

Value-Added Assessment Annotated Bibliography

- Adcock, E. P. (1995). *Value-added effective schools study for elementary schools: 1994 Maryland school performance assessment program results*. Research Report No. 36-9 95. Maryland: Prince George's County Public Schools, Research, Evaluation and Accountability.
This reports provides the results of a value-added study of the effectiveness of 119 Prince George's County elementary school's reading and mathematics programs.
- Andrejko, L. (2004). Value-added assessment: a view from a practitioner. *Journal of Educational and Behavioral Statistics*, 29(1), 7-9.
A school superintendent describes the use of the information from the Pennsylvania Value-Added Assessment System (PVAAS) in her school district.
- Astin, A.W. (1982). *Excellence and equity in American education*. Washington, D. C.: National Commission on Excellence in Education. (ERIC Document Reproduction Service No. ED 227 098)
Discusses the concept of value-added assessment and how it can promote equity in education by diverting attention away from mere acquisition of resources, focusing instead on their effective utilization.
- Bixler-Stoudt, T., Bohan, J., Lewald, K. & McCartney, P. (2005). The pursuit of a data informed school culture in Pennsylvania. *The Pennsylvania Administrator*, 8 (3), 15-1.
Outlines the various types of data available for school districts to use in Pennsylvania and formats to analysis said data. Included is a discussion of the use of value-added data (PVAAS) in schools.
- Bock, R.D., Wolfe, R. & Fisher, T. (1996). *A review and analysis of the Tennessee Value-Added Assessment System*. Nashville, TN: Comptroller of the Treasury. www.comptroller.state.tn.us/orea/reportstvaasp1.pdf
A review of TVAAS by outside reviewers found the system to be fundamentally sound and consistent with other hierarchical models in widespread use. The completeness of the data was found adequate, although improvements could be made.
- Bratton, S., Horn, S. & Wright, S. Paul. (1996). *Using and Interpreting Tennessee's Value-Added Assessment System*. Knoxville: University of Tennessee.
This paper focused on how teachers and principals may use the results of Tennessee's Value-Added Assessment System (TVAAS)/ Tennessee Comprehensive Assessment Principle (TCAP), and also listed some basic principles of the TVAAS statistics and of achievement testing. The paper is divided into four parts. First, Evaluating and Using TVAAS Results, guides readers' in interpreting results and putting them to good use. Second, Basic Principles of TVAAS talks about how TVASS works and what makes it new and different. Third, Standardized Testing and Alternative Assessments deals with educational evaluation issues that has more to do with assessment of learning in general. Finally, the last part covers an assortment of topics.

PENNSYLVANIA
Value-Added
Assessment System



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org

BIBLIOGRAPHY

Bryk, A.S., Thum, Y.M., Easton, J.A., & Luppescu, S. (1998). *Academic productivity of Chicago Public elementary schools*. Chicago: Consortium on Chicago School Research.

In the context of Chicago school reform, the productivity of individual schools cannot be measured fairly using system-wide reports of average test scores. The authors develop a school productivity profile that estimates the value a school adds to learning in a year. The initial study summarizes trends in reading. Subsequent studies will examine schools that appear to be especially effective.

California Department of Education, Office of Policy and Evaluation, Special Studies and Evaluation (1998).

The California Department of Education summarizes the value-added accountability models in use in Tennessee and in Dallas, and discusses their potential for California schools. The Tennessee system (TVAAS), was implemented in 1992 with support by the Tennessee Business Roundtable. Three hallmarks of the statistical model are that it controls for confounding variables by allowing each student to serve as his or her own control, it accommodates but does not over-react to missing data, and it protects against misclassification, particularly when there are few scores, by assuming that all teacher or school effects are an average of their school system until the weight of the data pulls their specific estimates away from the mean. Two reviews of the model commissioned by the Tennessee Comptroller have largely confirmed the soundness of the methodology. The Dallas system, also in use since 1992, uses a two-stage model. Stage one regression controls for confounding influences, and stage two hierarchical linear modeling controls for school level influences. Reviewers suggest that this system may not accurately account for confounding factors, and does not account well for missing data or regression to the mean. A value-added system could be implemented in California.

Ceperley, P.E. & Reel, K. (1997). The impetus for the Tennessee value-added accountability system. In J. Millman (Ed.), *Grading teachers, grading schools: is student achievement a valid evaluation measure?* (133-136). Thousand Oaks, CA: Corwin Press.

