

# THE EVOLUTION & USE OF GROWTH MODELS

By Christopher A. Cody, Joel McFarland, J. Eric Moore, & Jennifer Preston

Project 4.2 / August 2010

**POLICY QUESTION: What are growth models, and how can North Carolina use them to measure student performance?**

## GROWTH MODELS

The reauthorization of the Elementary and Secondary Education Act (ESEA) in 2001, known as No Child Left Behind (NCLB), and President Barack Obama's blueprint for its 2010 reauthorization have brought the idea of measuring student growth to the forefront of the education reform debate. In response, states across the nation have begun to explore how best to gauge the academic progress of students as they move through the public school system. Given the excitement around the concept of student growth, it is appropriate for us to explore the idea in more depth. In the following sections, this paper will present research on the evolution and use of growth models:

### Section I: Introduction

- A. Student growth and proficiency
- B. The importance of measuring growth
- C. North Carolina's use of growth information

### Section II: Models

- D. Types of student performance models
- E. Three categories of growth models
- F. Projection models

### Section III: Policy Uses & Recommendations

- G. The uses of growth models
- H. The uses of value-added models
- I. Recommendations for the use of growth models in policy decisions



## SECTION I: INTRODUCTION

### A. STUDENT GROWTH AND PROFICIENCY:

The Center for Assessment defines growth as simply “student change over time” (Gong, 2003). For the purposes of this report, the term “growth” broadly describes the amount of academic progress a student makes in a given timeframe, most commonly in a year of study. This concept complements the notion of “proficiency,” which measures a student’s performance against an externally established benchmark that signifies mastery of a subject (De Mello, 2009). These benchmarks often take the form of test scores, and proficiency measures how a student performs against a standard.

In situations where students have made tremendous growth, but have not yet attained proficiency, the two concepts may be in conflict.

**EXAMPLE:** A student enters the eighth grade at a fifth-grade reading level, but improves to a seventh-grade reading level by the end of the year. Although the student never reached proficiency for her level (eighth grade), she also made two years’ worth of growth in just one year.

Thus, neither growth nor proficiency should solely describe a student’s performance. Most of the models discussed in this paper marry the two concepts together for a more complete assessment.

#### RECAP:

- *Student growth: the amount of academic progress made in a school year*
- *Student proficiency: the measure of student performance against a defined level of achievement*

### B. THE IMPORTANCE OF MEASURING GROWTH:

Understanding students’ growth from year-to-year helps to pinpoint advanced or struggling learners and addresses the concern that proficiency assessment concentrates solely on children whose scores are clustered around a passing cut-score. Educators can then target specialized instruction to meet an individual student’s needs. For schools, ascertaining academic growth can help assess teacher impact on student learning. The presence of a cluster of high-growth students who share the same teacher may indicate that teacher to be particularly effective in helping students make academic progress. Similarly, growth data can inform reports on effectiveness at the school and district level. From year to year, the state may tap into growth information to assess cohort progress and to ensure each grade performs better than the previous one.

#### RECAP:

- *Growth provides a more detailed picture of a student’s progress*

### C. NORTH CAROLINA’S USE OF GROWTH INFORMATION:

Over the last fifteen years, North Carolina has used growth in its education programs, and the state plans to continue to do so in the future.

#### The ABCs of Public Education Accountability Program:

In 1998, North Carolina’s General Assembly passed a law amending its educational provisions to require the assessment of student improvement as a component of the evaluation of school employee performance. The language states, “The State Board, in consultation with local boards of education, shall revise and develop uniform performance standards and criteria to be used in evaluating certified public school employees.... These standards and criteria shall include improving student achievement....” (N.C. Gen. Stat. § 115C 335. 1998 5, s. 4). This line further codified the State Board of Education’s efforts to measure student performance as conceptualized in its accountability initiative, the New ABCs of Public Education, which had included growth measurements for grades K-8 during the 1996-1997 school year (North Carolina Department of Public Instruction, August 2009). Thus, by law, North Carolina must measure the academic growth of its students.

#### No Child Left Behind Accountability:

In 2005, United States Secretary of Education Margaret Spellings announced a program that would allow states to request to use growth information to meet reporting obligations required by No Child Left Behind (NCLB), the most recent reauthorization of the ESEA. North Carolina and Tennessee piloted the use of growth in their federally reported data. Due to its apparent effectiveness in measuring student success, nine states now assess growth to determine Adequate Yearly Progress under NCLB (United States Department of Education, 2008).

#### New Reform Initiatives:

In 2009, the Department of Education announced a new grant entitled Race to the Top, a product of the American Recovery and Reinvestment Act. In two rounds of funding, the competitive grant will award money to states to reform their educational systems, emphasizing reform of student performance evaluation. In North Carolina’s application, the state draws upon a resolution by the State Board of Education to commit itself to using student growth data to inform determinations of principal and teacher effectiveness (North Carolina Department of Public Instruction, 2010). The award (up to \$400 million) will partially rely on whether the state has the ability to measure student growth. North Carolina has been named a finalist for Round Two of Race to the Top.

The SMARTER Balanced Assessment Consortium, of which North Carolina is a governing member, is one of two major consortia to grow out of the Race to the Top initiative. These consortia have begun to look more in-depth at measuring and using growth information. Regardless of the outcome of Race to the Top, North Carolina will still be committed to the goals of the consortium. This movement suggests the educational culture to be trending toward

the use of student growth models in evaluating the effectiveness of its system from the teacher level upward. The next section explores performance models that incorporate growth measurement and describes two older models that do not.

**RECAP:**

- *North Carolina uses growth in its ABCs accountability program*
- *The state uses growth to satisfy No Child Left Behind requirements*
- *Growth plays a major role in the state's new reform initiatives*

**Improvement Models:**

Improvement models attempt to evaluate performance by measuring changes in achievement between cohorts. Improvement models compare the scores of one cohort, or class, of students in a grade to the scores of a subsequent cohort of students in the same grade (Hull, 2007). Such reports would take the form of "this year's fifth-graders scored seven percentage points higher in proficiency than last year's fifth-graders." Status and improvement models are capable of measuring the change in the percent of students meeting a certain benchmark (typically proficient), but they do not measure an individual student's growth from year to year (Hull, 2007).

