

WHITEPAPER

# Growth Matters

7 Key Criteria for Measuring Growth





What matters to you when it comes to assessment? At NWEA®, it's our favorite question to ask educators like you—and the answer is clear.

Time and time again, we hear the same response: "I need to know if my students are learning." Educators need tools that help them measure what really matters for their students' learning—and, most importantly, to measure it well. Measuring growth—not just proficiency—is key to meeting the needs of educators and kids alike. And the importance of measuring growth is finally being widely recognized in all corners of education.

Educators working directly in the classroom have long understood that each student is unique—with their own talents, challenges, and ways of learning—and that to apply the same one-size-fits-all approach to an entire population is bound to leave many kids behind. Grade-level proficiency measures might assess kids in the middle well, but they leave many students behind entirely—any kid working above or below grade level in any subject is underserved by such tests. Measuring growth is valuable because it gives teachers information that helps them support learning for every student, regardless of their proficiency status.

But measuring growth is only useful if it's done well. It's not about just comparing one proficiency score to the next one. It's about accurately identifying each student's current performance level—with the same precision for students well above or well below grade level as for those in the middle of the pack.

You deserve a growth measure that delivers valid data you can trust—and that you can use to support learning for every single student.

# 7 Features of a Great Growth Measure

We've been measuring growth for 40 years at NWEA, and have a large team of independent researchers dedicated to studying how to do it best. Our research has revealed seven essential criteria for a great growth assessment.

## A QUALITY MEASURE SHOULD:

- 01 be based on your educational standards**
- 02 use a scale with equal intervals over time**
- 03 measure a student's performance correctly, regardless of their grade**
- 04 have many possible questions, at many different levels of difficulty**
- 05 be unbiased and fair for all students**
- 06 have a clear purpose and be built to meet that purpose efficiently**
- 07 give you context for growth**

## 01 The test should be based on your educational standards

Each state has a set of educational standards that describe exactly what students should learn in every grade. Your curriculum plans are based on these standards. You teach students the content described in the standards, and you teach it in the order laid out by the standards—from early foundations for young children all the way to the most advanced ideas for teens.

So for your assessment to be useful, it needs to measure students' academic growth in gaining the knowledge described in your standards—as well as the level of understanding the student has achieved.

## 02 The test should use a scale with equal intervals over time

Imagine you measure your daughter's height at the beginning of the year and discover she's 4'3" tall. You measure her at the end of the year and discover she's 1.35 meters tall. Has she grown? It's impossible to tell by comparing those two numbers, because you used a different scale—feet and inches versus meters—each time you measured. In this case, you can translate meters into inches because there's a known relationship between those two scales. But what if you'd measured her once in ancient Egyptian cubits? There's no way to

translate that ancient measurement system into modern inches or meters.

This illustrates the importance of using the same scale over the entire time you're assessing progress. That's how you can see if someone has grown—you measure once, measure again later, and compare the results. The difference is the amount of growth. And, crucially, that measurement must take into account more than proficiency—it must truly identify a student's performance (regardless of grade level) and show their full learning progression.

In addition, you need your scale to use equal intervals over time, because that's what allows you to not just measure growth, but also see how that growth changes over time. After measuring performance multiple times, you want to know not just how much a student has grown between measurements, but details about that growth—is the student growing more slowly now than they were last year? Is she growing faster? About the same? This detail tells you important information about a student's learning progression and lets you know if you need to intervene now—before she's fallen behind.

So in order to give you a valid answer to “has this student grown, and how much?” your test should have a scale that uses the same equal intervals for every grade, from pre-kindergarten all the way through twelfth grade.

### 03 The test should measure a student's performance correctly, regardless of their grade

Let's go back to the metaphor of measuring height. Say you want to quickly and easily learn the height of every kid in a big third grade class, so you buy a new high-tech gadget that can digitally scan the group and detect their heights. But this magic device has one limitation: it can only detect heights between 4' and 4'6". Sure, most of the kids in your class will fall in that range, but the fact is, there are sure to be some extra-short kids and some especially tall kids who just don't fit in. If you go ahead and use that measuring gizmo, you won't have any clue how tall the shortest and the tallest kids are—so you certainly can't measure their growth over time.

