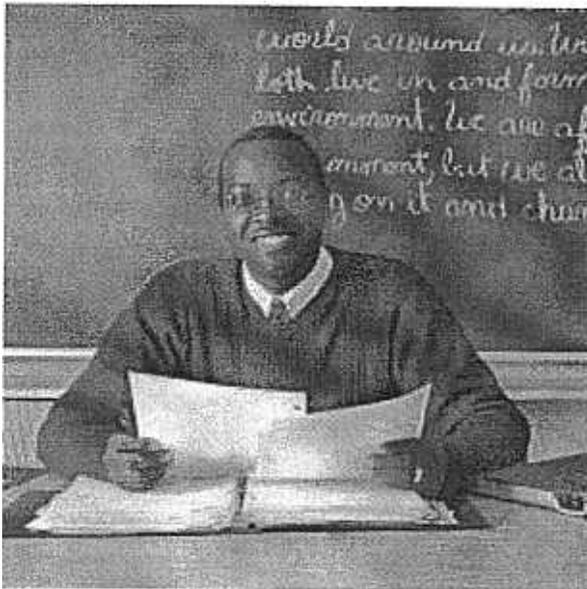


EVALUATING VALUE-ADDED



*Findings and
Recommendations
from the NASBE
Study Group on
Value-Added
Assessments*

**Executive
Summary**

NASBE
NATIONAL ASSOCIATION OF
STATE BOARDS OF EDUCATION

October 2005

The NASBE Study Group on Value-Added Assessments

Members

(all are state board of education members unless otherwise indicated)

Richard Farmer, Jr., Chairman-DE

Isis Castro-VA

Nancy Danhof-MI

Randy DeHoff-CO

Matthew Diethelm-AZ

Edward Donley-PA

Josephine Figueras-NJ

Priscilla M. Haden-WV

Kandy Imes-NE

Michael Kean-CA, *Associate Member, McGraw-Hill Education*

Mary Sue Murray-GA

Mike Nettles-NJ, *Ex officio member, Educational Testing Service*

Sharon M. Osborne-RI, *State Board Executive*

Karabelle Pizzigati-MD

Hilma Prather-KY

Clyde Reese-NJ, *Ex officio member, Educational Testing Service*

Debra Roberts-UT

G. R. "Sam" Schloemer-OH

Conrad (Wes) Snyder-MT, *Associate Member, University of Montana-Missoula*

John A. Tate III-NC

Richard Thompson-NC, *Associate Member, University of North Carolina*

Janet Tockman-ME

Presentations to the Study Group

Value-Added Assessments and Why They Matter to State Policymakers

Thomas Fisher, Assessment and Accountability Consultant

ABC's of Value-Added Models

Henry Braun, Educational Testing Service
Laura Goe, Educational Testing Service

Value-Added/Growth Models as They Relate to State Accountability Systems

Brian Gong, National Center for the Improvement of Educational Assessment

State Actions in Value-Added Assessments

Lois Adams-Rogers, Council of Chief State School Officers
Rolf Blank, Council of Chief State School Officers

Value-Added Assessment in Tennessee

William Sanders, SAS Institute, Inc.

Lessons Learned from Districts Using Value-Added Models

Nancy Doorey, Brandywine (Delaware) Board of Education
Steve Iovino, Warwick school District, Pennsylvania

State Data Systems

Deborah Newby, Council of Chief State School Officers
Richard Kesner, Northeastern University

Value-Added and Formative Assessment: Tension, Synergy

Dylan William, Educational Testing Service
Siobhan Leahy, Educational Testing Service

Teacher Organization Viewpoints

Rob Weil, American Federation of Teachers
Tom Blanford, National Education Association

U.S. Department of Education Viewpoint on Value-Added Model and Its Possible Inclusion in AYP Calculations

Kerri Briggs, U.S. Department of Education

Policy Implications of Value-Added Assessments

Mitchell Chester, Ohio Department of Education

Study Group Staff

David Kysilko

Marsha McMullin

Weaver Rogers

NASBE gratefully acknowledges the Educational Testing Service (ETS) for its support of the Study Group meetings.

© Copyright 2005
National Association of
State Boards of Education
All rights reserved
ISBN 1-58434-063-0

Copies of the full report are available for \$14.00 each + 10% for shipping and handling (\$4.50 minimum) from the National Association of State Boards of Education at 277 South Washington Street, Suite 100, Alexandria, Virginia 22314; (800) 220-5183.