**Growth Models:**

Recall that student growth is the amount of academic progress gained in a school year. The fundamental distinction between status, improvement, and growth models centers on whether prior achievement should be taken into account when assessing performance. Some growth models examine student development based upon a longitudinal record of student performance (Hull, 2007; Jennings and Corcoran, 2009). Other growth models also include gender, race and ethnicity, and socioeconomic status to qualify current status (Betebenner, 2009).

## SECTION II: MODELS

### D. TYPES OF STUDENT PERFORMANCE MODELS:

There are three ways of measuring student performance: status, improvement, and growth models. For a visualization of the models discussed below and how they have evolved over time, see Figure 1.

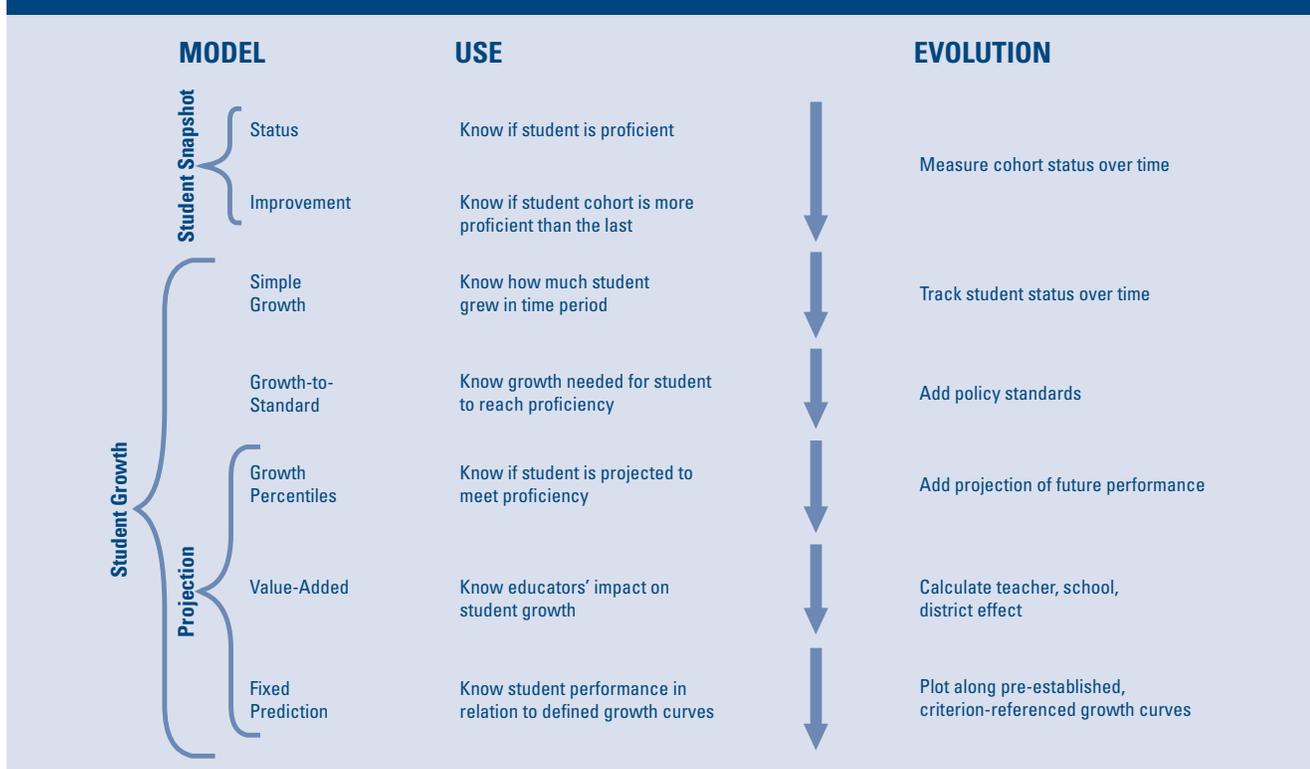
**Status Models:**

Prior to the 2005 revision of NCLB, North Carolina used status and improvement measures for federal reporting purposes. Status models examine student performance at a point in time (Betebenner, 2009). Essentially, a student's status may be reported generically, e.g. "proficient," "basic," etc., or in reference to cohort scores, e.g. "95th percentile." This status is a snapshot of the student and does not compare performance across years.

**RECAP:**

- *Status and improvement models gauge a student's performance on one measure of achievement*
- *Growth models measure student performance across time*

**FIGURE 1: EVOLUTION OF PERFORMANCE MODELS**



### E. THREE CATEGORIES OF GROWTH MODELS:

Most growth models tend to fall into three general categories: simple growth, growth-to-standard, and projection models. Each of these categories encompasses several variations depending on the model's purpose and data, and since different models continue to be developed, these categories may not capture all models.

**TABLE 1: THREE CATEGORIES OF GROWTH MODELS**

#### **SIMPLE GROWTH**

Function – Measures difference in student's scale scores from year to year

Usage – Can be used to calculate school, district, and state growth

#### **GROWTH-TO-STANDARD**

Function – Shows if students are on track to meet proficiency standards

Usage – Can be used to calculate school, district, and state growth and proficiency

#### **PROJECTION**

Function – Predicts students' academic levels based on previous test scores

Usage – Can be used to calculate growth and proficiency at the student, school, district, and state levels

#### **Simple Growth:**

Simple growth models explain growth based on changes in a student's scale scores as he moves from grade to grade (Hull, 2007). Scale scores are single numeric scores that show the overall performance on a standardized test.

**EXAMPLE:** If Student A scores 420 for the fifth grade assessment and 500 on the sixth grade exam, the student made 80 points of growth. The school's growth for sixth grade is then the average of all students' growth scores in Student A's cohort. This same process yields growth measures at the district and state level as well.

