CATALYST | MEASURING S E R I E S | GROWTH

THE CASE FOR GROWTH

Why Measure Student Learning?

A comprehensive e-book about measuring student growth.



Contents

1.	PART ONE: WHY ASSESS AT ALL?	3
	KEEP THE FOCUS ON STUDENTS: MEASURING GROWTH MAXIMIZES	
	EVERY CHILD'S CHANCE FOR ACADEMIC SUCCESS (INFOGRAPHIC)	4
2	PART TWO: WHAT IT TAKES TO MEASURE STUDENT GROWTH	
۷.	FAIRLY AND ACCURATELY.	5
3.	PART THREE: HOW GROWTH DATA CAN SUPPORT EVERY	
	STUDENT'S LEARNING NEEDS	8
4.	APPENDIX	14
	MAKE SURE YOUR STUDENT GROWTH DATA MEASURES UP	
	10 QUESTIONS TO EVALUATE STUDENT GROWTH ASSESSMENTS	14

A fundamental question—perhaps **the** fundamental question—educators ask is, "Are my students learning?"

[=]

This question reverberates throughout school systems, communities, and the nation. To answer it, educators use many tools, all of which combine to inform a teacher's understanding of student learning. Many educators choose to gain key information from assessments that measure student achievement and growth, and for good reason: just as you might use a yardstick to measure your child's physical growth, assessments can provide an objective, consistent measure of academic growth.

To help educators, parents, and policy makers drive learning for all students, this booklet shares practical insights and pragmatic learning around measuring student growth.

Part One: Why assess at all?

ANSWERING EDUCATIONAL OUESTIONS WITH DATA

Educators use assessments to answer educational questions with data. These assessments are designed for various purposes, depending on when they are given in the instructional cycle and the needs of the stakeholder who will be using the data. Assessments fall into three broad categories:

- Formative assessments provide information in the moment to help teachers adjust instruction; this process permits educators and students to collect critical information about student progress and to uncover opportunities for review, feedback, and adjustments to instruction.
- Interim assessments provide an objective measure of student achievement, progress, and growth over time; the results can be used to differentiate instruction, allocate resources, determine placement in special programs, and evaluate program effectiveness. These assessments are given at regular intervals through the school year.
- Summative assessments provide a summary of student achievement after completion of an instructional unit or course, including accountability tests at the end of a semester or school year.

MEASURING GROWTH REQUIRES MULTIPLE SNAPSHOTS

But how do you measure growth? In a time lapse film of a flower, the visual effect is created through multiple snapshots threaded together to show the plant's trajectory from bud to bloom. Student growth can be measured in a similar way, through a series of snapshots that measure achievement at a given time. Assessments, taken at different points through the year, can provide

these learning snapshots, and when viewed together, show the pattern of a student's growth.

To measure academic growth in an instructionally useful way—meaning the measure provides data that enables effective, skill-based grouping and differentiated instruction—an assessment must:

- be built on a stable scale
- measure students at, above, and below grade level parameters in order to get a true reading of achievement levels (a crucial component as it permits educators to provide effective, personalized learning paths via differentiated instruction and allows an accurate measure of growth over time)
- provide context for the data plus insight in the moment to help teachers adjust instruction

STUDENT GROWTH - AN URGENT CALL **TO ACTION**

When you consider a few facts about the state of education, the case for measuring growth is quite simple.