The same is true with academic measurement. If your test is only measuring whether your third graders have mastered third-grade skills, it won't tell you anything at all about the kids who are still struggling with second- or even first-grade skills. And it sure can't give you any useful information about the high-flyers who are working at a fifth-grade level.

If you aren't measuring those students, you can't understand what they know and what they're ready to learn next. You won't know if they've hit a wall in their learning, if they've missed out on a key concept, or if they've exhausted the challenges before them and are spinning their wheels.

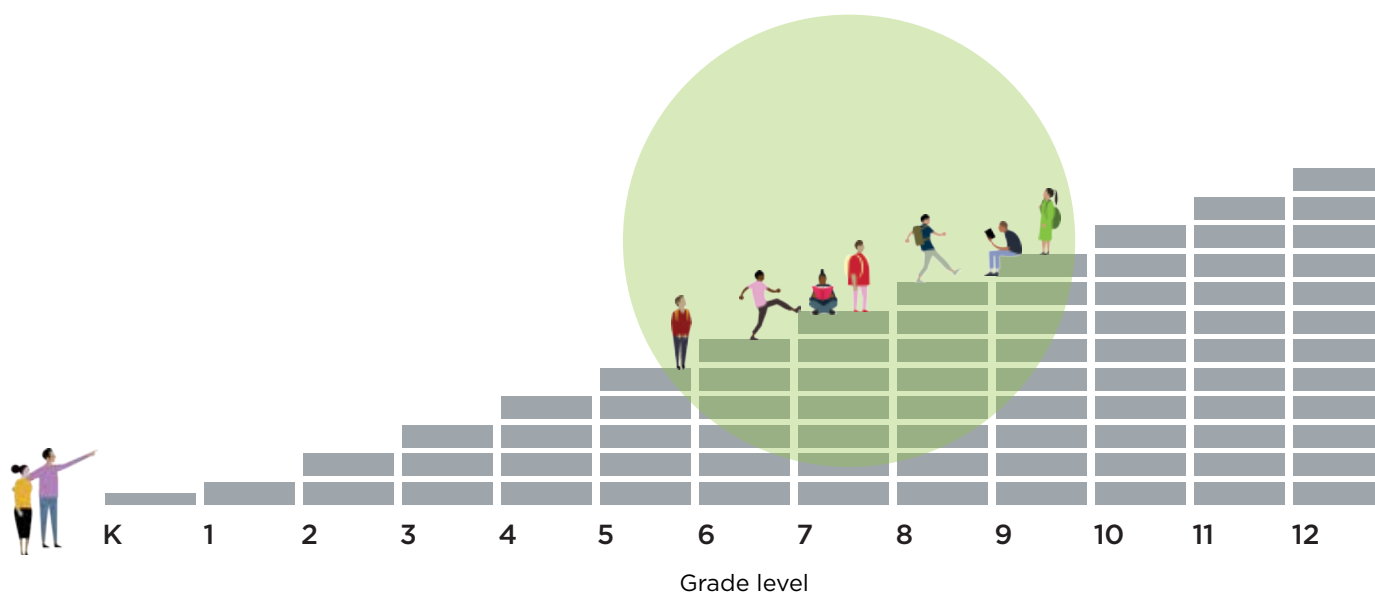
Equity in the classroom demands an assessment that accurately measures the growth of every student, regardless of the grade level they happen to be in.

### 04 The test should have many possible questions, at many different levels of difficulty

Computer adaptive assessments are ideal for measuring growth well because they can hone in—quite precisely—on what a student knows and is ready to learn. How? They dynamically adapt throughout the test in response to the student's answers. A correct answer generates a more difficult test item; an incorrect answer, an easier one. This ongoing adjustment allows the test to pinpoint exactly what the student knows.

But in order to actually work well, these assessments need a very large number of possible questions. We call this the "item pool"; it's the collection of questions the test draws from when trying to pinpoint exactly what a student knows. A deep item pool—lots of items about every possible topic, and at all possible levels of difficulty—means that the assessment can show the student enough unique questions to reveal the specifics of what they know and are ready to learn next. And they won't see the same question twice.

The other advantage of a deep item pool is it gets rid of that age-old fear: teaching to the test. If you're using an assessment with a small number of questions, educators may feel pressure to prep kids to answer just those questions well. But with a tremendous number of quality questions, teachers are free to teach the concepts their students need, knowing that the assessment will draw from a large number of questions to check students' understanding of the concepts.