While simple growth models measure growth based on individual student growth from year to year, they do not indicate if the growth occurring meets state standards of proficiency. Advancements in measuring student growth have produced more nuanced models that are attractive to states and school districts; as a result, there are no examples of the simple growth model in practice.

#### **Growth-to-Standard:**

Growth-to-standard models, often described as growth-to-proficiency models, show whether students are on track to reach proficiency (Hull, 2007). These models have become widely used due to the NCLB Growth Model Pilot program. Growth-to-standard models include categories of performance that describe how well students understand the knowledge and skills being assessed (Hull, 2007). In other words, growth-to-standard models add proficiency cut points to models such as simple growth. A student's growth is deemed adequate if it will lead to future proficiency (Betebenner & Linn, 2009). Growth-to-standard models attempt to depict growth to policy-mandated levels (Betebenner, 2009).

**EXAMPLE:** A state that uses a growth-to-standard model would test students annually in reading and mathematics in grades 3-8 and 10. The model then would require that students reach a defined proficiency within three years or by the eighth grade. Student growth targets would be calculated by subtracting a student's current year scale score from the scale score for proficiency three grades later and dividing by the number of remaining grades (O'Malley, 2008).

#### **Projection:**

Projection models predict how much academic progress a student will make in a given year based on previous test scores. The projected growth of a student is obtained by comparing a student's previous test scores to those of students with a similar academic history. If a projection model incorporates proficiency levels in determining a student's skills and knowledge, then the projections from the model can assist in determining how much a student needs to grow over certain years in order to reach proficiency. This process yields projected growth measures at the district and state level as well.

#### **RECAP:**

- *Simple growth models – change over time*
- *Growth-to-standard models – change required to reach proficiency standard*
- *Projection models – predicted change*

### F. PROJECTION MODELS:

Projection models are state-of-the-art, and, as such, this paper will go into more detail to explore their nuances than it does with simple growth and growth-to-standard models. For simplicity, the study will examine the projection category by reviewing the following three models: two current popular models (Student Growth Percentiles and Value-Added) and a third that may exist in the future (Fixed Prediction).

**TABLE 2: THREE TYPES OF PROJECTION MODELS****STUDENT GROWTH PERCENTILES**

- Design – Weight current test scores more heavily
- Usage – Demonstrate growth, not in test score point gains or losses, but in percentiles of gain

**VALUE-ADDED**

- Design – May include demographics
- Usage – Determines what factors were responsible for student growth
  - Produces value-added scores for teachers, schools, and districts

**FIXED PREDICTION**

- Design – Based on defined growth curves
- Usage – Predicts student trajectories across multiple years
  - Can yield value-added scores by showing movement between curves

**Student Growth Percentiles:**

The student growth percentiles model allows states to measure a student’s annual growth and project his or her future academic achievement. The student growth percentiles model was initially developed in Colorado. It was designed to define achievement based on one year’s growth or more in one year’s time and explain what constitutes proficiency in terms of adequate growth. The student growth percentiles model demonstrates growth, not in test score point gains or losses, but in percentiles of gain. The model uses a student’s gain or loss in test score points as his or her basis for growth calculations using quantile regression. This methodology uses all available test scores to estimate growth percentiles for every student. These growth percentiles are a normative measure of academic progress, comparing a student’s progress over time to that of students with a similar academic history (Betebenner, 2009). Student growth percentile scores have a simple interpretation in which a student that has an 84th growth percentile indicates that the student’s growth is as good or better than 84 percent of his or her academic peer groups. Academic peer groups are defined as students in a particular grade with a similar test score history.

Percentile growth projections are calculated for each student and emphasize more current test scores in the model. These projections indicate what it will take for a student to progress toward, reach, or maintain state-assigned proficiency/achievement levels. Growth adequacy is determined by whether a student’s growth is sufficient to reach defined

proficiency/achievement levels (Betebenner, 2009). The model could also be used to calculate value-added scores (discussed below) for teacher and schools, though there are no current plans to do so (Bonk, 2010). While Colorado is the pioneer for the student growth percentile model, Massachusetts, Arizona, and Indiana use the model as well.

**Value-Added Models:**

Some growth models incorporate a value-added component that measures how much students learn in a year, and then determines what factors were responsible for that growth. These models use complex statistical techniques to isolate the influence of individual school districts, schools, and teachers on student achievement (Meyer and Dokumaci, 2010). At a basic level, most value-added models compare a student’s projected growth with his or her actual growth (Hull, 2007). The models report the influence of a particular district, school, or teacher as a value-added score. For example, a teacher whose students scored higher than expected would have a high value-added score.

Several different value-added models have been implemented in states and school systems across the country. They differ in terms of the type of data involved and intended policy uses. This section will focus on three of the most widely known models: the Education Value-Added Assessment System, the Milwaukee model, and the Dallas model.

**Education Value-Added Assessment System (EVAAS)**

The Education Value-Added Assessment System (EVAAS) generates value-added estimates for teachers, schools, and districts, and also predicts how individual students will fare on future tests. The model was originally developed by researchers in Tennessee as the Tennessee Value-Added Assessment System (TVAAS), but is now commercially available to other states through the SAS Institute. The model has been implemented in several states, including North Carolina.

The EVAAS model has several features that distinguish it from other value-added models. First, EVAAS can use datasets with missing values. At minimum, the model requires data from three testing occasions. However, EVAAS is flexible about which tests it uses and performs statistical adjustments when values are missing. This attribute is important, since missing data are common. Students may be absent on testing days or data may be lost as students transfer between school systems and between states.

A second feature of EVAAS is that it utilizes only achievement score data; it does not make adjustments for a student’s demographic characteristics nor does it give weight to more recent test scores. This strategy is possible because the model focuses on variation in a student’s test scores over time. Since demographic characteristics likely do not change over time, variations within a student’s testing history cannot be attributed to these factors. In addition to the statistical reasoning, there is a philosophical basis for excluding student demographic data. The developers of EVAAS strongly believe that including such variables would create different learning expectations for

students based on their demographic characteristics. Under this rationale, including such variables would “let teachers off the hook” for producing lower test gains with students from traditionally disadvantaged backgrounds.