Table of Contents

Executive Summary	4
Preface	0
Overview of Value Added Assessment	
The Uses and Limitations of Value Added Assessment	19
Findings and Recommendations	7
Implementation Issues	29
Findings and Recommendations	34
Checklist for Policymakers Considering Value-Added Assessment	37
Endnotes	39

Executive Summary/ Findings & Recommendations

For decades, education policy-makers have wanted to go beyond traditional ways of analyzing student test score data that only looks at the particular status or "proficiency level" that students have attained. It is very worthwhile, many have said, to also see how far students have progressed over a given year or number of years. This is particularly important when students have started far behind their peers. A student who goes from nonreading to reading has made great progress, even if the final test scores still shows that he or she is reading below grade level. It is also important, in this age of Adequate Yearly Progress (AYP), for schools that are educating large numbers of disadvantaged students to be able to show that their students have made significant strides, even if many have not yet achieved a level of "proficiency." Finally, educators have been intrigued with the possibility of accurately determining the specific impact that a school or teacher has had on students' learning.

To address these issues, many education policymakers, researchers, and practitioners have turned to value-added assessments (or more accurately, "value-added models for analyzing assessment data"). Value-added assessment models are statistical approaches that use multi-year student test score data, and

student background characteristics in some models, to attribute student growth to schools, teachers, or both. In other words, value-added models attempt to determine how far a student has progressed compared to where the student started and to what degree that growth can be attributed to educational factors (as opposed to "external" factors such as socio-economic status, race, parents' educational levels, or innate ability).

Because value-added models are now beginning to be used with increasing frequency in states and districts across the country—and because there are many questions surrounding value-added both in terms of just what the concept means and how the models should be used—NASBE established its Study Group on Value-Added Assessments to address and make recommendations on a number of issues critical to education policymakers, including: How does value-added fit with other aspects of the state's testing and accountability system? How accurate and valid are the results? What are the best uses for value-added data and analysis in terms of school improvement, accountability, or other uses? What doesn't value-added do well? What do states need to think about if they are planning to use value-added analysis as a component of their assessment and accountability system?

The Uses and Limitations of Value-Added Assessment

Through its investigation, the Study Group examined a number of possible uses for value-added models, including:

Using value-added as a component of school accountability: If, as research has borne out to some degree, value-added assessments are able to distinguish the effects of teachers and schools on student achievement, then using value-added data as a component of a state's accountability system would seem to make sense, for example, as one indicator in a school's report card. Researchers caution, however, that there are very significant policy and technical hurdles that must be overcome in order to successfully implement such a system.

Using Value-Added for Teacher Accountability and Evaluation: Since value-added analysis is often portrayed as a way to distinguish the effects of individual teachers and classrooms on student achievement, it is not surprising that many politicians and others are tempted to use them for teacher accountability. As the Study Group heard over and over, however, for a host of statistical and other reasons this area must be approached with considerable caution, especially in terms of high-stakes decisions such as ranking teachers, merit pay, and promotion or dismissal.

This doesn't mean that value-added might not play some role in teacher evaluation, however. The Study Group's conclusion is perhaps best summed up by the most extensive evaluation of value-added models to date, by RAND Education, which concluded that while "the research base is currently insufficient to support the use of value-added models for high-stakes decisions," they do show "promise for lower-stakes diagnostic purposes" such as initially identifying possibly low- or high-performing teachers who can then be further evaluated to confirm results. In this way, value-added functions as a filter for detecting those teachers who would be subject to additional study through classroom observations, diagnostic tests, and portfolios. If further study confirms that a teacher is struggling, he or she can be counseled, provided professional development, or matched with a teacher who has been confirmed as highly effective.

Using Value-Added as a Tool for School Improvement: The Study

Group was enthusiastic about the use of value-added models as a data-driven component of efforts to improve instruction at the classroom, school, and district levels. Indeed, many believe that this is the most significant advantage of value-added models. Some of the key areas that can be informed by value-added analysis include: policy and program evaluation; identification of students in need; schoolwide and team planning; individualized professional development; and resource management.