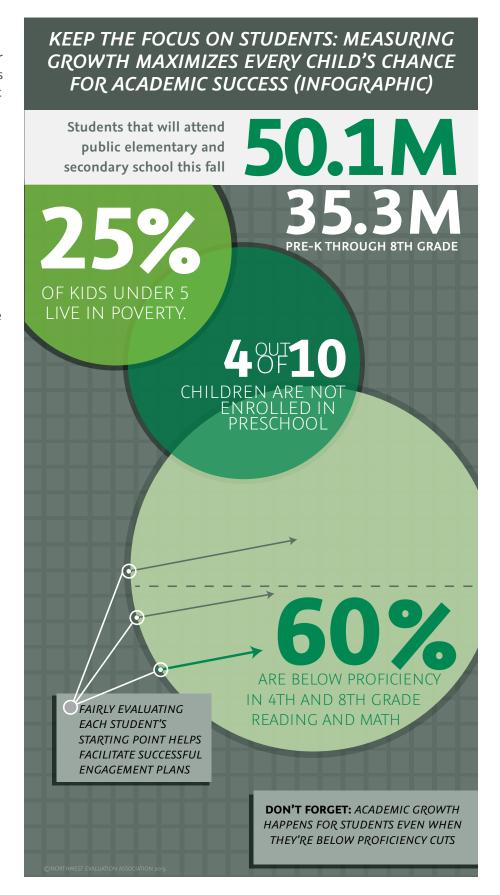
- 60 percent or more of 4th and 8th graders are not reading or doing math at grade level¹
- As we've seen from recent summative assessments in New York and Kentucky associated with the Common Core State Standards, the proportion of students scoring "below proficient" will rise as new, higher standards make proficiency harder to achieve².
- Many children live in poverty: 20.5% of 6-17 year olds and 25.9% of children 5 and under. Poverty is correlated with many challenges to student learning, including entry in school with severely reduced vocabularies, higher dropout rates, and dropping out at earlier grades³.

None of these statistics paints a flattering picture of education in America—nor do they reflect our beliefs about creating better lives for children, closing achievement gaps, ensuring future economic prosperity, or fostering civic engagement. To live up to our aspirations, we must give our teachers the tools they need to succeed.

If we know that up to 70% of students are not achieving "proficiency" on today's summative tests, and that tomorrow's will be based on the more demanding standards of the Common Core, it is imperative that teachers know the starting point and can follow the progress of each student. Then they can engage the student in charting an instructional path that moves toward success. Along the way, the student still may not be rated as proficient on summative assessments, but this does not mean that he or she is not growing academically.

In fact, often we see heroic growth. Consider the fifth grade student who starts the year reading at a second grade level. In June, his teacher rightfully celebrates his achievements because he is now reading like a fourth grader – still not "proficient," but growing rather remarkably.

Such growth must be made visible and made to count. Our students deserve recognition for their progress, and our society requires a clearer picture of what our schools accomplish every day.



Part Two: What it takes to measure student growth fairly and accurately.

Standing in front of you is a smiling child about whom you want to learn more. If you are her teacher, you may want to understand her reading level, so that you can guide her to the right books and help her progress. Or if you are the principal of her school, you might want to know how she compares to other students in the same grade. How will you obtain the information? How will you compare the results from this child to another? Measuring student learning growth well begins with accurately measuring student achievement.

However, to measure growth accurately and fairly requires thoughtful consideration of several issues, including:

- Aligning test items to content standards
- Creating a vertical scale of measurement
- Matching item difficulty level to student ability
- Gaining precision using a deep and expansive pool of items
- Ensuring fairness through empirical bias and sensitivity reviews
- Balancing the accuracy required with the need for which the data is gathered
- Providing context for growth

Let's see what's involved in each area.

ALIGNING TEST ITEMS TO CONTENT STANDARDS

When we ask, "What do you want to know?" about student academic achievement and growth, we are in the realm of content standards such as those adopted by a state or group of states (for instance, the Common Core State Standards).

Standards lay out a clear, consistent understanding of what students are expected to learn, and the sequence in which they are expected to learn it. Teachers then take these standards and develop their curriculum plans around them, often collaborating with their colleagues in adjacent grades to ensure that students gain prerequisite knowledge on which subsequent teachers can build. We use the standards to define concepts of what we want students to master, whether in reading, English language arts, mathematics, science, or other areas—all the way from beginning foundations to advanced expressions.

Assessment follows from this and is defined and bounded by the entirety of the state or Common Core standards—not just one narrow grade level within the standards. So the first requirement to measure student achievement and growth fairly and accurately is that the questions—or items, in assessment-ese—that make up a particular test must reflect the content of the standards.

However, education standards not only specify the content a student needs to know, but also the level of understanding expected for the mastery of the content of that standard. So in addition to needing to reflect the content of the standards, assessments need items that match, where possible, the level of understanding required to assure that what is being assessed reflects what is being taught.