### ***Dallas and Milwaukee Models***

The Dallas and Milwaukee school districts use value-added models that differ from EVAAS in their use of demographic data. In addition to test scores, these two models incorporate information on student characteristics, such as race and income level. In both cases, the researchers who developed these models have specified that the models are to be used for measuring teacher and school effectiveness, not for setting performance standards.

The Dallas model, which the city’s schools have used since the mid-1990s, includes “fairness” variables that measure student characteristics. Evaluators of the model have justified their inclusion by noting that they help measure teacher and school productivity more fairly and with more accuracy (Thum & Bryk, 1997). The evaluation drew a clear distinction, however, between measuring effectiveness and setting high achievement standards. Researchers agree that considering demographic factors when setting standards would be inappropriate (Meyer and Dokumaci, 2009; SAS, 2010).

In the early 2000s, Rob Meyer, an economist at the University of Wisconsin-Madison, developed another popular value-added model. Meyer originally developed the model for use in Milwaukee Public Schools, but the Chicago and New York City school systems have also used the model in some form. Meyer has explained that his model incorporates student demographic data for two reasons. First, they allow policymakers to identify and measure gaps in growth rates among different groups of students. Second, these variables ensure that value-added estimates are not biased against schools that “disproportionately serve students who, on average, exhibit relatively low growth (for example, low-income students)” (Meyer and Dokumaci, 2010). Under this argument, demographic data is included so that teachers and schools are not held accountable for factors outside of their control, such as having students who are, on average, more challenging to teach. Discussions of the Milwaukee model emphasize the use of demographic variables to diagnose growth rate gaps, rather than to set standards for expected achievement gains.

### **Fixed Prediction Model:**

Future growth models may attempt to predict student performance based upon previously established growth curves (Williamson, 2010). These curves would be similar to the curves the Centers for Disease Control and Prevention has created to predict height and weight trajectories for children based upon age and gender. Student growth curves would predict, given previous scores, a student’s score trajectory over several grades. Longitudinal studies of student performance using

vertical scales and linked testing standards would form the basis of these curves. For psychometric purposes, such standards should be based in real-world application and form a continuum of possible scores across all measured grades (vertical scaling) much like MetaMetrics, Inc. has attempted to do for reading by creating Lexiles. These two criteria would help ensure the curves would have meaningful anchors (e.g. reading level of an incoming college freshman) and be able to capture smoothly the progression of students’ learning over the course of their school experience. The stacked lines would represent percentiles of students at different levels of performance. A student’s current score and previous history can be plotted against the curves to predict his or her performance trajectory. Beyond that, movement between curves can contribute to value-added calculations previously discussed.

#### **RECAP:**

- ***Student growth percentiles – demonstrate growth in percentiles of gain***
- ***Value-Added – determines what factors were responsible for that growth***
- ***Fixed prediction - predicts student performance based upon previously established growth curves***

## **SECTION III: POLICY USES AND RECOMMENDATIONS**

### **G. USES OF GROWTH MODELS:**

The use of a growth model for policy purposes requires a significant shift in how stakeholders think about student achievement and what makes a school “good.” A “good” school might be one in which teachers cause a great deal of growth in their students, even if these students still fail to meet proficiency (Blank and Cavell, 2005). Growth models are best suited for use in school and district accountability programs, rewards for schools, and academic goal setting for students.

#### **School And District Accountability:**

One of the most common uses for a growth model is in school and district accountability measures. Rather than selecting a certain proficiency level that students must attain, state education officials determine the amount of growth that the students at a school must make (Jennings and Corcoran, 2009).

**EXAMPLE:** In Colorado, officials defined “normal growth” as one year’s worth of growth in one academic year, as well as “adequate growth” as the amount a student must grow in order to meet state standards (Technical Advisory Panel for the Longitudinal Analysis of Student Assessment, 2008).

One advantage to this approach is that it holds schools accountable for the performance of their higher-performing students. Many of these students begin a course already close to proficiency in that subject; no further development is expected from them. However, with an accountability scheme that uses growth, schools are expected to challenge these students to even higher levels of achievement (Jennings and Corcoran, 2009).

A second advantage of using a growth model is that it, more accurately than other performance models, shows that many students in the United States lack the skills and knowledge necessary to be successful (Harris, 2010). There is robust debate on the reasons for their academic deficiencies. The question remains: is it reasonable to expect students who start the school year several grades behind to be able to score proficient on state exams? Some would argue no, and, as a result, offer the use of growth models to measure the amount of improvement for these students, regardless of where they end the year. Growth model advocates argue that a student who grows three years in his or her reading level but fails to meet a reading proficiency level has still made significant progress (Harris 2010).

A disadvantage to the use of a growth model is how it may be perceived. Critics offer that growth models allow teachers and schools to accept the status quo when low-performing students fail to meet state standards (Blank and Cavell, 2005). They argue that proficiency models send a message that all students should achieve at the same high level of performance, while growth models “make excuses” for students who fail to do so. While these concerns are not unfounded, the consensus within the education community has moved toward a greater acceptance of the use of growth models for school-level accountability (United States Department of Education, 2008).

#### RECAP:

- *Educational leaders set required amount of growth*
- *Teachers and schools must improve the achievement of higher-performing students*
- *Students who begin school lacking skills or knowledge are not “penalized”*
- *Use of a growth model may relax emphasis on proficiency*

#### Rewards For Schools:

Independent of any accountability system, states can also use growth models to identify and reward schools whose students demonstrate high growth. Here in North Carolina, the ABCs Accountability Model combines a proficiency model with a growth model to evaluate school effectiveness. Schools receive certain designations based on the comparison of the growth of their students with the growth that was expected, as well as the percent of students that score an achievement level of III or higher on state exams.

