Committee on Education and Labor
Subcommittee on Early Childhood, Elementary and Secondary Education
Hearing:
"Examining Local Perspectives on No Child Left Behind"
Radnor High School
Radnor, PA: May 14, 2007

THE IMPORTANCE OF INCLUDING GROWTH MODELS IN NO CHILD LEFT BEHIND

Testimony by
Theodore Hershberg
Professor, Public Policy and History
Director, Center for Greater Philadelphia and Operation Public Education
University of Pennsylvania

THE IMPORTANCE OF INCLUDING GROWTH MODELS IN NCLB

Introduction

The best interests of the nation will be served by including growth models in the reauthorization of No Child Left Behind. They provide a much fairer way than the status (AYP) model used in NCLB to measure the performance of schools that differ greatly in their socioeconomic and demographic profiles. These models produce unprecedentedly valuable diagnostic data to help teachers improve their instruction and to help principals deploy teachers more strategically. They add a fair and accurate empirical component to improve the evaluation, remediation and compensation of educators, both teachers and administrators. Finally, because they provide a direct measure of teacher effectiveness — what students learn rather than an indirect measure such as whether a teacher is experienced or certified — they can better identify highly qualified teachers. I believe historians of school reform at the turn of the 21st century will identify growth models as the most important analytic breakthrough of the era.

1. **The Difference Between *Achievement* and *Growth***

Achievement describes the levels attained by students in their end-of-year tests. Whether referred to as proficiency, status, absolute or raw scores, these points on a vertical scale at a single point in time are best predicted by family background (income and values about education) (see Figure 1a).

G*rowth*, in contrast, describes the progress made by each student over the course of the school year and is best predicted by the quality of instruction (see Figure 1b). Good instruction is 15-20 times more powerful than family background and income, race, gender, and other explanatory variables in predicting student progress or growth.

When James Coleman (1966) and Christopher Jenks (1972) issued their famous studies concluding that the level of academic *achievement* is determined largely by factors beyond a school's control, they did not have the computer technology to permit the tracing of individual students over time nor the data sets to record their test scores in every subject and link this data to the teacher(s) who taught them

With this new technology and growth models, we now have a fair and accurate way to include student-learning results in educator evaluation, remediation and compensation (discussed separately below).

2. Using Achievement and Growth to Understand Shortcomings in No Child Left Behind

At the heart of this problem is that AYP focuses on *achievement* to the exclusion of *growth*. The four cells in Figure 2 help us identify and understand AYP's deficiencies. *Proficiency* (achievement), high and low, is tracked on the vertical axis, while *growth*, high and low, is tracked on the horizontal axis.

In the *bottom left* cell are schools that are clearly not serving the needs of their students — providing them with low proficiency and low growth — and thus deserve to be sanctioned.

Schools in the *top right* cell are performing wonderfully. They are doing what we want all schools to do: provide their students with both high proficiency and high growth. I think of this cell as responding to the challenges Tom Friedman identified in *The World is Flat*. Yet NCLB does nothing to encourage schools to reach these goals other than the absence of sanctions.

Schools in the *top left* cell are meeting their AYP goals — that is, they have high achievement — but low growth. Most often found in affluent communities where high-test scores go hand-in-hand with family income, these schools are often called "slide and glide" because they appear to be resting on the laurels of their students. It is important to understand that NCLB does nothing to hold these schools accountable for providing their students with the annual growth to which they are entitled. In a global economy characterized by fierce competition for demanding jobs that pay high salaries and benefits, this is a highly significant shortcoming.

Schools in the *bottom right* cell create high growth, but low achievement. They have succeeded in academically "stretching" or "growing" their students, but given how far behind these students were when they entered school, they have not yet been able to raise them to proficiency. These schools, while not bringing their students to AYP-required levels, are clearly helping students improve their academic performance, yet still face sanctions under current law.

NCLB reauthorization should remedy the shortcomings I have addressed here by embracing the philosophy of growth: all children, regardless of whether they are low, average, or high achieving, deserve *a year's worth of growth in a year*. Schools should be rewarded or sanctioned based on this principle.

3. Growth Models Provide Invaluable Diagnostic Information and Enable New Approaches to Educator Evaluation, Remediation and Compensation

In order to track student growth, states must have data systems that include a unique identifier for each student and each teacher and to record for every student the test scores in each grade and subject and the teacher(s) who taught them. NCLB reauthorization should mandate or provide incentives for states to develop such systems.

When collected at the classroom level, the data have uniquely powerful diagnostic value that reveal the *focus* of a teacher's instruction (on previously low-, average- or high-achieving students) and the *impact* of their instruction (highly effective, effective or ineffective). When students have two or three consecutive teachers from the last of these categories, they never reach the absolute level of accomplishment they would have achieved had they had teachers from the top two categories. When principals are provided with these diagnostic data, they can deploy their teachers so that students are never exposed sequentially to ineffective teachers.

