

A History of Value-Added Assessment and its Research-Based Implications for Policy and Practice on Teaching Effectiveness

Introduction

Over two decades ago, a small team of statisticians and researchers at the University of Tennessee, Knoxville initiated a new way to view the effectiveness of educators. Rather than focusing on the *achievement level* of students as a measure of effectiveness, the future SAS EVAAS team focused on the *progress* of students over time, following each individual student across subjects and grades. While the application to education represented a paradigm shift for educators and policymakers, the models themselves drew upon established statistical models, which overcame many non-trivial analytic complications concerning the use of student testing data to assess educators' effectiveness. Over the years, the EVAAS value-added approach—and the conclusions drawn from its research—have been reviewed, validated and confirmed by a variety of public and private sector experts.¹ This document summarizes key milestones and findings regarding value-added assessment.

The Early Years (1982 – 1999)

Led by Dr. Bill Sanders at the University of Tennessee, Knoxville, these early years focused on research that established many of today's basic beliefs about teaching effectiveness, such as the importance that teachers have on their students' progress and the fact that robust value-added estimates are not related to student characteristics.

1982: The First Study²

Key findings included:

- Differences in teaching effectiveness have a highly significant effect on the rate of student academic progress.
- The effects were greater in Math than in Reading Comprehension.
- Differences between highly effective classrooms and very ineffective classrooms, as measured by EVAAS multivariate longitudinal analyses, were very highly related to curricular specialists' classroom observations.

1984: The Second Study³

Key findings included:

- Principals forecasted most of the bottom-profiling teachers. For math (but not reading and language arts), they could distinguish between top and average teachers as well.
- Student IQ scores did not improve the precision of the school and classroom value-added estimates when all of each student's achievement scores were included in the analysis.

¹ See, for example: Lockwood J.R. and D.F. McCaffrey (2007). "Controlling for individual heterogeneity in longitudinal models, with applications to student achievement." *Electronic Journal of Statistics*, Vol. 1, p. 244. Also see: : McCaffrey, D. F., Han, B. and Lockwood, J. R. (2008). "Value-Added Models: Analytic Issues." Prepared for the National Research Council and the National Academy of Education, Board on Testing and Accountability Workshop on Value-Added Modeling, Nov. 13-14, 2008, Washington D.C.

² Reference: McLean, R.A., & Sanders, W.L. (1984). *Objective component of teacher evaluation: A feasibility study*. Working Paper No. 199. Knoxville: University of Tennessee, College of Business Administration.

³ Cook, David L. (1985) *An objective component for the evaluation of teaching*. Dissertation. Knoxville: University of Tennessee.

1986: The Third Study

Key findings included:

- No significant statistical interaction between race of teacher and race of student.
- White and black students both responded to teachers with high value-added measures regardless of their race.
- The ethnic composition of a school was a poor predictor of the *effectiveness* of a school.

1993: District Value-Added Reports

Available for the *first time in the nation*, district-level value-added reports are released to *all* Tennessee districts. With this release, it is possible to confirm that there is virtually no relationship between the ethnic composition of school districts and measures of cumulative academic growth.

1994: Students Changing Buildings⁴

Key findings included:

- When populations of students change buildings, there is a major impact on academic growth the *following* year.
- This was found to be true, regardless of the grade in which the building change occurred.
- The impact was found to be minimal when individual students changed to a different school in grades that were not the lowest grade of the receiving school.

1994: School Value-Added Reports

Available for the *first time in the nation*, school-level value-added reports are released to *all* Tennessee schools for grades 3 – 8. Supporting diagnostic reports were also made available, which enabled local educators to ascertain if appropriate academic growth was being realized for high-, middle- and low-achieving students.

1996: Teacher Value-Added Reports and another Seminal Study

Available for the *first time in the nation*, teacher value-added reports are released statewide to Tennessee teachers based on testing in grades 3 – 8 and high school math. Additionally, district and school high school math reports were released to all Tennessee districts and schools.

This release enabled a groundbreaking study on **Cumulative and Residual Teacher Effects**,⁵ which is still cited extensively today. Key findings included:

- Teacher effects were found to be cumulative and additive with very little evidence of compensatory effects. In other words, if a student had two very ineffective teachers in a row for the same subject, then there is very little evidence that a subsequent teacher could compensate for that loss in progress.
- An effect of a classroom could be measured – either positively or negatively – on future student achievement.

⁴ Sanders, W.L., A.M. Saxton, J.F. Schneider, B.L. Dearden, S.P. Wright, and S.P. Horn (1994). "Effects of Building Change on Indicators of Student Academic Growth." *Evaluation Perspectives*, 4(1), 3, 7.

⁵ Sanders, William L., and June C. Rivers (1996). *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement*. Knoxville: University of Tennessee Value-Added Research and Assessment Center.

1997: Published Models⁶

- The statistical methodology underlying the multivariate, longitudinal methodology used in PVAAS was published in the open literature.

1999: Study on Sequence of Teachers⁷

- The sequence of teachers that a student had *greatly* affected the probability of a student passing a high stakes test.

Since The Early Years (2000 – Present)

Moving from a university to SAS Institute Inc. enabled the EVAAS team to expand its services beyond the State of Tennessee and to improve its delivery to include a user-friendly, hosted web application, which could provide a variety of reports beyond the value-added estimates to support educators and policymakers. In the early years of No Child Left Behind and then subsequently with Race to the Top, there is an ever-growing awareness and importance placed on identifying effective teaching.