CREATING A VERTICAL SCALE OF MEASUREMENT

So now you have assessment items that match the standards your staff is teaching. The next thing required to measure accurately and fairly is a scale that identifies the difficulty of the items. Here, an analogy may be useful.

Suppose you measure your son's height today and you learn that he's 5'9". Nine months from now you measure his height again, and you learn that he's 1.82 meters tall. Did he grow? The answer is yes (about 2.5 inches). You have a choice of tools to measure him, but both are built on a scale with a consistent unit of measure. Accurately measuring student growth and achievement requires the same tool.

There are different varieties of scales used in academic assessments. One option is to measure grade level performance, with each grade on its own scale. This creates problems in linking the data—similar to the yardstick versus meter stick to measure height. Researchers have found that such linked scales are unstable for measuring growth over time. They lack a common reference point or a consistent unit of measure. To overcome these obstacles, many assessments use a vertical scale that spans the grade levels, from Pre-K through high school, preferably with equal intervals marked along its length. With one long "measuring stick" in use, no scale linking is required, and this means that measuring growth increases accuracy and reliability.

MATCHING ITEM DIFFICULTY LEVEL TO STUDENT ABILITY

After creating a vertical scale, you will need to match item difficulty to student ability. Returning to our analogy of measuring height, this step helps determine whether we can put a mark on the wall which really is level with the top of each child's head. Here, it is important to examine some assumptions about students.

If we assume that all nine-vear-olds are between 45 and 51 inches tall and restrict our measurement ability so that we can only place marks inside these limits, we would be unable to accurately measure a large proportion of children at that age, because a child's height does not necessarily fall within this range. Some are shorter, some much taller – that's how they grow.

Academic measurement follows the same principle. Assessments which restrict items to grade level standards alone have an important role in providing information that school systems, states, and our nation need. Summative assessments given for state accountability purposes are explicitly built for this purpose.

However, when assessments are restricted in this way, we are not able to measure with precision students who are performing above or below gradelevel—and this represents many, if not most, of our students. Without exposing students to items on other parts of the scale, we are unable to accurately assess their growth and achievement. If all of the items cover basic arithmetic and the student has mastered algebra, all we really know is that he or she has mastered basic arithmetic. This has important implications for our national aspirations around creating equal opportunity for all students. To meet this objective, teachers need to locate a student's current achievement level, and plot an instructional program that enables the student to move ahead. This requires a wide range of item difficulty to be used, which leads to the next requirement.

GAINING PRECISION USING A DEEP AND EXPANSIVE POOL OF ITEMS

If the assessment presents items to students at their level of achievement, then the number of items presented and the closeness between the students' level and the difficulty level informs the level of precision you can expect from the assessment data.

If there is only a question or two at the student's level, regardless of his or her grade placement, the precision will be limited. When there are many items that are on, above, and below the student's level, then you gain a finer grain of detail on the student's current achievement. The more questions students get right and wrong, coupled with how close the questions are in difficulty to the student's actual achievement level, the more precise the results. This requires not only many questions at each difficulty level along the scale, but it requires that appropriate questions be presented to each student. Computer adaptive testing (CAT) makes the management of presenting the items manageable and scalable.

ENSURING FAIRNESS THROUGH EMPIRICAL BIAS AND SENSITIVITY REVIEWS

Congratulations: You now have a deep pool of items aligned to content standards, calibrated on a vertical scale and spanning the full range of students you want to measure. But you are not yet finished building your fair and accurate measure of growth. In fact, you are about to undertake what is perhaps the most important step—ensuring fairness through systematic, empirical reviews for bias and item performance.

Students come to school from myriad backgrounds—cultural, socio-economic, ethnic, religious, and more. In addition, not all students may have had the opportunity to learn the material to be tested, or the material may be presented in such a way as to privilege a certain background. These factors all contribute to the potential for bias in an assessment.

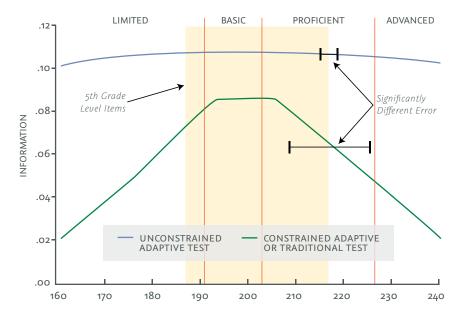
Because this issue cuts to the core of creating equitable access to educational resources, assessment developers use practices including Differential Item Functioning (DIF) and bias and sensitivity reviews to reduce bias in the instruments they create. The American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME) publish standards on this for test makers, to support this effort and to provide consistency in approach across developers.