TVAAS was at the heart of Tennessee's 1992 education reform legislation, prompted by a funding lawsuit. The Business Roundtable demanded accountability provisions for schools in exchange for business support of tax increases for school funding. The reform plan prescribed penalties for schools and districts that repeatedly failed to make gains equal to national norms.

Council of Chief State School Officers (2005). *Policymakers' Guide to Growth Models for School Accountability: How Do Accountability Models Differ?* Washington, D.C.

This paper address the potential use of growth models for school accountability and the possibility of adding a growth model to existing systems to provide additional information about educational performance of school and groups of students. For example, a growth model can have the purpose of predicting whether and when a school will meet a projected proficiency goal.

Crane, J. (2002). *The promise of value-added testing*. Progressive Policy Institute. Retrieved from www.ppionline.org

No Child Left Behind brings a new emphasis on accountability. Value-added assessment would provide a better picture than annual yearly progress for measuring progress, because of its focus on gains.

Cunningham, L. (1997). In the beginning. In J. Millman (Ed.), *Grading teachers, grading schools: is student achievement a valid evaluation measure?* (75-80). Thousand Oaks, CA: Corwin Press.

PENNSYLVANIA
Value-Added
Assessment System



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org

BIBLIOGRAPHY

The Dallas Independent School District built statistical and evaluative capacity throughout the 1960's and 1970's under a remarkably capable superintendent. William Webster, head of research and evaluation, and his staff are leaders in evaluation. Dallas is a leader in educational accountability systems. The value-added assessment system used in Dallas was built over a number of years and consistently refined to fit into (and at times guide) the state accountability scheme.

Darlington, R.B. (1997). The Tennessee value-added assessment system: a challenge to familiar assessment methods. In J. Millman (Ed.) *Grading teachers, grading schools: is student achievement a valid evaluation measure?* (163-168). Thousand Oaks, CA: Corwin Press.

In this evaluation of TVAAS, the author asks two questions: does the system work well enough to be worthwhile; and how does it compare to regression methods? He concludes that the system does work well enough, and that it is superior to regression models in most real world contexts, where significant data is missing. Regression models are well understood by educators, and thus would be preferable in an ideal school system with complete data, but no such system exists. Mixed models are not well known in education circles, being more common in agricultural statistics. Nonetheless, the mixed model seems superior because of its handling of missing data.

Dembosky, J., Pane, J. Barney, H, & Christina, R. (2006). *Data driven decision-making in southwestern Pennsylvania school districts*. Retrieved from http://www.rand.org/pubs/working_papers/2006/RAND_WR326.pdf.

This study by the RAND Corporation examines the preparedness of school districts to use data driven decision-making, including data provided by PVAAS.

Gormley, W. T., Jr., & Weimer, D. L. (1999). *Organizational report cards*. Cambridge, MA: Harvard University Press.

Offers the first comprehensive study of organizational reports. It discusses the circumstances under which they are desirable alternatives to other policy instruments, how they should be designed, who is likely to use them and for what purpose, and what role if any, government should have in their creation.

Graphical summary of educational findings from the Tennessee Value-Added Assessment System (TVAAS) (1997). Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Compilation of research findings presented in a series of graphs organized in 1) answer to frequently asked questions about factors thought to influence student academic gain; 2) summarize statewide academic gain trends; and 3) address the problem of variability.

Graphical summary of educational findings from the Tennessee Value-Added Assessment System (TVAAS) (1995). Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Similar to graphical summary, 1995 but adds a section for interpreting TVAAS reports.

Harris, J. (2006). Using PVAAS data to improve student achievement. *Pennsylvania Educational Leadership*, 26 (1), 63-69.

Discusses the use of PVAAS data by the district data team. The author concludes that by using PVAAS data along with other school data, the team was able to better identify the academic growth of their students and respond accordingly.

Harville, D. A. (1976). Extension of the Gauss-Markov theorem to include the estimation of random effects. *The Annals of Statistics*, 4, 384-395.

PENNSYLVANIA
Value-Added
Assessment System



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org

BIBLIOGRAPHY

Harville, D. A. (1995). A review of the Tennessee Value-Added Assessment System (TVAAS). Unpublished manuscript: Iowa State University. Describes and reviews methodology and application of TVAAS.

Hippert, L. (2007). Measuring student growth. *PASA Flyer*, 49 (4), 1-6.