**EXAMPLE:** For several years, teachers and staff at North Carolina schools that achieved expected growth, or growth that is higher than expected, received a salary bonus of up to \$1500 (North Carolina Department of Public Instruction, 2009). The General Assembly has recently eliminated the program in response to the national economic downturn.

#### RECAP:

- *States and districts can provide financial awards to schools and teachers of students who demonstrate higher than required levels of growth*

#### Academic Goal-Setting For Students:

Many growth models include a predictive element that allows a teacher to see how much growth a student is predicted to make in his or her class. Educators can then use this information to create growth plans that establish the amount that students should grow by certain times of the year, for example, by when benchmark testing occurs. In this way, growth models can be used to provide formative data to teachers. Information from growth models can also be used to ensure that students take appropriately challenging coursework that will allow them to make the growth predicted.

#### RECAP:

- *Growth models that can make predictions on student performance can be used to set academic goals*



## H. USE OF VALUE-ADDED MODELS:

Value-added models are subject to intense scrutiny in the current education environment. While value-added models may be used for high-stakes programs, they can also be used for less controversial purposes.

### High-Stakes Uses:

Some proponents of value-added models advocate for their use in merit pay programs for teachers, as well as for evaluations for educators. These uses are considered to be high-stakes because they determine the employment and compensation of individual teachers. This analysis also examines challenges to the use of value-added models for high-stakes purposes.

### Merit Pay:

Value-added models can be used to provide merit pay to educators. The most controversial variation of merit pay involves actual salary increases for teachers who are considered effective. Supporters claim that merit pay rewards teachers for improved student outcomes (Eckert and Dabrowski, 2010). The most well known example of such a program can be found in the nation's capital.

**EXAMPLE:** In April 2010, D.C. Public Schools and the Washington Teacher Union agreed to a teacher contract that includes a merit pay component for teachers who elect to be paid according to performance. The merit pay plan includes not only higher pay for teachers, whose students perform at high levels, but also lower pay and other consequences for teachers whose students do not have high levels of achievement. The funding for merit pay for those teachers will be provided by private foundations (Simmons, 2010). While D.C. Public Schools has not moved completely to a merit pay system, the new contract does represent a significant shift in that direction, especially as teachers' unions have traditionally been opposed to any type of merit pay in contracts.



Some states and school districts have chosen to take a less controversial approach to merit pay: they use value-added models only for awarding bonuses to effective teachers. While this type of bonus program does provide additional funding for effective teachers, it does not generate as much controversy because the teachers' salaries are not entirely based on student achievement.

### RECAP:

- *School districts can determine teacher salaries based on an educator's value added to student achievement*

### Teacher Evaluation:

In addition to merit pay, value-added models can be used for teacher evaluations and tenure decisions. Most education experts offer that teacher evaluations should never be based solely on value-added measures (Goe, 2008). Teachers also work to develop citizenship, ethics, self-esteem, and communication skills in their students; standardized tests do not assess students' development in these areas (Goe, 2008). At the same time, most education officials acknowledge that the current teacher evaluation and tenure process does not place enough emphasis on actual student achievement; far less than 90 percent of American students are receiving a quality education even though over 90 percent of teachers in the nation receive "satisfactory" ratings on teacher evaluations (Eckert and Dabrowski, 2010).

### RECAP:

- *School administrators can base a percentage of a teacher's evaluation on his or her effectiveness as measured by value-added scores*

### Challenges to High-Stakes Uses of Value-Added Models:

States and school districts that have implemented merit pay or teacher evaluation based on value-added models have encountered several major challenges: teachers in non-tested areas, bias against teachers of high-achieving students, negative effects on teachers, and competition between teachers.

### Teachers in Non-Tested Areas:

For merit pay programs, one solution to assessing teachers of non-tested areas is to award school-wide bonuses to staff at schools in which student growth exceeds projections. Essentially, high school-wide growth indicates that the value added by the school is also high. North Carolina's ABC program used this system to award schools that met or exceeded expectations for growth. Under other types of teacher improvement initiatives, such as the Teacher Assessment Program, teachers are eligible to receive salary increases for their personal added value to education, as well as the school's added value (Milken Family Foundation, 2009).

### ***Bias against Teachers of High-Achieving Students:***

As mentioned above, growth models have the potential to better assess the performance of high-achieving students. A practical limitation to this advantage is that state tests cannot include above-grade level items. Students who meet, and exceed, the expectations for their course cannot demonstrate their true performance level (Eckert and Dabrowski, 2010). Thus, there is a ceiling effect on the value-added measures for their teachers, which can affect their merit pay and evaluations.

### ***Nonrandom Student Assignment***

Students with varying ability levels and propensities to learn are not randomly assigned to teachers. Parental residence choice affects the distribution of students among districts, and parents or principals may influence the assignment of students to particular teachers. This nonrandom student assignment may undermine the validity of teacher value-added measures. Economist Jesse Rothstein (2008) suggests that this may distort teacher incentives to teach certain types of students. He claims, “Teachers operating under high-stakes value-added model based accountability and incentive systems can be expected to lobby their principals to be assigned the ‘right’ students who will predictably yield high value-added scores, and principals will presumably alter their assignment rules to direct these students toward favored teachers.”

### ***Negative Effects on Teachers:***

Some caution against the use of value-added models for salary determinations and performance evaluations because they have such critical ramifications on the economic well-being of teachers and their families. Studies have shown that the calculations in some value-added models vary from year to year. As a result, measures of teacher effectiveness may change even when one considers the same teacher, the same group of students, and the same assessment. If measures of teacher effectiveness have implications for job security and pay, it may be problematic to rely on a measure that may fluctuate simply due to changes in calculation (Eckert and Dabrowski, 2010).