BALANCING THE ACCURACY REQUIRED WITH THE NEED FOR WHICH THE DATA IS GATHERED

Even with carefully constructed vertical scales and a deep item pool, no single assessment can determine a student's absolute achievement. Measurement error occurs whenever anything is measured. Moving from an achievement score to a growth score increases the need for measurement precision. However, there are costs to get extra levels of precision—both financially, and in time and resources. So, what level of accuracy do you need? The answer lies in examining how the data will be used. For example, if the assessment is conducted several times a year to measure growth, then the amount of growth evidenced between two measurement intervals needs to be similar to or larger than the measurement error. If not, then the students' true growth will be masked underneath the error.

CURVE OF MEASUREMENT ERROR

COMPARING UNCONSTRAINED AND CONSTRAINED ADAPTIVE TESTS



Cross-grade computer adaptive assessments produce more precise information for students performing outside of grade level than do grade-level tests. If the adaptive assessment is designed to measure proficiency as well, it can produce more information than a fixed form test, even near the proficient cut score, because it can hone in with its item selection. However, if the test adapts only within a limited range along the scale, once beyond the range, the amount of information gathered once again drops steeply and measurement error increases. For adaptive assessments that provide access to items anywhere along the scale, the amount of information generated for all students is high, and the measurement error is low.

PROVIDING CONTEXT FOR GROWTH

Now that achievement and growth are accurately measured with limited error, a world of instructional opportunity opens—as long as there is accompanying information that provides a context.

If a teacher receives a middle school student's mathematics score without knowing whether the student knows how to add fractions, the score isn't very helpful. Rather, the teacher would benefit from knowing what the student's score is in

relation to all the other students in her classroom as well as all other students in that grade in the district, the state, and the nation as a whole.

When we accurately measure student achievement and growth, we can also compare that score to a standard. A cut score can be set and students can then be accurately compared to it. Additionally, normative data can establish context for a student's score since they provide a reference to performance for a given group. Two typical comparisons are growth compared to peers and growth trajectories. With growth, the definition of peers is more nuanced because the amount of student growth depends not only on the grade and subject, but on the starting achievement level and how much instruction they receive (i.e., more days of instruction generate more growth).

Another context is established by looking at the student's growth trajectory and determining if he or she is on a path to meet a given achievement standard in time. For instance, is a student starting at his or her initial achievement level, and growing from there much as he or she did, on a path to be ready for college when they graduate? Given the current focus on college and career readiness, this is a key question for today's educators.

Part Three: How growth data can support every student's learning needs

HELPING ALL STUDENTS LEARN

Accurate measurement of student achievement along a vertically aligned scale allows educators to accurately determine academic growth for all of their students.

Setting standards, defining expectations, and making judgments about students' levels of achievement and growth is a process that data can inform; however, these are all based on individual or collective values and should be considered separately from the requirement to measure accurately. The insight of teachers, parents,

administrators, and the students themselves are all critical to the interpretation and usefulness of the data, helping us all know if all of our students are learning.

Educators know that their students have the potential to learn and grow during the school year. They also understand that students start the school year at different levels of understanding and readiness. In order for educators to maximize the growth potential of all their students, they need accurate student growth data to inform their learning plans for each student.

Growth data can help all stakeholders answer important questions about student learning, such as:

- As a teacher, how will I know what kind of progress my students are making toward learning goals, and how can I adjust my instruction to meet their needs?
- As a school principal, how can I ensure that the students in my school are tracking toward key milestones? How can I offer the best professional development to support our teachers in their use of timely growth data?
- As a district administrator, how can I evaluate our district's programs for improvement planning? What's working best? What should we stop doing?
- As a parent, how do I know my child is progressing?
- As a student, how do I know if I am learning and what do I need to work on next to reach my goals?

Let's take a look at how the thoughtful use of assessment data can help each stakeholder answer these questions.