## 05 The test questions should be unbiased and fair for all students

Remember the old riddle about the man and son who were rushed to the hospital after a car accident? The doctor walked into the boy's room and cried, "I cannot operate on this boy—he's my son!" The riddle asked people to say how this was possible, and they were often at a loss to come up with the answer—the doctor was the boy's mom. For a long time, many people unconsciously pictured all doctors as male—and they didn't realize their bias until it was pointed out.

The fact is, everyone has unconscious biases—it's part of human nature (albeit a part we're always striving to improve). But if those biases creep into assessment questions, fairness and equity go out the window. And creating test questions that are equally accessible to every student—regardless of cultural, socio-economic, ethnic, and religious background—requires rigorous and systemic review.

A number of organizations publish stringent standards that help test makers eliminate bias and take a consistent approach to writing unbiased items. Forgive the litany of acronyms, but this is crucial for equitable assessment: you need a growth measure that uses Differential Item Functioning (DIF), as well as bias and sensitivity reviews that follow standards set forth by The American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME).

## 06 The test should have a clear purpose and be built to meet that purpose efficiently

The most important question on any assessment is the one you, the educator, ask yourself before you give the test: "Why am I doing this?" Any test you give should have a clear purpose and be built to fulfill that specific purpose with accuracy and precision—taking exactly as much time as needed: no more, no less.

**Choose an assessment that measures exactly what's needed to give you valuable data with maximum efficiency.**

The more precise a test, the more accurate the score it gives—and the more time it takes to administer. In attempting to protect valuable instructional time, it can be tempting to reach for the shortest assessment available. But if that assessment gives you unreliable or inaccurate data, it's a waste of time all around.

So when your purpose is to measure growth—and use the information you glean to support student learning—choose an assessment that measures exactly what's needed to give you valuable data with maximum efficiency: no more and no less testing than needed for the purpose.





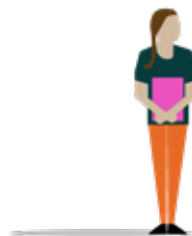
## 07 The test should give you context for growth

Once you accurately measure performance and growth, a world of instructional opportunity opens—but only if that growth is put into a useful context. Your assessment data needs context in order for educators at every level to take action based on it.

A teacher benefits from knowing what the student's score is in relation to all the other students in the classroom. A principal benefits from knowing their school's relative position within a district, and a district supervisor finds it useful to place a school's performance in the state and national context.

**Your assessment data needs context in order for educators at every level to take action based on it.**

Another context is established by looking at student growth trajectories and determining if they're on a path to meet a given achievement standard in time. For instance, is a student on track to be ready for college upon graduation? Is a school on track to meet proficiency benchmarks on state accountability measures? Knowing this context helps you make key decisions about resource allocation and instruction—at the classroom, building, and district level.



## Quality Growth Measurement Can Transform Education

Imprecise measurement can have devastating consequences—students flailing without adequate help, lost opportunities for enrichment, disengaged students, and discouraged staff. Great schools insist on the best measures to understand all their students, not just those in the middle of the bell curve.

We believe measuring growth is the key to helping students learn, which is why our flagship interim assessment—MAP® Growth™—was built specifically with this purpose, and has been continually refined

over the decades to deliver the highest quality data in the industry. And that's why educators in nearly half the districts in the US—as well in countries around the world—trust the data they get from MAP Growth to drive key decisions for their students.

But we don't stop there. The NWEA suite of assessments offers multiple quality-built measures for growth, skills mastery, academic proficiency, and school improvement. Together, they provide you a multi-faceted view of student and school performance and help you enable transformative learning for all kids.



## Growth measures you can trust

**Discover how MAP Growth can help you answer the fundamental question—  
“Are my students learning?”—as well as inform instructional decisions  
to accelerate that learning for every student.**

**EXPLORE MAP GROWTH**



NWEA is a not-for-profit organization that supports students and educators worldwide by providing assessment solutions, insightful reports, professional learning offerings, and research services. Visit [NWEA.org](https://nwea.org) to find out how NWEA can partner with you to help all kids learn.

© NWEA 2017. MAP and NWEA are registered trademarks, and MAP Growth is a trademark, of NWEA in the US and in other countries. Illustrations on pages 4 and 6 © Adam Simpson and Heart 2017.