In these years, the SAS EVAAS team has sustained its roots in research. This research has continued to provide insights from the earlier findings as well as break new ground.

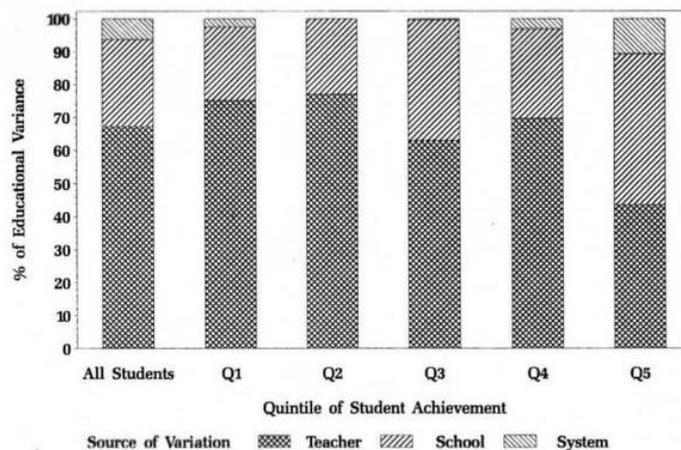
Importance of Educational Entities

The relative importance of educational entities on the rate of student academic progress can be thought of in three levels:

- Districts OR
- Schools within Districts OR
- Classrooms within Schools within Districts?

As illustrated below in Figure 1, most of the differences in the rate of student progress can be attributed to the third entity: classrooms within schools within districts. This reinforces the importance that teachers have on their students' academic opportunities.

FIGURE 1: RELATIVE PERCENTAGE OF EDUCATIONAL VARIANCE IN FIFTH GRADE MATH



⁶ Sanders, W. L., Saxton, A. M., and Horn, S. P. (1997). *The Tennessee Value-Added Accountability System: A Quantitative, Outcomes-Based Approach to Educational Assessment*. Pages 137-162 in J. Millman (Ed.), *Grading Teachers, Grading Schools: Is Student Achievement a Valid Evaluation Measure?* Thousands Oaks, CA: Corwin Press.

⁷ Rivers, June C. (1999) *The Impact of Teacher Effect on Student Math Competency Achievement*. Unpublished doctoral dissertation, University of Tennessee Knoxville.

Teaching Effectiveness

Continued research on teaching effectiveness has yield important insights on a variety of characteristics:

Years of Teaching Experience

- Teacher effectiveness is related to years of service.
- On average, there is measureable improvement for up to 10 years.
- Other researchers have found improvement only up to 3 years.
- Teachers who leave after 1 or 2 years of experience are *less* effective than those who stay.

Academic Record

- For lower grade elementary teachers, very little predictability of their effectiveness has been found from their own college academic record.
- It was found that there are measurable relationships between teachers' academic records, as students, with their effectiveness as middle and high school math teachers.

Building Changes⁸

- When teachers change buildings, the effectiveness measured in the school before the move was found to be similar to the effectiveness measured after the move.
- The effectiveness measured the second year after the move was found to be more highly correlated with the effective measurement prior to the move than the measurement the first year after the move.

Stability⁹

- Value added estimates, based on three years of data for beginning teachers, indicate that highly effective teachers will tend to be effective 3 – 5 years later.
- Only about one half of beginning ineffective teachers will become average teachers in the future.
- With our methodology, the repeatability correlation is about a 0.70 – 0.80 for three year teacher value-added estimates

Projections to Future Tests

- Individual student projections offer an opportunity to minimize inequities that often occur in student placement to more advanced courses.
- It has been found that the use of these projections have identified students who were successful in a more advanced course, yet traditionally would not have been placed in these courses.
- Projections being made to all types of tests
 - Measures of college readiness, such as the SAT/PSAT, ACT/PLAN/EXPLORE, and Advanced Placement tests
 - National tests, such as NRTs
 - State tests, such as CRTs and end-of-course tests

⁸ Sanders, W. L., Wright, S. P., and Langevin, W. E. (2009). The Performance of Highly Effective Teachers in Different School Environments. Chapter 8 in M. G. Springer (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*. Washington, D. C.: Brookings Institution Press.

⁹ : White, J. T., Wright, S. P., and Sanders, W. L. (2011) Unpublished report.

SAS Projects and Research Efforts

The interest in gaining insight to teaching effectiveness continues beyond the traditional value-added estimates for districts, schools and teachers. SAS collaborates with a variety of its education partners to further knowledge in this field. Examples of recent EVAAS projects and research efforts include the following:

Program Evaluation

- Striving Readers
- EAP tutoring program
- SES tutoring programs
- Teacher Advancement Program (TAP)

Linking teacher effectiveness to teacher preparation programs

- Annual Higher Education Report Card in Tennessee
- Exploratory analysis on what other measures contribute to teacher effectiveness

Relationship between value-added and student surveys

- Pilot project in a large North Carolina district currently underway

Relationship between value-added and expenses

- Per pupil expenditure for districts
- Program costs

Relationship between value-added and teacher retention rates

- Which teachers stay and which teachers leave

Relationship between value-added and the special education population

- Exploratory analysis in Pennsylvania in development.