PLANNING MEANINGFUL INSTRUCTIONAL PATHS FOR STUDENTS USING GROWTH **DATA—SUPPORTING STUDENTS, PARENTS, AND TEACHERS**

At the beginning of the school year, teachers and students can work together to set individual growth goals. A stable, grade-independent scale allows teachers and students to gauge the growth needed to meet or exceed proficiency expectations for that school year, allowing them to set growth goals and milestones. Pairing growth data with formative assessment practice (a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes) has been demonstrated to move students even more rapidly towards their learning goals.

Measuring student progress mid-year provides teachers, students, and parents with a snapshot of learning accomplished so far, helping everyone know if the student is on track to meet the mutually established growth goals.

End-of-year growth data lets everyone know if the student reached their goals and is expected to meet or exceed the accountability goals in their state.

Growth data can help teachers identify areas to target in teaching, and enable administrators and teachers to group students flexibly based on demonstrated need.

But, what does this look like in practice?

Here's an example of what this could look like with MAP® data.

Chart title 195 Has New Meaning for Louis—Support for Students and Teachers

Mr. Waller reviews all of his 3rd grade class' growth projections for math. He notices Louis has a projected growth RIT score of 183, and based on his state's cut score ranges, Louis needs a score of 195 to meet proficiency.

In a goal-setting conference, Mr. Waller and Louis dig deeper into the data and uncover goal areas of strength and opportunity. Agreeing on a mid-year RIT goal of 190 and an end-of-year RIT goal of 195, they choose a few learning statements from DesCartes: A Continuum of Learning® to develop into specific student goals. Learning statements represent specific concepts and skills on which the student is assessed. Several of the learning statements are from the 191-200 RIT range in order to set challenging goals closer to proficiency.

Sample learning statements from the 191 – 200 RIT Range:

Because he has helped set his goals, Louis will monitor his own progress and stay engaged in his learning.

Skills and Concepts to Enhance	Skills and Concepts to Develop	Skills and Concepts to Introduce
181	191	201
Compare & Classify 2-D	Compare & Classify 2-D	Compare & Classify 2-D
and 3-D Figures	and 3-D Figures	and 3-D Figures
 Identifies and names multiple shapes (e.g., square, rectangle, triangle, circle) Classifies polygons by sides and vertices Identifies and names a cube Identifies and names a sphere 	 Identifies parallel lines Identifies points on a circle Identifies diagonals of a polygon Identifies and names a polygon Identifies and names a pentagon Identifies the number of faces on rectangular prisms Identifies and names a cylinder Identifies and names a sphere Sorts 2-D shapes and objects according to their attributes Creates a new shape by combining different shapes, or identifies the different shapes that were used to make the original shape 	 Identifies intersecting lines Identifies parallel lines Identifies right angles Identifies and names a parallelogram Identifies and names a polygon Identifies and names a hexagon Identifies and names an octagon Classifies polygons by sides and angles Classifies cubes by their properties (e.g., edges with equal lengths, faces with equal areas and congruent shapes, right angle corners) Identifies and names a cylinder Classifies cylinders by their properties (e.g., base shape, lateral surface shape, vertices)

Teaching Across Grade Levels—Support for Principals and Teachers

Ellington Elementary School allows for flexible scheduling of the school day. Teachers notice that the three 5th grade classes have up to five separate groupings for math. Based on this data, the principal decides to implement flexible groupings across the 5th grade level.

Now, the three math teachers collaborate on teaching math at the same time during the day, with each 5th grade teacher focusing on one to two groups at a time. Student groups are adjusted regularly as the students make progress.

Growth Scores Help Evaluate Program Success—Support for District Administrators

The Washington School District can use the end-of-year growth data to evaluate how Ellington Elementary School's reconfiguration of instructional groups impacted student performance, informing decisions on future program development, and professional development for other schools in the district.

TRANSITIONING TO HIGHER STANDARDS **USING GROWTH DATA AS A BRIDGE**

Within the context of changing standards and new expectations, measuring growth provides essential information to all of the education stakeholders. With the transition to Common Core State Standards, many expect (and have seen) that the number of students deemed to be proficient will drop for at least the first few years of implementation. During transitions like this, educators need to have an accurate indication of students' achievement along a stable continuum in order to provide context and insight on the progress of each student. Growth data can help to identify instructional needs and provide a more complete view of student learning—even when the number of students expected to be proficient drops by 20 or 30 percent.

