP Growth and other solutions and tools in the MAP Suite of assessments on NWEA.org.

MAP Growth assessment design

MAP Growth tests are aligned to state curriculum standards in a way that supports the cross-grade vertical scaling of the assessment. Instructional areas are established according to standards to reflect the articulation of content across grades. Tests are also grade banded (e.g., K–2, 2–5, 6+) to ensure that item format and test content are suitable for students from different grade levels. Once the test design is established, items are selected from a large item bank to create a test-specific item pool that contains items aligned to the curriculum standards. Each item pool includes multiple item types such as basic multiple-choice items, common stimulus, and more sophisticated technology-enhanced items. Item pools are deep enough to implement longitudinal exposure constraints and make certain that a student never sees the same item more than once during a long period of time. Every test must pass a computer adaptive test simulation of eight consecutive test administrations, which ensures the adequacy of the item pool to support the intended test design and satisfy performance metrics.

MAP Growth scores are reported on the Rasch Unit (RIT) scale that ranges from about 100 to 350 with a mean of 200 and a standard deviation of 10. The RIT scale has two characteristics: it is a vertical scale with equal-interval units. The vertical scale characteristic means that the scale applies to all terms within a grade (i.e., fall, winter, spring, and summer) and across all grades K–12. Scores in the same subject area from two or more time points are comparable because they are on the same scale. For example, grade 3 mathematics scores in the fall may be compared to grade 4 mathematics scores in the spring. The equal-interval characteristic of the RIT scale allows the measurement of growth over two or more time points (i.e., learning) by calculating the difference in scores from the same vertical scale. Thus, the vertical scale and its equal-interval units are central to measuring growth using the change in scores across multiple time points.

MAP Growth efficiently and precisely measures student achievement in mathematics, reading, language usage, and science. Spanish test forms are available in reading and mathematics. The adaptive nature of the assessment produces better measurement precision throughout the range of the scale than a traditional fixed-form assessment of the same length. A fixed-form test could only obtain precision to a similar extent with a longer and more time-consuming test. By design, the efficient test length of MAP Growth yields more time for other learning activities than a fixed-form test. The specific level of measurement error is reported for each student.

In-school and remote test administration are possible. Both are handled by a trained test proctor who is available to assist students with technical challenges and other issues that may arise during testing. A proctor is in the room with students during in-school administrations, whereas remote proctoring is facilitated with a video, chat, email, and other communication channels. Non-adaptive practice tests are available online to familiarize students with the types of questions and item types used in MAP Growth.

Target population and test fairness

Three aspects of fairness in testing are (a) fairness in treatment during the testing process, (b) fairness as a lack of measurement bias, and (c) fairness in access to the measured construct (AERA et al., 2014). Some of the ways MAP Growth instantiates these characteristics include item and test design, statistical analysis, bias and sensitivity review, accommodations, and eliminating threats to validity.

MAP Growth is designed to be fair and equitable for all K–12 students. The adaptive nature of MAP Growth tests meets students where they are in their achievement and selects items of a suitable difficulty. Students get about half of the items correct, regardless of their ability. This feature maximizes the information the test provides about student achievement, providing insight into what a student knows and does not know. Due to its adaptive nature, everyone experiences a test of

similar length and has a comparably small amount of measurement error.

A second way that MAP Growth attains fair and equitable measurement is through development and analysis procedures. Test development procedures include reviews of item content to ensure that items reflect diversity and inclusion and do not include information that would differentially affect performance of students because of their race, sex, or cultural background. Item development review is followed by statistical analysis techniques specifically designed to detect unfair items. Any item that shows statistical signs of unfairness is reviewed for bias by an external diversity panel, and the item may be removed from the pool.

Universal design and test accommodations are incorporated into MAP Growth to make the test accessible for students. Some embedded universal design features include volume amplification, a line reader, and zoom. Non-embedded features such as a Spanish dictionary are also available. Designated features assigned to students by trained teachers include text-to-speech, color contrast, human reader, native language translation, and separate testing setting, among many others. A variety of accommodations are available to students with an Individualized Education Plan (IEP) or 504 accommodation. These include assistive technology, extended time, screen reader, and refreshable braille.

Students motivated to perform on tests tend to obtain better test scores than students not engaged in the test (Wise & DeMars, 2005). Proctors facilitate test sessions to cultivate student engagement. In addition, MAP Growth includes functionality for encouraging positive test-taking behavior. Specifically, the system can detect a lack of test-taking engagement in real time. It will attempt to re-engage unmotivated students through messages and proctor intervention. The purpose is to help students give their best effort and obtain a score that more accurately reflects their achievement in a subject.

Test security procedures not only address the integrity of the test itself and the test items, but also the privacy of student data. Role-based access, data encryption, and system redundancy with disparate geographic locations are among the tools used for test security. During test administration, a lockdown browser may be used to inhibit students from obtaining unauthorized help in answering test items. Proctor monitoring of the test session also encourages students to put forth an independent effort on the test.

