Collective Bargaining & Member Advocacy

nea
NATIONAL
EDUCATION
ASSOCIATION
nea.org
Great Public Schools
for Every Child

1201 16th St., N.W. | Washington, DC 20036-3290 | Phone: (202) 822-7080 | Fax: (202) 822-7833

WHAT YOU SHOULD KNOW ABOUT VALUE-ADDED ASSESSMENT

A Value-Added model (VAM) is one type of growth model in which states or districts use student achievement data, and sometimes other student background data, as statistical controls in order to isolate the specific effects of a teacher, school, or program on student academic progress. VAMs aim to answer the question: On average, did the students' change in performance meet the growth expectation [based on what can be gleaned from his/her past performance]? (Goldschmidt, 2005)

The adequate yearly progress (AYP) goals and associated sanctions established under the No Child Left Behind Act of 2001 have spawned tremendous interest in statistical models that evaluate teacher effectiveness on the basis of student test scores. However, growth models are complex, have rigorous data requirements, and require a degree of human resources and psychometric expertise that most states and districts are not equipped with. It is equally disconcerting that there are still several unanswered questions about the reliability and validity of these models and the estimates they produce. While there is no doubt that VAMs have the potential to provide some insight into student achievement, until these questions can be answered satisfactorily, it is irresponsible to utilize teacher effect estimates for high stakes decisions related to teacher performance or pay.

No consensus in the research community—There is no consensus in the research community as to whether VAMs can accurately isolate the effects of a single teacher. "Performance measures based on student achievement data are meant to measure teachers' contributions to student achievement as opposed to other attributes of students or other aspects of teacher performance…However, measuring student achievement and growth consistently and precisely is at best difficult, if not infeasible." (Buddin, R., 2007)

Teacher input may be an important influence on achievement, but it is not the only influence
"[T]reating the output of a value-added analysis as an accurate indicator of a teacher's relative contribution to student learning is equivalent to making a causal interpretation of a statistical estimate…In the absence of randomization, causal interpretations can be misleading. In reality, the classroom placement of students and teachers is far from random. In most districts, parents often influence where their children go to school and even to which class and teacher they are assigned. Similarly, teachers may select the school and classroom where they are placed. Thus, the students assigned to a particular teacher may not be representative of the general student population with respect to their level and rate of growth in achievement, parental support, motivation, study habits, interpersonal dynamics and other relevant characteristics. It is very difficult for the statistical machinery to disentangle these intrinsic student differences from true differences in teacher effectiveness." (Braun, 2005)

Other issues that may impact estimations of teacher effectiveness
- **Missing test data** – VAMs rely on student test scores across multiple years.
- **Timing of tests** – Test scores are compared from year to year so it is possible that a student may have experiences during summer recess that impact achievement estimates.
- **Contributions of the prior year's teacher/school** – Studies are inconclusive about how long a teacher's effect lasts. Estimations of current teacher effectiveness may be impacted by the lasting or shrunken effects of an earlier teacher.

- **Differences in what tests measure** – If the statistical method ignores the fact that different tests (in the same subject area) may measure different skills, the resulting estimates of teacher performance may contain errors that favor some teachers over others. (Martineau, 2006)

VAMs may work against collaborative teaching —
Educational communities that value collaborations, team teaching, interdisciplinary curricula, and promote student autonomy and active participation in educational decisions may find little use for such information. A model that regards teachers as isolated, independent actors and students as passive recipients of teacher "effects" may not be adequate in some contexts. When the fit between the model and the phenomenon it seeks to represent is poor, validity is threatened. (Kupermintz, 2003)

Using achievement measures may encourage an unbalanced approach to teaching
Research on the response of schools and teachers to the requirements of the federal No Child Left Behind Act shows that educators report focusing more attention on reading and mathematics than untested subjects and focusing their attentions on students just below proficiency. (Hamilton. L, 2007)

