



Center for
K–12 Assessment
& Performance Management

An independent catalyst and resource for the improvement of measurement and data systems to enhance student achievement.

Exploratory Seminar:
Measurement Challenges Within
the Race to the Top Agenda
December 2009

GROWTH IN STUDENT ACHIEVEMENT: ISSUES OF MEASUREMENT, LONGITUDINAL DATA ANALYSIS, AND ACCOUNTABILITY

This policy brief is based on a presentation by Damian W. Betebenner (National Center for the Improvement of Educational Assessment) and Robert L. Linn (University of Colorado, Boulder) at the Exploratory Seminar: Measurement Challenges Within the Race to the Top Agenda, December 2009. Download a copy of the final paper written by Dr. Betebenner and Dr. Linn, as well as the other papers presented at the seminar, at <http://www.k12center.org/publications.html>.

We are swimming in student achievement data, but often find ourselves lacking the information necessary for making crucial decisions. Techniques to turn data into information are available (e.g., growth analysis), but they are not a panacea. They bring attention to a wide variety of issues related to student growth including measurement, longitudinal data analysis, and inferences made for accountability purposes.

Measurement/Psychometric Considerations

When measuring growth, first we need to ask, Growth in what? In education, the construct to measure needs a specific reference, such as reading or math. Then we need to ask, How much growth are we looking for (i.e., What is the magnitude of it)? Having a normative frame of reference is critical in making this kind of judgment. If you are told a relative has grown three inches over the last year, your response would be very different if the relative were three years old than if the relative were 70 years old. Context is essential, and without it, understanding is impossible. Providing the context for understanding such differences is essential to explain the magnitude of growth to teachers and parents.

There are at least three approaches to calculating growth magnitudes:

- *Performance standards* anchor discussions about growth relative to discrete reference points (i.e., proficiency levels).
- *Vertical scales*, used by several states, presume that it is possible to look at change across multiple grades and that the scales have interval properties.
- *Learning progressions*.

Created by Educational Testing Service (ETS) to forward a larger social mission, the Center for K–12 Assessment & Performance Management has been given the directive to serve as an independent catalyst and resource for the improvement of measurement and data systems to enhance student achievement.

Despite shortcomings in each of these methods, people expect an answer about the magnitude of academic growth. However, magnitudes lacking a norm-referenced component will not be informative. It is important that whatever method chosen be embedded within a norm. Despite their unpopularity with criterion-referenced and standards-based advocates, norms can provide a useful framework for interpreting student growth, providing a transparent and fair way to communicate what is. Moreover, norms can be integrated with criterion referenced targets to indicate *what should be*.

Longitudinal Data Analysis Issues

Instead of beginning by asking about the technical aspects of measuring growth over time, we should be asking questions about what we want to know from the different approaches. Questions could include the following: How much annual growth are students making in reading? Are they making sufficient growth to reach/maintain desired achievement targets (growth-to-standard and growth model pilot program)? Are minority students and other subgroups making as much progress as other students? How much did the teacher/school contribute to students' growth over the last year (value-added)?

Knowing what questions stakeholders want answered is critical because the longitudinal analysis techniques often address different questions, which, in turn, lead to different uses and outcomes. The Race to the Top decisions and policymaking in some states have put a number of large issues on the table, but if the adopted analyses do not answer the questions people want answered about student academic growth, the analyses will fall short of stakeholder demands.

The techniques that could be used include

- *Gain scores* (they need a suitable scale).
- *Cross-tabulation* based on prior and current achievement levels (value tables, transition matrices).
- *Regression-based approaches* (growth-curve analysis, fixed/mixed effects models, growth norms).

Accountability Issues

High-stakes accountability in education has led to a vigorous use of large-scale assessment results in determinations of education quality, particularly at the school and teacher level. To much criticism, No Child Left Behind (NCLB) made extensive use of status results to judge schools. Growth results represent an improvement over status results but are not a panacea in investigations of education quality.

Growth measures basically are descriptive, but accountability has skewed discussions of growth from description toward determining responsibility for student achievement, or causality. Good descriptive models are very useful: they can be interpreted, they are informative, and they can serve many purposes. Unfortunately, growth models have been turned into inferences in the value-added debate. We do not have a value-added model as much as we have a value-added inference. The term teacher effect, for example, is so heavily laden with attempts to attribute responsibility for growth that it is hard to sell this use of value-added as a descriptive measure. The use of causal attributions in value-added

models may be popular for accountability discussions, but it is suspect and subject to considerable technical challenges unless it employs random assignment.

The Colorado growth model, one of the federal pilot programs, created a way to use student growth percentiles to depict student progress, relying less on attributions and more on descriptions. This separation of description from responsibility has been accepted broadly by the public, especially teachers' unions. The analysis uses quantile regression to calculate the relationships between current and prior achievement. Also, student growth percentiles can be criterion referenced to show growth to standards or end targets set by the state.

Ultimately, all models have problems, but sophisticated uses of models can yield results that are useful to policymakers and researchers. For example, when Colorado used its growth model to determine the influence of poverty at the school level, it found that some high-poverty schools had substantial growth, a finding that was similar to findings in Tennessee and Ohio.

For More Information

For more information on this subject, please see the paper by Dr. Betebenner and Dr. Linn:

Betebenner, D.A., & Linn, R. L. (2010). *Growth in student achievement: Issues of measurement, longitudinal data analysis and accountability*. Retrieved from <http://www.k12center.org/publications.html>.