### ***Competition between Teachers:***

The use of value-added models may pit teachers against each other in a competition for higher pay and better performance evaluations. This would occur in situations in which financial rewards are given to the top few teachers, rather than to any teacher that reaches a certain benchmark of student growth. In its best form, teaching is a collaborative process in which educators share best practices and work together to increase their effectiveness (Goe, 2008). The use of value-added models for salary determinations and job evaluations could shatter this spirit. Researchers have offered the possibility of calculating the value-added of certain groups of teachers, such as fourth grade teachers who “team-teach” their students, or all the United States History teachers at a high school. While this adjustment would make it more difficult to determine the pay or evaluation of a single teacher, it would encourage collaboration in teaching.

### **RECAP:**

- ***Challenging to determine value-added for teachers in non-tested areas***
- ***Difficult to demonstrate growth in high-achieving students***
- ***Nonrandom student assignment may undermine the validity of value-added scores***
- ***Opaqueness in how value-added models produce scores for individuals***
- ***Teachers may compete to increase their own value-added measures***

### **Low-Stakes Uses:**

Value-added models can also be used for a variety of low-stakes purposes. Many of these uses are less controversial and may be easier to implement as a result. There are five such low-stakes uses: identification of teachers in need of professional development, design of alternate licensure programs, evaluation of professional development programs, evaluation of teacher preparation programs, and evaluation of teaching styles and curriculum design.

### ***Identification of Teachers in Need of Professional Development:***

School administrators can use value-added models to identify ineffective and effective teachers, not for sanction or reward, but simply for professional development opportunities. Teachers with low added value are identified for additional assistance from administrators and curriculum specialists. Administrators have also used value-added models to identify effective teachers to serve as mentors and provide in-house professional development for their peers. When designing individual growth plans, teachers can use their value-added measures to set quantitative goals and select steps for improvement (Goe, 2008).

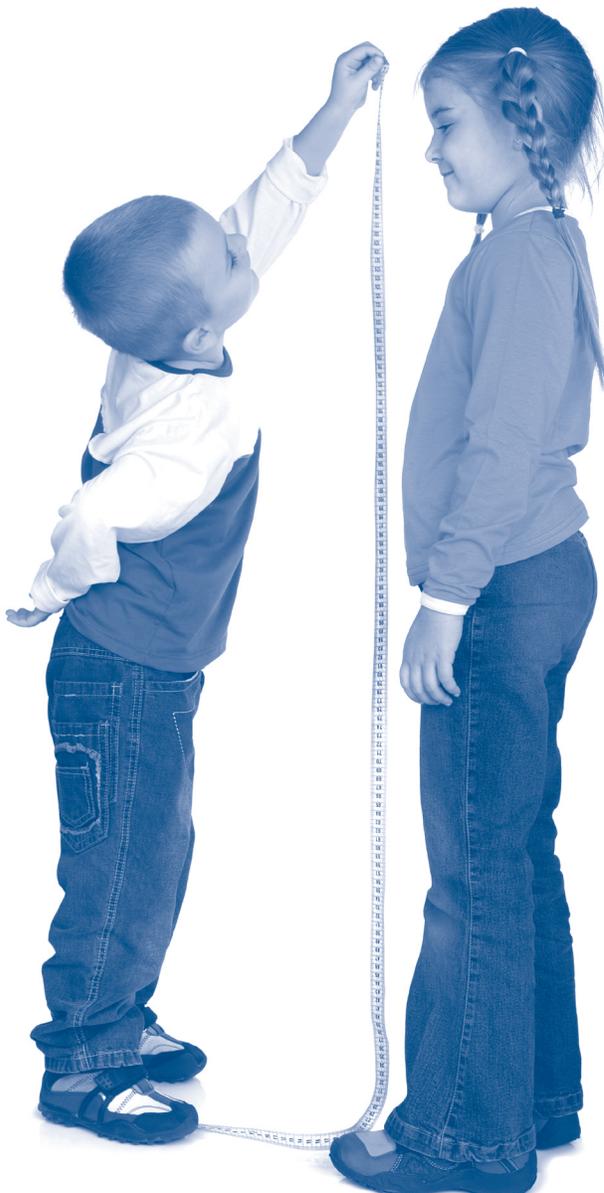
### ***Design of Alternate Licensure Programs:***

In many states, a teacher’s certification changes based on the number of years that he or she has spent in the classroom. Most advocates of the use of the value-added models do not support the elimination of this system, but the addition of innovative types of certification or pathways to certification through value-added measurements. In Georgia, state officials created a “Master Teacher Certificate” for certified teachers with at least three years of experience and demonstrated success in raising student achievement. Other states could follow suit and create new categories of licensure for teachers with high added value (Georgia Professional Standards Commission, 2010). Another possibility is to allow teachers to advance their licensure status more quickly through years of demonstrated student achievement than through years in the classroom.

### ***Evaluation of Professional Development Programs:***

Some school districts have begun to use value-added models to identify professional development needs and to determine whether particular professional development programs are

effective. For example, in the Houston Independent School District, administrators determine a teacher's added value for different groups of students, such as low-income, racial or ethnic minorities, and low-achieving. The district then provides targeted support on effective ways to educate students from groups for which teacher's added value is low (Houston Independent School District, 2010). Schools and districts can compare the value added by teachers who have participated in different forms of professional development. The comparison allows them to select the professional development models that are more effective for their teachers. Program evaluation staff can also use value-added models in their work on professional development programs (Goe, 2008).



#### ***Evaluation of Teacher Preparation Programs:***

The state of Louisiana has pioneered a method of using value-added models to identify effective and ineffective teacher preparation programs. This system allows the Louisiana Board of Regents to quantitatively evaluate the state's teacher training programs, and it also provides a way for colleges and universities to track the student achievement of various cohorts of teachers from their teacher preparation programs. There are several statistical concerns about such an evaluation system that would merit attention before more widespread adoption. It would be difficult to track the effectiveness of teachers who leave the state; there could be certain personal characteristics or characteristics about their training and education that might make them more likely to move out-of-state. It is also challenging to control for the various elements that determine whether a teacher is hired for a "good fit" position at a "good fit" school in a "good fit" district. All of these factors may influence their value-added scores (Goe, 2008).