Using value-added data for improving teacher training:

Given the uneven quality of graduates from the nation's teacher training institutions, states have long struggled with finding mechanisms to hold preparation programs accountable for results, or at least to get a clearer picture of the effectiveness of new teachers and to help shore up deficiencies in their training. While this use of value-added analysis is still in its infancy, two states (Louisiana and Ohio) have initiated projects that

use value-added data to link the effectiveness of new teachers with their preparation programs.

Using value-added or growth measures as a component of AYP calculations: Many educators and policymakers feel that strict, yearly adherence to proficiency targets as used under the No Child Left Behind Act unfairly punishes those schools dealing with large numbers of disadvantaged students, because they are required to catch up, often very quickly, to schools that start with higher achieving students. Many states are now looking into ways to add growth or value-added measures to AYP calculations, and the Study Group strongly urges the U.S. Department of Education to provide this flexibility while still maintaining the framework of proficiency for all students that is the basis of the law.

In considering these uses for value-added assessment models, the Study Group made the recommendation below:

Findings and Recommendations on the Uses and Limitations of Value-Added Models

Recommendation 1. Value-added assessment is not designed for high-stakes use in teacher evaluations. The Study Group recommends that value-added information not be used for high-stakes teacher evaluation involving either rewards or punishments. We believe that educators should recognize that value-added assessment is a "tool," but it is not the "total"—and indeed that the data can only with certainty identify about the top 10 percent and bottom 10 percent of teachers. Researchers also note that the number of studies supporting the value-added methodology is relatively limited and that there is a serious lack of validity studies. Thus, users are strongly encouraged to exercise caution in any use of value-added assessments so that the proper use of the technique does not go beyond the capabilities of the tool.

Recommendation 2. Value-added assessment has significant potential—when used in conjunction with other measures and supports—as a tool to improve teaching. The potential of value-added data to 1) differentiate the very most and least effective teachers and 2) show individual teachers more precisely in which areas and with which students they are and aren't being successful can be a valuable tool in helping all teachers improve.

Recommendation 3. Value-added assessment information can be a powerful school improvement tool in promoting more effective practices and resource allocation. If a state or district uses value-added information, the Study Group recommends that the information be used for diagnostic purposes geared toward strengthening school performance and teaching techniques. Indeed, perhaps the most potent potential of the value-added model is through its power to start the dialogue about improving teacher effectiveness. At the same time, principals and school teams can use value-added data to identify effective or ineffective programs and curricula, identify individual students or groups of students who need additional support to reach their potential, and target school resources to areas in need.

Recommendation 4. Once-a-year value-added analysis does not substitute for other ongoing testing information that helps teachers adjust and target instruction to meet the needs of their students. While the Study Group believes that value-added assessment models have significant potential, the Group also recommends that schools avoid reliance on information from just a yearly testing program. Teacher-designed tests, commercial diagnostic measures, and other interim assessments done throughout the school year are crucial to keeping schools on target and can allow for mid-course corrections as needs arise. Waiting for results from one statewide testing program is inadequate for improving classroom instruction.

Recommendation 5. Value-added models show potential for tracking and improving the effectiveness of teacher training institutions. Using value-added data with newly licensed teachers has the potential to provide policymakers with information on which preparation programs are producing the most and least effective young teachers. Preparation programs can also use this information, especially if used alongside interviews with the teachers, to pinpoint weaknesses in their curriculum and other aspects of their program. However, as with other areas, the Study Group does not believe that value-added assessment is appropriate at this time as a stand-alone, high-stakes measure for evaluating teacher-training institutions.

Recommendation 6. The Study Group urges the U.S. Department of Education to allow growth indicators as a component of AYP calculations within the No Child Left Behind framework. While the rigorous targets for attainment required by NCLB serve an important purpose in keeping a focus on the ultimate goal of high achievement for all students, adding growth is needed not only out of fairness, but because together these indicators provide the most accurate picture of the effectiveness of schools. In addition, failure to use growth as one indicator of success could end up making it even more difficult to retain effective teachers in disadvantaged schools.

Implementation Issues

Because value-added assessments are so new to the education enterprise, because they tend to be extremely complex in their mechanics, and because there are potentially a number

of high-stakes implications for their use, state boards of education and other decisionmakers will find there are a host of issues that must be considered before, during, and after implementation of value-added systems. While some of these are technical issues that

must be "checked off" as the state proceeds, the Study Group found that other technical, system-related, and political issues have the potential to be "deal breakers" or to render the value-added models far less effective than they otherwise might be.