This article presents an overview of PVAAS and its use in helping schools look at student growth.

Ladd, H., & Walsh, R. (2002). Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review*, 21(1) 1-17.

Value-added systems must distinguish the effectiveness with which schools operate, controlling for differing resources. North and South Carolina school effectiveness measures are based on overall gains in test scores, and do not accurately account for how efficiently a school is operated because they include effects outside the control of school personnel. While these measures provide useful information, because they do not control for school resources, including the student mix, that are outside the control of schools, they should be used cautiously for accountability purposes. Results show that schools serving higher performing students are more likely to be deemed effective than schools serving lower performing students. The authors determine that two-fifths of this correlation is due to measurement error. However, even correcting for measurement error, effectiveness remains correlated with socio-economic status and average test scores. The likely result is that the best teachers will be drawn to the better performing schools.

Lissitz, R.W. (2005). *Longitudinal and Value-Added Modeling of Student Performance*. Maple Grove, MN: JAM Press.

This book is based on the very well received conference of the same name held on the University of Maryland Campus on November 7 and 8, 2005. This book presents a variety of chapters regarding the theory and application of Longitudinal (Growth) Modeling and Value Added determinations of Student Achievement. This book is found to be stimulating to academics, psychometrics personnel, as well as to school practitioners who are concerned with the monitoring of student performance across time and the organization of schools to utilize this information to encourage maximizing student performance across time. Concerns include statistical theory, estimation issues, and a variety of approaches to modeling that have direct application to this school problem. NCLB has emphasized the status of Cross-sectional Cohorts through the analysis of AYP measures. This is one approach to the problem of measuring school performance. This book is concerned with alternatives that will permit schools to model the performance of individual students with the hope that all students might eventually have their performance maximized as they progress through the school experience. This goal requires the field to develop new ways to measure such progress and new ideas for the use of such measures by the schools.

Lockwood, J.R., Louis, T. & McCaffrey, D. (2002). Uncertainty in rank estimation: implications for value-added modeling accountability systems. *Journal of Educational and Behavioral Statistics*, 27(3), 255-70.

The performance of rank (percentile) estimators is investigated in a basic, two-stage hierarchical model that captures the essential features of value-added systems currently in use. The authors conclude that percentile estimators do not necessarily perform well enough to be desirable for use as a basis for evaluations. Substantial information, which is not always present, is needed for acceptable performance of systems based on percentile rankings.

PENNSYLVANIA
Value-Added
Assessment System



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org

BIBLIOGRAPHY

PENNSYLVANIA
Value-Added
Assessment System



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org

- Marsh, J, Pane, J., & Hamilton, L. (2006). Making sense of data-driven decision-making in education. Evidence from recent RAND research. Retrieved from http://www.rand.org/pubs/occasional_papers/2006/RAND_OP170.pdf
This paper seeks to clarify the ways in which multiple types of data, including value-added data are being used in schools and districts.
- McAdams, D. (2002). Enemy of the good. *Education Next*, 2(2). Retrieved from www.educationnext.org.
Even though value-added analysis is flawed, the author, a former Houston school board member, is committed to it. It is better than the other measures. However, it must be applied with flexibility and supervisors must have the freedom to use their judgment. High-performing organizations measure almost everything, because measurement changes behavior. Constant measurement focuses attention. While it is obvious that testing can be misused, there is little evidence of value-added being misused, and much evidence that it focuses attention on learning.
- McCartney, P. (2004). The Pennsylvania Value-Added Assessment System (PVAAS): A Tool for Educators to Improve Instruction. *Insight*, 24 (3), 6-7.
A brief overview of the Pennsylvania Value-Added Assessment System pilot.
- McLean, R. A., Sanders, W. L. & Stroup, W. W. (1991). A unified approach to mixed linear models. *American Statistician*, 45, 54-64.
Explanation of mixed-model methodology.
- Meyer, R. (2002). *Value-added indicators: do they make an important difference? Evidence from the Milwaukee public schools*. Wisconsin Center for Education Research. Paper presented at the Annual Meeting of the American Education Research Association, New Orleans, April 2, 2002.
The Milwaukee Public Schools initiated annual assessment of students in 2001 and is now using value-added analysis to track school performance, program efficacy, and other policies. Math achievement data from 7th and 8th grades is used to illustrate the approach. The purpose of using value-added analysis is to isolate the contribution of schools from other factors, such as prior achievement and student, family and neighborhood characteristics. The value-added math productivity of Milwaukee middle schools varied widely, with some generating almost two times the growth of the average school. Most schools the served students with lower prior achievement produced better than average gains.
- Meyer, R.H. (1996). Value-added indicators of student performance. In Hanushek, E.A. & Jorgenson, D.W. (Eds.). *Improving America's Schools: The Role of Incentives* (197-224). Washington, D.C.: National Academy Press.
The author's premise is that level indicators such as average test scores, while valid for some purposes, are not valid measures of school performance. He describes various types of statistical models that can be used to determine gains in student achievement, discusses the contexts in which particular models can be effective, and examines the factors that affect reliability. He compares value-added assessments to other, more common, indicators of student achievement, and suggests that inappropriate indicators of achievement drive flawed policy decisions. He recommends that students be tested at every grade level and that other data be collected to enable value-added assessments of student gains.
- Meyer, R.H. (1997). Value-added indicators of school performance: A primer. *Economics of Education Review* 16(3), 283-301.