More specifically, growth data allow us to:

 Assess student learning before, during, and after the transition to new standards. When the assessment measures students with items above-, on-, and below grade-level content, educators receive accurate data on where students are currently performing in relation to standards. They can then chart an instructional plan that moves their students toward new milestones and engages students in goal setting. They can also compare achievement on the current state standards to what is needed on the new standards.

- Support formative conversations and instructional shifts. The use of growth data creates a space for meaningful conversation among all stakeholders about what is and what is not working in the transition to new standards. Examining growth within the context of other factors fosters continued collaboration to get all students on an appropriate learning trajectory. It helps to answer the questions:
 - Did the student growth in Mr. Waller's class stand out in some way? Why?
 - Did a certain grade level or school exceed its growth projection? Why?
 - Can the students in my class use the growth data to own their learning?

- Evaluate programs. Schools and districts can use growth data to evaluate curricula and intervention programs, inform changes in instructional practice, and target professional development. Examining programs that create significantly more growth compared to other programs helps inform decisions on where to invest resources. Using only achievement level student data does not provide enough information to know if programs work.
- The use of growth data can help identify what's working and point to potential program exemplars for scalability. It helps us answer the questions: Did the students in our new math program experience higher rates of growth than other students?
 - On which academic standards are students exhibiting less than sufficient growth?
 - Where do our teachers need to focus instructionally?
 - What kind of professional development will assist our district in targeting areas of insufficient growth?
- Provide support for teachers and principals to become data and assessment literate. Educators should be supported in their professional growth on data literacy, especially in this time of national transition. A recent report from the Data Quality Campaign⁴ found that only four states have comprehensive plans to assist teachers in using data; only 16 states require teachers to demonstrate data literacy as a condition of certification. These statistics underscore the reality that many teachers and principals are not getting formal or informal learning opportunities to gain expertise with data. When teachers and principals are not able to speak the language of data, the opportunity for teacher and principal improvement may be lost.

But what does this look like in practice?

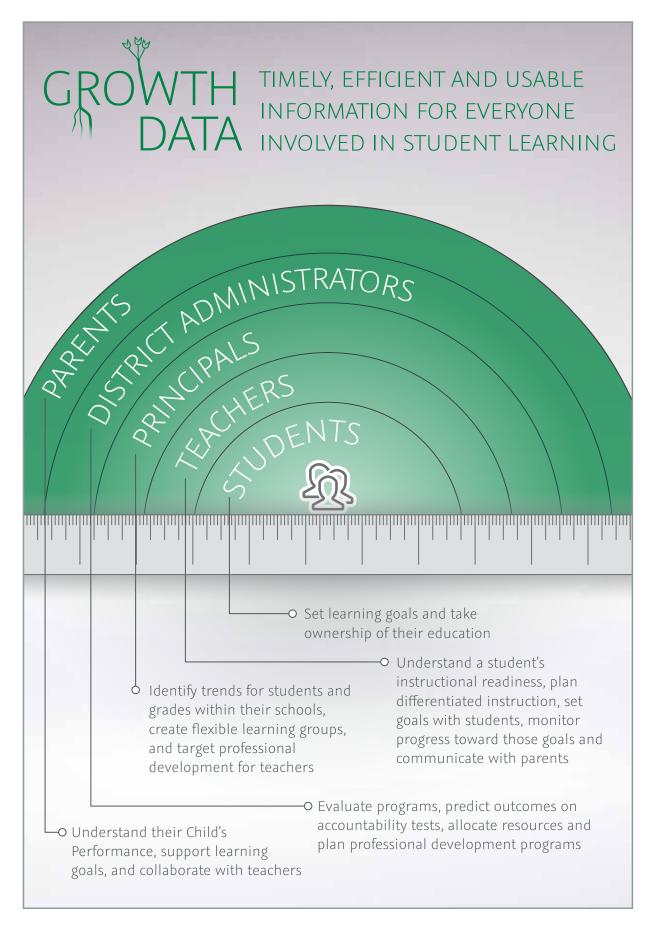
EVALUATING COMMON CORE PROGRAMS AND TARGETING TEACHERS' PROFESSIONAL DEVELOPMENT TO **ENSURE STUDENT GROWTH**