Key Testing and Data Issues Affecting Implementation

Value-added models use results from state or district testing programs, along with lots of student and other data, for the basis of their analysis. Therefore, value-added analyses are believable only when the underlying testing program has technical quality, produces valid scores, and maintains an accurate longitudinal database. Holding to these standards is likely to be a challenge for many states. The Study Group identified the following as some of the key areas of concern:

- **Testing Issues:** Policymakers must first ensure that state tests have a high degree of validity and reliability. For value-added, it is also important that tests have "stretch," that is, that they measure the full range of the material that is being covered. There must also be multiple test scores over several years for individual students. Finally, states must ensure that random measurement errors are kept to a minimum, especially if the value-added results are to be used as part of an accountability system.

- **Individual Student Identifier:** Value-added models require that each student has a unique student identifier (ID). About half of all students currently have or are working on establishing individual student IDs.

- **Statewide Data Collection Systems:** For value-added methodologies to work well, states need a robust data collection and management system that functions seamlessly from the school building to the state level. Such systems must be able to link program, course, and student data and enable users to efficiently exchange data electronically. As part of this

system, states need reasonable processes to identify data that are in error, to spot check certain information randomly, and to conduct site visits to audit the accuracy of data at the local level.

State Education System Considerations

Value-added assessment models will make a number of demands on state education agencies that policymakers should attend to for the program to be effective. These include:

- **Staff Training:** Many of the experts consulted by the Study Group emphasized that value-added approaches are ill advised without a simultaneous commitment to a significant amount of training. Teacher and principal training on value-added assessment will need to fully inform staff about the basic concepts behind value-added assessment and what the data mean. Training should also provide principals and teachers with opportunities to practice analyzing real value-added assessment data, as well as putting the analysis to work in improving instruction.

- **Measurement Expertise:** States and districts will need to have a significant measurement and statistical capacity either in-house or contracted on a consulting basis. States should also consider using expert advisory panels both for evaluating the initial development plans for the value-added system and for periodically reviewing the quality and effectiveness of the system once it is up and running.

- **Costs:** While some value-added developers note that the "big cost" of testing is in the purchase, administration, and scoring of tests, which states

already do, purchasing or developing a value-added system on top of the existing testing program can add significantly to an education budget. This is especially true when other needs such as developing a robust data management system, massive school personnel training, and ensuring statistical and psychometric expertise are included.

Public Engagement and Political Considerations

Many Study Group presenters emphasized the need to maintain trust and involvement with affected stakeholders—particularly teachers—throughout any effort to implement value-added analysis. It is clear from many educational innovations over many years that any number of players can cause a program to fail due to lack of understanding, ignorance of the process, or mistrust. Value-added assessment, rightly or wrongly, has developed the reputation as a high-stakes policy instrument. Thus, extra care must be taken to develop trust and buy-in. Following are some of the particular political and public relations areas that must be considered.

- **The Overall Impact of Changes to the Testing or Accountability System:** Adding yet another significant change to a testing or accountability system can be demoralizing to teachers and confusing to the public. Because value-added systems may require significant change and enhancements to existing accountability systems, states should carefully consider the purposes of the existing system and think through the adjustments to current efforts that will be needed to incorporate a value-added approach. In addition, potential users should recognize that implementing a value-added system cannot be done quickly.

- **State board of education planning and communications:** State boards will need to act as a voice of reason to help the public, school personnel, and lawmakers understand both the benefits and challenges of value-added models. Perhaps most importantly, state boards will need to exercise leadership and insist that decisions about value-added assessment should be made using reason and good judgment and not in response to persuasive sales approaches by vendors who are

marketing commercial value-added testing systems.

- **Relations with Teachers:** Teachers, in particular, may be cynical about seeing yet another education reform heading their way. They may be suspicious that the "real" reason behind value-added analysis is punishing or rewarding teachers, and they may be wary of the amount of time it will take to understand and use value-added results. On the other hand, teachers may see the growth orientation of value-added assessment

as an improvement over current accountability programs and they may be very interested in the potential of the system to improve instruction for classes and individuals. All in all, the Study Group believes that teachers (and principals) will respond much more positively to value-added models if the system is not seen as a stick, but primarily as a way to assist educators in improving instruction and promoting more growth in student learning—in short, that it can be an important tool in helping teachers do their jobs.