BIBLIOGRAPHY

PENNSYLVANIA
Value-Added
Assessment System



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org

National Association of State Boards of Education (NASBE) (2005). *Evaluating Value-Added*. Alexandria, VA.

This paper looks at the questions surrounding value-added model concepts and practical uses. It addresses and makes recommendations on issues such as the alignment between value-added and other aspects of the state's testing and accountability system, the accuracy of value-added results, the best uses for value-added data, the programs associated with value-added, and the areas states need to think about if they are planning to use value-added analysis as a component of their assessment and accountability system.

Noell, G. & Burns, J. (2006). Value-added assessment of teacher preparation: an illustration of emerging technology; *Journal of Teacher Education*, 57(1), 37-50.

Broad-based empirical outcomes assessment is an increasingly evident part of governmental services and this trend is particularly apparent in education. The clearest manifestation of this trend in education has been the advent of high-stakes, broad-based testing and accountability programs in K-12 education. Although this assessment regime has not yet been used to assess the efficacy of teacher preparation programs, the data management capacity and statistical technology is now emerging to make this possible. This article presents data from the first year of a pilot study examining the methodological and practical issues involved in implementing a value-added assessment of teacher preparation based on a massive multivariate longitudinal database. The pilot data are discussed in relation to the literature pertaining to value-added assessments in K-12 education. Selected research needs and practical concerns related to the use of value-added models for the assessment of teacher preparation are discussed.

Noell, G. (2005). Assessing teacher preparation program effectiveness: A pilot examination of value added approaches. Retrieved from www.asa.regents.state.la.us/TE/digest.pdf.

This paper is a plot work examining the feasibility of using Louisiana's educational assessment data in concert with the Louisiana Educational Assessment Data System (LEADS). The degree of matching across years and the degree of matching between the LEADS data and achievement data suggest this approach is viable. The reliability of individual level estimates of teacher efficiency across 12 months, different student groups and test forms were promising. As the number of years achievement data increased, the contribution of demographic factors rapidly decreased to low levels. Some statistically significant differences were obtained between new teacher groups and experienced certified teachers for student overcomes after controlling for prior achievement, demography variables, and classroom context variables.

O'Brien, D. & Ware. A. (2002). Implementing research-based reading programs in Fort Worth Independent School District. *Journal of Education for Students Placed at Risk*, 7(2), 167-95.

The Fort Worth Independent School district implemented a direct instruction reading program in kindergarten, 1st and 2nd grades in 61 schools. Two different programs were implemented. This article reports on the history, implementation and evaluation of the programs. Outcome evaluation used statistical analysis to look at the effect of the program on district performance, to compare schools, programs and classrooms, and to identify sources of variation in student gain. Student gains were evaluated using a value-added regression methodology. Simple comparisons of the two programs with a control group would not be adequate because of large differences in the demographics of school populations. The student level regression analysis used a spring reading test score as the independent variable. Dependent variables included the fall reading test score, race, free or reduced price lunch status, gender, special education status and English proficiency. In 50 of 60 program years, holding control variables constant, score gains for the direct instruction programs were larger than the control school gains.

BIBLIOGRAPHY

Peevely, G. & Ray, J. (2001). Does equalization litigation effect a narrowing of the gap of value-added achievement outcomes among school districts? *Journal of Education Finance*, 26(4), 463-76.