Intended uses of results

The intended uses of results in a comprehensive assessment system may be classified as instructional, predictive, or evaluative (Perie et al., 2009). The particular use depends, in part, on who is using the information. MAP Growth reports provide information for decision-making at multiple levels of an education agency, ranging from the individual student to the state department of education. Individual student scores may be aggregated to the school, district, or state level to summarize performance for larger groups of students. The level of aggregation should match the level of decision-making. Classroom decisions should use aggregates of individual student scores. School-level decisions should involve classroom summary statistics. District-level decisions should involve summary statistics at the school level.

A variety of MAP Growth reports are available, with each serving a different purpose. MAP Growth reports also provide aggregate score summaries at multiple levels of decision-making. The MAP Growth reports portfolio summarizes each MAP Growth report as a guide for users. In addition, premium reports provide a fine-grained level of detail such as information about student performance on grade-level standards.

Instructional uses of MAP Growth scores typically happen in the classroom by teachers and students. For individual students, scores

provide information about their achievement at a particular point in time and the extent of their learning since the last test administration. It is a direct way for students to receive feedback on their progress in school, reflect on learning, and set goals for the next term. Teachers may use the data in a similar way to monitor student learning and set goals for their students. At the classroom level, teachers may use MAP Growth scores to reflect on their teaching practices for the term and plan for the next term. When combined with other data sources such as formative assessment, the information supports differentiated instruction and tailoring learning activities for students.

NWEA provides instructional connections that teachers may incorporate into their lessons. These include partnerships with instructional content providers, including a partnership with Khan Academy called MAP® Accelerator™. These resources typically use MAP Growth scores to place students in learning content and curricula offered by other companies. The wide availability of these instructional connections is another way that MAP Growth encourages fairness in access to the measured construct.

Many of the instructional uses of MAP Growth are supported by NWEA professional learning opportunities, as shown in Table 1. Modules cover formative assessment practices that are independent of MAP Growth and are based on principles of assessment literacy. Other modules involve MAP Growth–specific training on a variety of topics including the interpretation of score reports and setting growth goals.

TABLE 1. NWEA PROFESSIONAL LEARNING MODULES PERTAINING TO MAP GROWTH

MODULE	DESCRIPTION
Responsive Teaching and Learning	A series of modules focused on building assessment-empowered classrooms that engage and ignite student learning. These are not specific to MAP Growth or any other NWEA assessment.
MAP Growth: Basics	Get ready for a successful implementation, gain a solid understanding of what makes MAP Growth unique, and dive into resources that can help inform the process moving forward. Learn how to administer the assessments for the first time and begin to understand the importance of engaging students and optimizing learning.
MAP Growth: Applying Reports	Learn to access, interpret, and apply rich data. Then plan how to use your data to inform ongoing work, with a particular focus on flexible instructional grouping and goal setting with students.
MAP Growth: Informing Instruction	Dive deeper into classroom applications of MAP Growth. Support differentiated instruction and meet the needs of students through responsive instructional plans based on your MAP Growth results. Tailor your learning even further with learning centers focused on topics such as using formative assessment and integrating Khan Academy resources.
MAP Growth: Focusing on Growth	Learn how to apply MAP Growth data in goal setting and data conversations to improve student learning. Experiment with using a variety of data conversation tools and protocols, identifying strengths and areas for growth, and looking for patterns at various levels—including student, classroom, grade, school, and district.

MAP Growth reports provide norm- and criterion-referenced information to give context to scores and facilitate score interpretation. MAP Growth norms provide a way to compare student performance to a nationally representative sample of test takers for the fall, winter, and spring terms (Thum & Kuhfeld, 2020). Achievement status norms provide percentiles for scores at a particular point in time for each grade level. It is a snapshot of student performance. Growth norms describe a student's change in test scores relative to a national population. In particular, the Conditional Growth Index (CGI) and Conditional Growth Percentile (CGP) are two growth measures that describe a student's performance compared to similar students who had similar amounts of instructional time and prior test scores.

MAP Growth norms use a multilevel growth model that has individual students nested within schools. The model-based norms have different interpretations for individuals and schools. Individual norms allow for a student's achievement status and growth to be compared to similar students, whereas school norms allow for mean scores from one school to be compared to average scores for similar schools. The variability of scores at the individual student level and the variability of school means is a key distinction between student- and school-level norms. The variability of growth for a group of students (i.e., a school) is smaller than the variability in growth for individual students. As such, percentiles in these two cases will be different because each represents a different level of the model. School percentiles should be obtained from school norms and not individual norms.

Criterion-referenced information comes in the form of predictions for summative assessments and connections to external measures. NWEA linking studies connect MAP Growth scores to state summative assessments (Hu, 2021). Each study is detailed in a report that is available online. Linking study results are combined with the statistical model used in the MAP Growth norms to predict proficiency on a state assessment using MAP Growth scores from fall, winter, or spring. The information enables teachers to determine if a student is on track for proficiency by the end of the school year. The prediction is additional information a teacher may use to plan instruction and meet learning objectives. It is also used for implementation of state policies such as Read by Grade 3 legislation.