#### ***Evaluation of Teaching Styles and Curriculum Design:***

Further statistical work is needed before the last use of value-added models can be widely implemented: the identification of effective teaching strategies (such as project-based units) and curriculum design styles (such as teaching history from the present-day backward). It has been difficult to measure the effectiveness of different teaching strategies. Value-added results, if used with robust statistical controls, may be able to isolate the effects of certain teaching styles such as project-based learning, student groups, or traditional lecture (Goe, 2008). This type of analysis would be difficult to complete because many teachers vary how they teach throughout the year. However, it might be easier if the analysis was simplified and value-added models could be applied to the amount of teacher value-added to one discrete unit of curriculum. Such an approach would also allow teachers to receive formative information on their added value while they are still teaching, thus allowing them to make modifications to instruction.

#### **RECAP:**

- ***School leaders can identify teachers in need of support***
- ***Effective teachers may pursue accelerated routes to licensure***
- ***School districts can select effective professional development programs***
- ***Teacher preparation programs can evaluate performance of their graduates***
- ***Teachers can identify styles and curricula that are effective for their students***



### I. RECOMMENDATIONS FOR THE USE OF GROWTH MODELS IN POLICY DECISIONS:

Choosing a model to adopt is not solely a mathematical decision, but one that involves significant understanding of the purposes and policy implications of each. The following are recommendations to consider when adopting a growth model:

1. Growth and proficiency: Neither growth nor proficiency should solely describe a student’s performance. The combination of the two describes a student’s academic progress and the level of mastery achieved.
2. Teacher evaluations: Value-added measures should never be the sole basis for teacher evaluation. A qualitative assessment, such as a principal evaluation, would capture other, non-tested aspects of teaching.
3. Professional development: Training should help teachers interpret growth information and use it to inform instruction. Professional development should also train principals and administrators to use growth information to place students and teachers appropriately.
4. Low-stakes: Educators should take advantage of the often-overlooked low-stakes uses of growth models.

### CONCLUSION:

State education agencies and the federal Department of Education continue to grapple with the concept of growth and how to use growth measurements to improve education. Each growth model has its own strengths and limitations, and it is likely that robust debate will continue over which model is most appropriate for use in the states, as well as at the federal level. States continue to explore innovative ways to use growth, both for high- and low-stakes programs. As states do so, in-depth evaluations of their initiatives will provide useful information to guide the discussion on growth. The decision on which growth model to use in a state hinges largely on the policies and programs for which it will be used.

**TABLE 3: GROWTH MODEL STRENGTHS AND LIMITATIONS:**

#### STRENGTHS

- Design – Examines students in a longitudinal perspective
- Not limited to only examining how high or low a student’s performance is or how close it is to achievement thresholds
  - Uses initial levels of achievement which are fairer bases for the comparison of teachers, schools, and local education agencies
  - Takes into account students’ different starting points at the beginning of the school year
- Usage – Fits the view that education should help students progress from one level of knowledge and skills to a higher level

#### LIMITATIONS

- Design – Measurement error
- Test score scaling problems:
    - Tests may not be designed for the specific purpose of comparing student progress across grades
    - “Ceiling effect” – questioning if tests have the ability to measure growth of high-performing students
  - Nonrandom student assignment may undermine the validity of value-added measures
- Usage – Difficult to express growth models in terms for public consumption due to the advanced statistical analysis involved
- Lacks transparency
  - Relaxes assumptions about the relationship between proficiency rates and school quality.
  - Limits a schools incentive to move low-performing students all the way to proficiency.