Value-added student achievement levels of mostly small, rural schools that participated in the Tennessee school finance case were compared to districts that had not been litigants. The litigants asserted that inequitable funding resulted in lower achievement levels, and implied that higher expectations required equalization of resources. When “full funding,” as mandated by the state was achieved, litigants’ budgets increased, but were still less than other districts. During the phase in of full funding, the gap between litigants and non-litigants did not significantly close, and the direction of change was mixed.

Raudenbush, S. W., & Bryk, A. S. (1988). Methodological advances in analyzing the effects of schools and classrooms on student learning. *Review of Research in Education*, 15, 469.

Roderick, M, Bryk, A., & Jacob, B. (2002). The impact of high-stakes testing in Chicago on student achievement in promotional gate grades. *Educational Evaluation & Policy Analysis*, 24(4), 333-57.

Chicago has linked school accountability with high-stakes consequences for students by requiring mandatory levels of performance on standardized tests to achieve promotion from 3rd, 6th and 8th grades. Using a three-level hierarchical linear model, the authors compare student achievement gains in grades targeted for promotion both before and after implementation of the policy, and conclude that student achievement increased substantially in grades 6 and 8 after implementation of the policy. Results for grade 3 were less conclusive. In reading, the lowest-achieving students showed the greatest gains, while in math the opposite was true.

Ross, S.M. (2001). *Creating critical mass for restructuring*. Appalachia Educational Lab., Charleston, WV. [AWT03075] Office of Educational Research and Improvement, Washington, DC. [EDD00036]

From 1995 to 1999, Memphis City Schools implemented a successful systemic reform program. However, in 2001, the program was discontinued. The author documents the success of the program in its early stages, as evidenced by observations and by gains in value-added scores. As the reform program was expanded in its later years, however, schools were added that were more resistant to change and support resources became strained. Schools that joined the reform effort later reported feeling pressure to adopt “favored” reform models. In addition, the superintendent and assistant superintendent, who had spearhead the program announced their departures. The author concludes that for any reform effort to be successful, it must be tailored to school needs, have teacher buy-in, be supportable by district staff, originate in schools rather than in district offices, and be backed by credible, high-quality training.

Ross, S.M., Sanders, W., Wright, S., Stringfield, S., Wang, L., Weiping, A., & Albert, M. (2001). Two- and three-year achievement results from the Memphis restructuring initiative. *School Effectiveness and School Improvement*, 12(3), 323-46.

Two-and three-year achievement results from the Memphis restructuring initiative. *School Effectiveness and School Improvement* 12(3), 323-46. This study examined the results of three years of school reform initiatives in Memphis. The authors recount the history of recent Memphis reform initiatives and summarize the various reform models chosen by certain Memphis schools. Using TVAAS scores, the study compares the percent of expected gains attained by reform schools versus non-reform schools. After three years, reform schools generally gained relative to matched non-reform schools. Stronger results were shown in the first group of schools to adopt reform models.