The most popular VAM is also one of the most problematic —
The Educational Value Added Assessment System (EVASS), operated by Dr. William Sanders, is the oldest and most popular value added model in use. The EVASS model assumes that teacher effects are constant over time and does not adjust for student background characteristics. EVAAS does not attempt to model the interaction between where a school starts and how much it grows. There are tremendous data requirements (linked student data over time) as well as tremendous computing capacity requirements. Further, due to proprietary estimation procedures, broad applications of this model independently by states are not possible. Hence cost is an additional factor. Further, using models that contain complex (and proprietary) computations which are inaccessible to stakeholders may make it harder to build consensus and a sense of confidence around the validity of the results. (Goldschmidt, 2005)

VAMs may be expensive to adopt
- These models utilize multiple years of test data that must be comparable from year to year.
- Contracting to acquire psychometric expertise is often necessary. There are challenging technical issues to be resolved in creating a growth formula that meets a state's policy needs.
- Data system requirements need to be addressed. To measure growth, there must be a capacity to track individual student scores from one year to the next (and sometimes from one district to another in the state). This capacity often requires a statewide student identification system.
- Training is required to build capacity among the teachers, administrators, media, legislators, and general public to understand the additional complexities that occur when using data from more than one point in time. Even changing to the simplest of growth models will require a significant retooling of training materials. (Goldschmidt, 2005)

VAM numbers don't say much about achievement —
"These analyses treat the classroom as a 'black box'… and do not tell us why some classrooms are more effective than others, nor do they give us a very good picture of the potential improvements in student achievement that might be produced if we combined particularly effective instructional conditions into powerful instructional programs. (B. Rowan, 2002)."

Researchers have warned against using value added estimates for high stakes purposes
"VAM results should not serve as the sole or principal basis for making consequential decisions about teachers. There are many pitfalls to making causal attributions of teacher effectiveness on the basis of the kinds of data available from typical school districts. We still lack sufficient understanding of how seriously the different technical problems threaten the validity of such interpretations. (Braun, 2005)"

Braun, H., *Using Student Progress to Evaluate Teachers: A Primer on Value Added Models,* Technical report, Princeton, NJ: Educational Testing Service, Policy Information Center, 2005. (Available at: www.ets.org/Media/Research/pdf/PICVAM.pdf)

collectivebargaining@nea.org **Fact Sheet**

Buddin, R., D. McCaffrey, S. Nataraj Kirby, and N. Xia. *Merit Pay for Florida Teachers Design and Implementation Issues,* Santa Monica, CA: RAND Corporation, *Working Pape*r WR-508-FEA, August 2007.

Goldschmidt, P., et al, *Policymakers' Guide to Growth Models for School Accountability: How Do Accountability Models Differ*? The Council of Chief State School Officers, Washington, D.C., October 2005.

Hamilton, L., B. Stecher, J. Marsh, J. McCombs, A. Robyn, J. Russell, S. Naftel, and H. Barney, *Standards Based Accountability under No Child Left Behind: Experiences of Teachers and Administrators in Three States,* Santa Monica, CA: RAND Corporation, MG-589-NSF, 2007.

Kupermintz, H. "Teacher effects and teacher effectiveness: A validity investigation of the Tennessee Value Added Assessment System'," *Educational Evaluation and Policy Analysis,(25)*3, 287-298, 2003.

Martineau, J., "Distorting Value Added: The Use of Longitudinal, Vertically Scaled Student Achievement Data for Growth Based Value Added Accountability," *Journal of Educational and Behavioral Statistics,* Vol. 31, No. 1, 2006.

Rowan, B., R. Correnti, and R. J. Miller, "What Large-Scale Survey Research Tells Us About Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools," The Consortium for Policy Research in Education, November 2002.

To learn more about Value Added Assessment, contact the NEA Collective Bargaining and Member Advocacy Department at 202-822-7080 or mailto:collectivebargaining@nea.org.
October 2007