## WORKS CITED

- Betebenner, Damien W. "Norm- and Criterion-referenced Student Growth." *Education Measurement: Issues and Practice*, 28, 4 (2009): 42-51.
- Betebenner, Damien W. and Robert L. Linn. *Growth in Student Achievement: Issues of Measurement, Longitudinal Data Analysis, and Accountability*. Center for K-12 Assessment And Performance Management. Washington, D.C. December 2009. Accessed at <http://www.k12center.org/publications.html> on 4 June 2010.
- Blank, Rolf and Lori Cavell. *Memorandum to Chief State School Officers: Growth Models Primer*. The Council of Chief State School Officers. Washington, D.C. 16 March 2005. Accessed at <http://www.ccsso.org/content/pdfs/Growthmemo.pdf> on 27 May 2010.
- Bonk, William. Personal communication. 25 June 2010.
- De Mello Victor, et al. *Mapping State Proficiency Standards Onto NAEP Scales: 2005-2007: Research and Development Report*. National Center for Education Statistics. US Department of Education. Washington, D.C. October 2009. Accessed at <http://nces.ed.gov/nationsreportcard/pdf/studies/2010456.pdf> on 4 June 2010.
- Eckert, Jonathan M. and Joan Dabrowski. "Should Value-Added Measures Be Used for Performance Pay?" *Phi Delta Kappan*. 91, 5 (May 2010). 88-92. Accessed through JSTOR on 2 June 2010.
- General Assembly of North Carolina. *N.C. Gen. Stat. § 115C 335*. 1998 5, s. 4. Raleigh, NC. 1998. Accessed at [http://www.ncga.state.nc.us/enactedlegislation/statutes/html/bysection/chapter\\_115c/g\\_s\\_115c-335.html](http://www.ncga.state.nc.us/enactedlegislation/statutes/html/bysection/chapter_115c/g_s_115c-335.html) on 7 June 2010.
- Georgia Professional Standards Commission. About Master Teacher Certification. Atlanta, GA. 15 September 2009. Accessed at [http://www.gamasterteachers.org/MT\\_about.aspx](http://www.gamasterteachers.org/MT_about.aspx) on 10 June 2010.
- Goe, Laura. *Using Value-Added Models to Identify and Support Highly Effective Teachers*. National Comprehensive Center for Teacher Quality. Washington D.C. May 2008. Accessed at <http://www2.tqsource.org/strategies/het/UsingValueAddedModels.pdf> on 2 June 2010.
- Gong, Brian. *Models for Using Student Growth Measures in School Accountability*. Center for Assessment at the Council of Chief State School Officers. Washington, D.C. 15 November 2003. Accessed at <http://www.nciea.org/publications/GongGrowthModels111504.pdf> on 4 June 2010.
- Harris, Douglas N. "Clear Away the Smoke and Mirrors of Value-Added." *Phi Delta Kappan*. 98, 1 (May 2010): 66-69. Accessed through JSTOR on 7 June 2010.
- Houston Independent School District. The Next Generation of ASPIRE. Houston, TX. 2010. Accessed at <http://portal.battelleforkids.org/ASPIRE/Home.html?sflang=en> on 2 June 2010.
- Hull, James. *Growth Models: A Guide to Informed Decision Making*. The Center for Public Education. Alexandria, VA. 9 November 2007. Accessed at [http://www.centerforpubliceducation.org/site/c.kjJXJ5MPlwE/b.3619055/k.8D9A/Growth\\_Models.htm](http://www.centerforpubliceducation.org/site/c.kjJXJ5MPlwE/b.3619055/k.8D9A/Growth_Models.htm) on 13 June 2010.
- Jennings, Jennifer L. and Sean P. Corcoran. "Beware of Geeks Bearing Formulas: Reflections on Growth Models for School Accountability." *Phi Delta Kappan*. 90, 9 (May 2009): 635-639. Accessed through JSTOR on 7 June 2010.
- Meyer, Robert and Emin Dokumaci. *Value-Added Models and the Next Generation of Assessments*. Center for K-12 Assessment & Performance Management. 2010. Accessed at <http://www.k12center.org/rsc/pdf/MeyerDokumaciPresenterSession4.pdf> on 24 July 2010.
- Milken Family Foundation. *Understanding the Teacher Advancement Program*. Santa Monica, CA. 2009. Accessed at [http://www.tapsystem.org/pubs/tap\\_brochure.pdf](http://www.tapsystem.org/pubs/tap_brochure.pdf) on 24 May 2010.
- North Carolina. *Race to the Top Grant Application*. Section D.2.iii. Raleigh, NC. 2010.
- North Carolina Department of Public Instruction. *ABCs/AYP 2009 Accountability Report Background Packet*. Raleigh, NC. 2009. Accessed at <http://www.ncpublicschools.org/docs/accountability/reporting/abc/2008-09/backgroundpacket.pdf> on 9 June 2010.
- North Carolina Department of Public Instruction. *Evolution of the ABCs. Accountability Services*. Raleigh, NC. August 2009. Accessed at <http://www.ncpublicschools.org/docs/accountability/reporting/abc/2008-09/abcevolution.pdf> on 7 June 2010.
- O'Malley, Kimberly. "Review of Student Growth Models Used by States." *Pearson Measurement Education Bulletin*. 2008.
- Rothstein, J. *Student Sorting and Bias in Value Added Estimation: Selection on Observables and Unobservables*. (Working Paper No. 170). National Bureau of Economic Research. Cambridge, MA. 2008.
- Simmons, Deborah. "D.C. Teacher Contract Includes Merit Pay." *The Washington Post*. Washington, D.C. 8 April 2010. Accessed at <http://www.washingtontimes.com/news/2010/apr/08/merit-pay-deal-included-in-dc-teacher-contract/> on 9 June 2010.
- Technical Advisory Panel for the Longitudinal Analysis of Student Assessment. *Colorado's Academic Growth Model: Report of the Technical Advisory Panel for the Longitudinal Analysis of Student Assessment Convened Pursuant to Colorado HB 07-1048*. State of Colorado. Denver, CO. 13 February 2008. Accessed [http://www.cde.state.co.us/cdeassess/documents/res\\_eval/FinalLongitudinalGrowthTAPReport.pdf](http://www.cde.state.co.us/cdeassess/documents/res_eval/FinalLongitudinalGrowthTAPReport.pdf) on 20 May 2010.
- Thum, Yeow M., and Anthony S. Byrk. "Value-Added Productivity Indicators." *In Grading teachers, grading schools: Is student achievement a valid evaluation measure?* Ed. Jason Millman. Thousand Oaks, CA: Corwin Press, 1997: 100-110.
- United States Department of Education. *Archived: Growth Models: Ensuring Grade-Level Proficiency for All Students by 2014*. Washington, D.C. April 2008. Accessed at <http://www2.ed.gov/admins/lead/account/growthmodel/proficiency.html> on 7 June 2010.
- Williamson, Gary. Personal communication. 27 May 2010.



By Christopher A. Cody, Joel McFarland, J. Eric Moore, and Jennifer Preston

The Financial and Business Services Area is in its fourth year of the Research Intern Program. The Program is designed to help build a quality research program within DPI to supplement and supply data for discussions related to procedural, process, and policy changes. This year's program included students from the Duke University master's program in Public Policy, a graduate of the University of North Carolina at Chapel Hill master's program in Public Administration, and a doctoral student from North Carolina State University in Public Administration. The intern program is managed by Christi Chadwick (919-807-4029) and Kayla Siler (919-807-3824) | [intern\\_research@dpi.state.nc.us](mailto:intern_research@dpi.state.nc.us)

NC DEPARTMENT OF PUBLIC INSTRUCTION :: **June St. Clair Atkinson, Ed.D., State Superintendent** :: 301 N. Wilmington Street :: Raleigh, NC 27601-2825  
In compliance with federal law, NC Public Schools administers all state-operated educational programs, employment activities and admissions without discrimination because of race, religion, national or ethnic origin, color, age, military service, disability, or gender, except where exemption is appropriate and allowed by law.

Inquiries or complaints regarding discrimination issues should be directed to: Dr. Rebecca Garland, Chief Academic Officer  
Academic Services and Instructional Support :: 6368 Mail Service Center, Raleigh, NC 27699-6368 :: Telephone: (919) 807-3200 :: Fax: (919) 807-4065