PENNSYLVANIA
Value-Added
Assessment System



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org

BIBLIOGRAPHY

PENNSYLVANIA
Value-Added
Assessment System



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org

- Sanders, W. (1998). Value-added assessment. *School Administrator*, 11(3), 24-27.
"Value-added" methodology is defined as "a statistical method of determining the effectiveness of school systems, schools and teachers in sustaining academic growth for student populations." This methodology allows each student to serve as his or her own control, rather than explicitly accounting for race and socioeconomic status. The validity of this approach is confirmed by the fact that demonstrated cumulative school gains are unrelated to racial and socioeconomic factors.
- Sanders, W. (2000). Value-added assessment from student achievement data: opportunities and hurdles. *Journal of Personnel Evaluation in Education*, 14(4), 329-39.
In this transcript of a speech given on receiving an award named for Dr. Jay Milman, the author describes TVAAS and addresses common criticisms of value-added assessment. He describes the pattern, especially in inner city schools, in which previously lower scoring students make progress, but earlier higher scoring students fail to do so. The goal should be realistic growth for all students. While measurement alone will not bring change, information has to be available to guide development and growth.
- Sanders, W. L., & Horn, S. (1993). *An overview of the Tennessee Value-Added Assessment System (TVAAS) with answers to frequently asked questions*. (booklet) Knoxville, TN: Univ. of Tennessee.
- Sanders, W.L. & Horn, S.P. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, 8(3), 299-311.
TVAAS is described, including the background for its development and the advantages of using a statistical mixed model approach. The authors defend the practice of using each student as his or her own control rather than trying to explicitly account for all other variables that affect student achievement, noting that it is impossible to account for everything that might have an effect on a student's achievement in a given year.
- Sanders, W. L., & Horn, S. P. (1995). Educational assessment reassessed: The usefulness of standardized and alternative measures of student achievement as indicators for the assessment of educational outcomes. *Educational Policy Analysis Archives*, 3(6).
- Sanders, W.L., Saxton, A., & Horn, S.P (1997). The Tennessee Value-Added Assessment System: a quantitative, outcomes-based approach to educational assessment. In J. Millman (Ed.), *Grading teachers, grading schools: is student achievement a valid evaluation measure?* (137-162). Thousand Oaks, CA: Corwin Press.
TVAAS is described by its creator as a "repeated measures, multivariate response analysis" that allows the inclusion of all the information for each student regardless of the degree of missing information. It uses each student as his or her own control, because it would be a "hopeless impossibility" to include all data for each child so that all confounding influences could be accounted for. The history of TVAAS implementation in Tennessee and the efforts made to educate teachers and other groups is summarized.
- Sanders, W. L., Saxton, A. M., Schneider, J. F., Dearden, B. L., Wright, S. Paul, and Horn, S. P. (1994). Effects of building change on indicators of student academic growth. *Evaluation Perspectives*, 4(1).

BIBLIOGRAPHY

Sanders, W. L., Wright, P. S., & Rivers, J. C. (2006). Measurement of academic growth of individual students toward variable and meaningful academic standards. In Lissitz, R. W. (Ed.). *Longitudinal and value added models of student performance*. Maple Grove, MN: Journal of Applied Measurement Press.

Sanders, W.L., Wright, P., Ross, S. & Wang, W. (2000). Value-added achievement results for three cohorts of Roots and Wings schools in Memphis: 1995-1999 Outcomes. www.successforall.com

This is an extension of a prior study of Roots & Wings schools in Memphis. Twenty-two Roots & Wings schools were examined, with four, three and two years of implementation history. Among the cohort with the longest history of implementation, achievement gains measured by the TVAAS (using the percent of national norm gain across all grades) significantly exceeded comparison non-reform schools. The results were positive, but less pronounced for cohorts with fewer years of reform experience.

Schacter, J., Schiff, T. Thum, Y.M., Fagnano, C., Bendotti, M., Solmon, L., Firetag, K. and Milken, L. (n.d.). *The impact of the Teacher Advancement Program on student achievement, teacher attitudes, and job satisfaction*. Milken Family Foundation. Retrieved from www.mff.org/pubs/impact_of_tap.pdf

This study compared gains in student achievement of schools implementing a reform model developed by the Milken Foundation to similar schools not implementing the reform model. Student level scores for Stanford 9 Total Reading, Language and Mathematics were used in a value-added model that measured the percent of the gap to a specified target achievement level that was in fact achieved. All reform schools made improvement and gained significantly more than control schools. Those schools that implemented the reforms the most zealously gained the most.

Stecher, B., & Naftel, S. (2006). *Implementing standards-based accountability (ISBA): study design, state context, and accountability policies*. Rand Education. Paper presented at the symposium "Implementing No Child Left Behind: New Evidence from Three States" at the annual meeting of the American Educational Research Association, San Francisco. CA. April. Retrieved from http://www.rand.org/pubs/working_papers/2006/RAND_WR380.pdf.

Sterbin, A. (2001). Rozelle Elementary School: A longitudinal analysis 1995-2000. University of Memphis. Retrieved from www.mrsh.org/ipr.html

In a case study of one Memphis elementary school, the author used value-added scores, among other factors, to assess the success of Modern Red School House comprehensive school reform model as implemented in the school. Comparing longitudinal value-added scores from the subject school with those of similar schools, all Memphis schools, and all Tennessee schools, he found that value-added scores rose, especially in the early years of reform implementation, far more than in all comparison groups, and remained higher throughout the study than comparable schools and all Memphis schools.

Stewart, B. (2006). *Value-Added Modeling: The Challenge of Measuring Educational Outcomes*: Carnegie Corporation of New York.

The report summarizes the history of value-added models (VAM) for measuring student academic progress over time and the discovery of VAM findings from Sanders to researchers like Robert Mendro. The author discusses the ways in which VAM is being used, and the criticism about the accuracy of VAM measurement.