Findings and Recommendations on Implementation Issues for Value-Added Assessments

Recommendation 7. States considering a value-added system should develop and complete an "Implementation Checklist" as part of the initial planning process. Due to the complexity of value-added models, the demands they place on state testing and data systems, and the effects they can have on the overall education system, it is imperative that policymakers conduct a complete and honest evaluation of their current structures to ensure they meet the requirements for a value-added assessments. (See the Study Group's checklist on page 37 as a starting point.)

Recommendation 8. Value-added assessment requires massive training for state, district, and school personnel. The Study Group believes that any state or local district that wishes to use value-added assessment needs to prepare for massive training of teachers and others in how to use the information generated. Teacher training on value-added assessment will need to inform teachers about the basic theory and methodology of value-added assessment. Training should also provide teachers with opportunities to practice analyzing real value-added assessment data, and develop skills regarding how to use the results with faculty and parents.

Recommendation 9. Users of value-added assessment must consider confidentiality and transparency issues. The Study Group recommends that states and districts considering value-added assessment develop appropriate policies to keep teacher ratings confidential, and limit their availability to the teachers themselves and authorized local school officials. At the same time, confidentiality must be balanced with useful reporting to educators and the public in order to attain sufficient transparency. Legal assistance may be needed to determine wording for policies in this area.

Recommendation 10. Teacher education programs should incorporate more training in educational measurement and the use of data into their curricula. The Study Group believes it is crucial that training regarding the basic concepts of status and growth measures, value-added assessment models, and the development and use of formative assessments be provided to prospective teachers in higher education. Those preparing to teach should also have opportunities to practice using data for decisions prior to classroom employment.

Recommendation 11. State policymakers and a wide range of stakeholders, including teachers, should be involved early in planning for value-added assessment. Legislative leaders and legislators serving on education committees are two groups that need to understand the general rationale of value-added assessment and understand what the model can and cannot do. In addition, educators need to know what concerns and complaints might be anticipated from parents or teachers so that they will be prepared to respond if they are contacted by constituents.

Recommendation 12. Educators should exercise due diligence in evaluating commercial companies offering value-added assessment products. The Study Group recommends that all policymakers who are considering using a value-added system be fully aware that they are facing an entrepreneurial environment. It may be true that many, if not most, of these vendors are trustworthy, but they are still vendors interested in selling something for a profit. It is also true that each vendor has limitations and expertise that need to be evaluated for appropriateness to the educational environment and the timeline of a given state or local school district. Consumers of value-added information and services need to do thorough due diligence before signing agreements to work with firms or individuals.

Recommendation 13. Financial costs must be considered carefully when considering value-added assessment. The Study Group recommends that state boards and other potential value-added assessment consumers be alert to the potential costs of this approach. While some experts note that the "big cost" of assessment is in the purchase, administration, and scoring of tests (which states already do), including a value-added system on top of a testing program can add significantly to an education budget. Users should do a cost/benefit analysis to determine whether the information they will receive justifies the price.

Recommendation 14. Value-added assessment needs continued pilot testing, research, and validation work. The Study Group is convinced that value-added assessment is a highly promising—although still immature—approach. While preliminary results show significant potential to improve education, continued pilot testing and research is needed. The Study Group believes that only with appropriate validation studies and other research will educators and policymakers get maximum benefit from the methodology.

Concluding Thoughts

In a sense, educators currently face a "measurement emergency." They face increasing pressure from NCLB and other state or local accountability programs to demonstrate increased student learning. This pressure encourages the use of testing information to judge the performance—the value being added—of teachers and schools. With educators, policymakers, and the public becoming more measurement and out-

come-oriented, it is natural that public school personnel and policymakers would seek additional ways to judge progress and use testing data.

The Study Group believes that value-added models—with their promise of attributing student growth to schools, teachers, or both—have the potential of offering a way to analyze student achievement in a more individual, robust, and understandable way than previous methods. As it is refined as a model, and if it can be adequately

validated, it can lead us toward new information about the teacher skills, curriculum components, or program initiatives that are particularly effective in improving student learning. But, perhaps the most important feature of value-added assessment is that it serves to keep everyone's focus on student growth and learning momentum, which is the essence of the schooling experience. After all, is it not the main role of educators to take children from where they find them and then "add value?"