Linking studies also connect MAP Growth scores to other measures, with each study serving a different purpose and use of test scores. One study linked MAP Growth to a national composite of state summative test scores to support the use of MAP Growth as a universal screener in reading and mathematics (He & Meyer, 2021). As another example, score reports provide information on MetaMetrics' Lexile® and Quantile® measures, which opens vast reading and mathematics resources to families and educators. NWEA college readiness benchmark information uses MAP Growth scores to predict future performance on the ACT® and SAT® (Thum, 2017; Thum & Matta, 2015). The link between MAP Growth and ACT/SAT scores drives the NWEA College Explorer tool and helps students plan for higher education.

Finally, school administrators, district personnel, state policymakers, and researchers may be interested in evaluating the efficacy of a curriculum, policy, or intervention. MAP Growth scores may serve as a common outcome for these evaluations. For example, MAP Growth data have been used to research summer learning loss and achievement disparities (Atteberry & McEachin, 2021; McNeish & Dumas, 2021), test score gaps between advantaged and disadvantaged students (von Hippel & Hamrock, 2019), the potential disadvantages schools may impose on Asian American students (Yoon & Merry, 2017), and personalized learning (Pane et al., 2015).

References

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological measurement*. AERA.

Atteberry, A., & McEachin, A. (2021). School's out: The role of summers in understanding achievement disparities. *American Educational Research Journal, 58*(2), 239–282. https://doi.org/10.3102/0002831220937285

He, W., & Meyer, J. P. (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

Hu, A. H. (2021). *MAP Growth linking studies: Intended uses, methodology, and recent studies*. NWEA. https://www.nwea.org/content/uploads/2022/01/MAP-Growth-Linking-Studies_Uses-Methodology-Recent-Studies-2021-11-15.pdf

McNeish, D., & Dumas, D. (2021). A seasonal dynamic measurement model for summer learning loss. *Journal of the Royal Statistical Society, Series A, 184*(2), 616–642. https://doi.org/10.1111/rssa.12634

Nordengren, C. (2021). *Focusing squarely on students: A theory of change for NWEA professional learning*. NWEA. https://www.nwea.org/resource-center/white-paper/29610/Focusing-Squarely-on-Students-A-Theory-of-Change-for-NWEA-Professional-Learning_NWEA_White-paper.pdf/

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

Pane, J. F., Steiner, E. D., Baird, M. D., & Hamilton, L. S. (2015, November). *Continued progress: Promising evidence on personalized learning*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR1365.html

Perie, M., Marion, S., & Gong, B. (2009). Moving toward a comprehensive assessment system: A framework for considering interim assessments. *Educational Measurement: Issues and Practice, 28*, 5–13. https://doi.org/10.1111/j.1745-3992.2009.00149.x

Thum, Y. M. (2017). *MAP college readiness benchmarks: An addendum with preliminary results keyed on the SAT.* NWEA Research Report. https://www.nwea.org/content/uploads/2020/10/MAP-SAT-College-Readiness-Benchmarks_NWEA_linkingstudy.pdf

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

Thum, Y. M., & Matta, T. (2015). *MAP college readiness benchmarks: A research brief.* NWEA. https://www.nwea.org/content/uploads/2015/08/MAP-College-Readiness-Benchmark_Study-AUG15-Revised.pdf

von Hippel, P. T., & Hamrock, C. (2019). Do test score gaps grow before, during, or between the school years? Measurement artifacts and what we can know in spite of them. *Sociological Science, 6*, 43–80. https://doi.org/10.15195/v6.a3

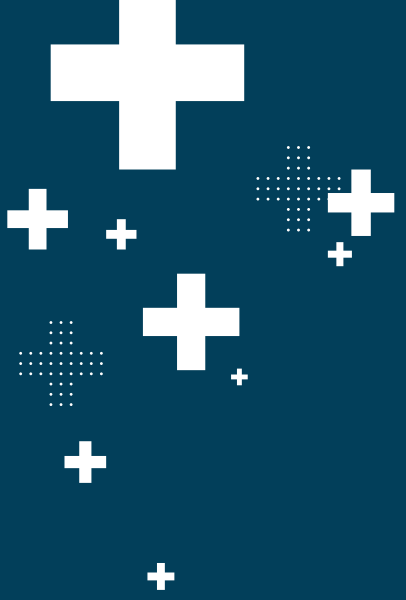
Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment, 10*(1), 1–17. https://doi.org/10.1207/s15326977ea1001_1

Yoon, A. J., & Merry, J. J. (2017). Understanding the role of schools in the Asian-white gap: A seasonal comparison approach. *Race Ethnicity and Education, 21*(5), 680–700. https://doi.org/10.1080/13613324.2017.1365053

Acknowledgments

It is with great appreciation that we recognize the many people at NWEA who contributed to this theory of action: Jacob Bruno, Ted Coe, Miah Daughtery, Laurence Dupray, Amanda Fritz, Wei He, Ann Hu, Tami Hunter, Nate Jensen, Cindy Jiban, Chase Nordengren, Kelly Rivard, Adam Withycombe, and Adam Wolfgang.

Suggested citation: Meyer, J. P., & Dahlin, M. (2022). *MAP Growth theory of action*. NWEA.



nwea

MAR22 | KAP8104