PENNSYLVANIA
Value-Added
Assessment System



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org

BIBLIOGRAPHY

Stronge, J. & Tucker, P. (2000). *Teacher Evaluation and Student Achievement*. Annapolis Junction, MD: NEA Professional Library.

Thum, Y.M. (2003). *No Child Left Behind: Methodological challenges & recommendations for measuring adequate yearly progress*. CSE Tech Report 590, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles.

A statistical accountability measure that focuses on value-added performance can be used to measure Annual Yearly Progress (AYP) under NCLB. AYP can be conceived as a comparison of a unit's growth rate to the rate expected if the unit is to reach proficiency by the NCLB deadline. While there are various types of measures called value-added, only a model based on student level gain gives us a true map of change. (Estimating gains scores are preferable to regressing post-test scores on pre-test scores.) Even when the unit of accountability is the school, the unit of analysis should be the student. Reports should show, and accountability regimes should take account of, the precision of the productivity estimate.

Thum, Y.M. (2002). *Measuring student and school progress with the California API*. National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles.

The California Public Schools Accountability Act required the state board of education to establish a numerical index (API) for measuring school performance and the performance of groups of students within schools. The author concludes that the API misrepresents student and school performance and presents an alternative based on a Bayesian meta-analysis of results from multilevel models of student test scores. This new approach retains the main features of the API and is feasible given the current API database. The approach is illustrated using data from the Long Beach Unified School District.

Thum, Y.M. & Bryk, A.S. (1997). Value-added productivity indicators. In J. Millman (Ed.), *Grading teachers, grading schools: is student achievement a valid evaluation measure?* (137-162). Thousand Oaks, CA: Corwin Press.

Dallas is a leader in school accountability, especially notable for the degree of political support and educator buy in for the system. While the system rules work reasonably well at controlling for student mobility and preventing manipulation, and the adjustment for fairness variables is appropriate, there are significant questions about the validity of the model and it should be used with caution.

Vaughan, A. (2002). Standards, accountability, and the determination of school success. *The Education Forum*, 66(3), 206-13.

A review of the history of the modern standards movement from 1989 to the present includes examples of state successes in raising student achievement and narrowing the achievement gap. Policy makers now realize that evaluations based on average test scores simply result in rewarding prosperous schools. Value-added evaluation models such as those in place in North Carolina, Tennessee, Prince Georges County, Maryland, Colonial School District in Pennsylvania, and Dallas, Texas may be the best hope for transforming schools.

Viadero, D. (2006). U.S. Pilot of AYP 'Growth' Models Advances; *Education Week*.

To make adequate progress, or AYP, schools and districts currently must meet annual targets for the percent of students who score at least at the proficient level on state tests, both for the student population as a whole and for subgroups of students who are poor, speak limited English, have disabilities, or come from racial or ethnic minorities. Federal officials are sending proposals from eight states on to the next round of screening for a new pilot program allowing states to consider students' academic growth in measuring whether schools and districts meet performance targets under the No Child Left Behind Act.

PENNSYLVANIA
Value-Added
Assessment System



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org

BIBLIOGRAPHY

Webster, W.J. (1998). A comprehensive system for the evaluation of schools. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA., April 13-17, 1998.

The author describes the three-tier accountability system in place in the Dallas public schools. School effectiveness indices are one component of the system. The inclusion of these indices results in a valid and fair way to compare the effects of the school on student learning, apart from factors out of the school's control.

Webster, W.J. & Mendro, R.L. (1995). Evaluation for improved school level decision-making and productivity. *Studies in Education Evaluation*, 21, 361-399.

Dallas' three-tier accountability system couples accountability at the school and district levels with school accountability indices. School and district improvement plans support the first two components of the accountability system. The school improvement indices used by Dallas schools separates school effects from non-school effects in a more precise way than measures used by the Texas Education Agency. The only fair method for holding schools accountable for improvement is to adjust measures of outcomes for factors that impact those outcomes but are out of school control. This can be done through multiple regression and through hierarchical linear modeling. The indices predict student levels of accomplishment and set the desired level of improvement based on these predictions. School-level consequences are imposed based on achievement or non-achievement of predicted levels of student improvement.

PENNSYLVANIA
**Value-Added
Assessment System**



333 Market Street
Harrisburg, PA 17126
(717) 606-1911

www.pde.state.pa.us
pdepvaas@iu13.org