Growth models can also make an important empirical contribution to teacher evaluation, remediation and compensation. As recent reports from RAND, the National Association of State Boards of Education and the Educational Testing Services (ETS) make clear, growth models can be used to identify the highest and lowest performers, but should never be used as the sole or principal criterion of teacher effectiveness. The data yielded by growth models should be used as part of a balanced system (*inputs,* or observation, and *outputs,* or student learning results), with

multiple measures such as those contained in the sophisticated teaching frameworks developed by Charlotte Danielson covering planning and preparation, classroom environment, instruction and professional responsibilities, as well as appropriate *safeguards,* such as review panels composed of teachers and administrators, to ensure fair treatment for individual educators.

The Congress should also add to the definition of a "highly qualified teacher" those identified as effective by growth models — that is, the lack of credentials notwithstanding, the fact that the students in their classrooms are learning at appropriate levels should be sufficient to earn the "highly qualified" designation.

4. Fixing AYP Without Abandoning Proficiency Through "Growth to Standards"

The essence of the "Growth-to-Standards" approach is to identify schools that are putting their students on growth trajectories to reach proficiency in the future and to credit these schools for that achievement.

Schools could do this by using a growth model that converts the static achievement scores of their students to dynamic growth scores. If students currently performing below their AYP targets are on track to reach proficiency by the time they graduate, they would be counted among those meeting their AYP target in the current year. If a school were to place enough of these students on growth-to-standards trajectories, it could meet its AYP goal for the year. Using a growth-to-standards approach, in other words, would reduce the proportion of schools failing AYP, but without abandoning the commitment to proficiency.

This approach may be criticized for the same reason that the existing definition of AYP is criticized: it creates what many call a "perverse incentive" for educators to focus like a laser beam on one group of students to the exclusion of all others: those close to but below proficiency. Schools choose to ignore students far below proficiency as well as those whose scores already exceed proficiency, the argument goes, because the prime directive in NCLB is for schools to hit their annual AYP targets.

While this is clearly the logic of the incentive, we do not yet know if this is supported in fact. The growth-to-standards approach described above, like AYP, might simply illuminate the pattern — the gains made by those who start just below proficiency are coming at the expense of those who start the year above it — rather than exacerbate it.

We know this pattern long pre-dates NCLB and has been widespread in poor communities, whether in inner-cities or Appalachia. It explains, for example, the observation made by elementary school teachers that the proportion of precocious students in kindergarten and first grade is sharply reduced by fifth and sixth grades. Faced with so many low performing children, the explanation goes, teachers focus on the bottom of the student distribution so that previous low-achievers get high growth while previous high-achievers get low growth. Sustaining this focus in the early years explains why so few high achieving, low-income children are found in middle school.

When Dr.William Sanders applied his growth-to-standards approach to all Tennessee schools in the 2002-03 school year, he learned that 13 percent more schools would meet their federal goals if this alternative means of calculating AYP were accepted by the U.S. Department of Education. But when Sanders looked more closely at its effects — he examined nine Memphis schools all of whose students were minority and low-income (on free and reduced price lunch) — he discovered some troubling results. While some schools met their AYP through the growth-to-standards alternative without denying any of their students adequate yearly growth, others did so at the expense of students who had achieved at higher levels in the past. Seeing no sense in a trade-off that benefits one group of poor minority kids at the expense of another, Sanders proposed a "net" approach: schools would receive credit for students placed on a growth-to-standards trajectory and debits for formerly higher achieving students denied adequate growth in the process.

The U.S. Department of Education has given approval to Tennessee to use this approach in determining if schools meet their AYP goals. NCLB should provide incentives to expand the use of this model in other states.

6. A Cautionary Note: Not All Growth Models Are Equal
This is not the place to discuss the complex statistical issues embedded in the use of different kinds of growth models, such as the "projection" model used in growth-to-standards or the "expectations" model used to evaluate the effectiveness of individual teachers. Some models that are described as "simple and transparent" are actually statistically flawed and will yield specious and erratic results. Suffice it to say that much attention must be paid to the details in order for growth models to be used fairly and effectively.

ATTACHMENTS

Figure 1a: Achievement
Figure 1b: Growth
Figure 2: Identifying AYP's Shortcomings

Appendix: Theodore Hershberg, "Value-Added Assessment and Systemic Reform: A Response to the Challenge of Human Capital Development," *Phi Delta Kappan* (December, 2005). Paper prepared for the Aspen Institute's Congressional Institute, *The Challenge of Education Reform: Standards, Accountability, Resources and Policy* (Cancun, Mexico: Feb. 22-27, 2005).