The Shaw School District implemented a new 4th grade reading program that emphasizes the Common Core shifts in reading, including more focus on informational text and eliciting evidence from text. The end-of-year assessment data revealed that the majority of 4th graders did not change from 3rd grade in terms of their broad achievement levels (i.e. needs improvement, proficient, etc.); however, the growth data demonstrated the majority of individual students made significant gains within their achievement levels. Instead of abandoning the reading program, the district decided to continue its use for 4th grade and extend it to the incoming 5th graders in order to chart its effects on longitudinal growth.

Disaggregating the data presented specific standards or goal areas that experienced lower rates of growth than other standards. The district targeted teacher professional development to address the standards on which students were exhibiting insufficient growth.

FINAL THOUGHTS—KEEPING THE FOCUS WHERE IT BELONGS: ON STUDENT **LEARNING**

Between managing accountability requirements and navigating shifts in standards, there is much for teachers, parents, and administrators to balance in the work of educating the next generation. Student growth data can be an invaluable and efficient tool in keeping the focus on students—what they need to grow academically and step into the lives they envision for themselves and their roles as citizens. By providing each student with the right instruction at the right time during the school year, growth data can help teachers develop a personalized learning path and engage students as partners, instilling an internal "personal navigation system" that leads to lifelong learning.



APPENDIX

MAKE SURE YOUR STUDENT GROWTH DATA MEASURES UP 10 QUESTIONS TO EVALUATE STUDENT GROWTH ASSESSMENTS

CLARITY OF PURPOSE

1. Be clear why you are assessing – are you measuring growth over time to inform instruction for all students? Communicate progress? For teacher and principal evaluation? For program evaluation?

QUALITY OF DATA AND ASSESSMENT

- 2. Does the test provide accurate scores for every student, regardless of achievement level, that measure the student's desired target?
- 3. Does the way the assessment questions are aligned to the standards make sense both from a content and depth of knowledge perspective?
- Has the assessment scale been validated and stabilized over time?
- 5. Will the assessment provide growth norms that allow comparison between similar students? Groups? Schools? Districts?
- **6.** Are the norms updated on a regular basis? Were the assessment items used on the test field tested? Have they been evaluated for bias?

USING THE DATA

- Have you considered how this assessment will complement your existing testing program?
- 8. Will you assess at reasonable intervals throughout the academic year in order to gain better insight on instructional impact?
- Can the assessment data efficiently provide both immediate insight and long term growth data?

WHAT ABOUT THE STUDENT?

10. How can you use the assessment data to engage students with learning goals, and how can families participate in supporting those learning goals?

ENDNOTES

- 1. Source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).
- 2. Source: Khadaroo, Stacy T., "New York test scores hint at hard road ahead for Common Core," August 8, 2013 Christian Science Monitor Web
- 3. Source: U.S. Census Bureau. 2011. "CPS 2011 Annual Social and Economic Supplement," Table POV01.
- 4. Source: Data Quality Campaign, Data for Action, 2012: Focus on People to Change Data Culture

COPYRIGHT© 2013 Northwest Evaluation Association. All rights reserved. Measures of Academic Progress, MAP, and DesCartes: A Continuum of Learning are registered trademarks and Northwest Evaluation Association and NWEA are trademarks of Northwest Evaluation Association.



Founded by educators nearly 40 years ago, Northwest Evaluation Association (NWEA) is a global not—for—profit educational services organization known for our flagship interim assessment, Measures of Academic Progress (MAP). More than 6,800 partners in U.S. school districts, education agencies, and international schools trust us to offer pre—kindergarten through grade 12 assessments that accurately measure student growth and learning needs, professional development that fosters educators' abilities to accelerate student learning, and research that supports assessment validity and informed policy. To better inform instruction and maximize every learner's academic growth, educators currently use NWEA assessments with nearly 8 million students.

Visit **NWEA.org** or call **866-654-3246** to find out how NWEA can partner with you to help all kids learn.

