

# STUDENT ASSESSMENTS



“Anyone who opposes annual testing is an apologist for a broken system of education that dismisses certain children and classes of children as unteachable,” said Rodney Paige, secretary of education. That may be, but what kind of test should we use to evaluate how well children are doing academically over time? The firestorm over assessments gets hotter as more fuel is added by frustrated teachers and scared parents.

While efforts to assess student learning have always been plagued with difficulties, comparing student learning between classrooms, schools, or districts has proven nearly impossible. The norm-referenced tests that have been the common standardized measure for decades measures students in relation to their peers. By definition, half of the students will come up as “below average” on these tests, making raising student test scores on any large scale a futile endeavor.

This is not to say that norm-referenced tests are valueless—just that they have focused and limited value. Further, if the only performance standard is the “average,” there is no objective way of knowing whether that average reflects what and how much students should actually be learning.

In response to this, high stakes “criterion-referenced” assessments were created, such as the Washington Assessment of Student Learning (WASL). However, this raises another conundrum. Assuming that a valid and reliable high stakes assessment can be created—and this is not a safe assumption—if standards are set high, the vast majority of students may fail them, thus creating tremendous political pressure to either scrap the standards altogether or lower them to currently attainable levels.

Into this conflict comes another option for comparing achievement levels: value-added assessment. In-

stead of comparing a student with others as norm-referenced tests do, or solely against an established standard, value-added assessment compares a student with herself. The measurement is how far the student has progressed in one year compared to where she started.

Furthermore, by statistical analysis of data from groups of students, conclusions can be drawn about the impact of a particular teacher, school, or district on student learning.

***Instead of comparing students to each other or to an established level of proficiency, value-added assessment compares students to themselves, to determine if they are advancing academically.***

## What is value added assessment?

Value-added assessment is *not* a new or different type of test. Rather, it is a model used to statistically analyze test data to determine the influence of teachers, schools, and school districts on student learning. Instead of comparing students to each other or to an established level of proficiency, value-added assessment compares students to themselves, to determine if they are advancing academically, and, if so, at what pace.

The value-added assessment statistical model uses test scores accumulated year to year from each student

to track change in achievement. This allows creation of academic “growth charts” for each student’s progress, measuring the “value added” to the knowledge the student already had. Like a physical growth chart, the curve is rarely even—the record will show flat spots or spurts of accelerated learning. By calculating statistically significant variances in a group of students’ test scores, determination can be made as to how well a particular teacher, school, or district is educating a particular student.

The most prominent value-added assessment model was developed by Dr. William Sanders, a former statistics professor at the University of Tennessee. His method, called the Tennessee Value-Added Assessment System (TVAAS), uses mixed-model methodology, a type of statistical analysis developed originally for use in agriculture.

Essentially, value-added assessment is like a pre-and-post-test given to students to determine what they have learned during a particular course of study. Unlike those tests, however, value-added assessment seeks results that can be compared across classrooms and years. Thus, the value-added model must be overlaid on tests that have the following characteristics:<sup>1</sup>

- The tests must be highly correlated with curricular objectives
- The tests must have sufficient stretch to measure the progress of both previously low and high scoring students
- The tests must demonstrate appropriate reliability

Many schools use value-added assessment with the readily available standardized achievement test results. These are norm-referenced tests, comparing one student with the average performance of all students. But a sufficiently reliable and consistent criterion-referenced test could be used instead, comparing students with an established standard of achievement. Naturally, the better the test at measuring student achievement, and the closer it correlates to established learning goals, the more valuable the results of the analysis will be.

Because students’ scores are compared with their own prior test scores, external factors such as socioeconomic status are blocked out. Gathering the data over several years accommodates for statistical variations, while the TVAAS model is constructed in a way to allow for variables such as missed tests, transferred students, skipped grades, and other complications.<sup>2</sup>

## Advantages to value-added assessment

Because value-added assessment focuses on students' rate of advancement, it provides a means of objective feedback on how well a particular teacher, school, or district is doing. This is an unprecedented development in assessment. The resulting feedback allows identification of high achieving teachers, schools, or districts to more closely analyze the reasons for success. It identifies underachieving teachers, schools and districts as well, making it possible to hold them accountable for ensuring student progress, without faulting them for circumstances beyond their control.

Value-added assessment opens up the possibility of tying teacher and school funding to performance. It makes it possible for high-performing schools and teachers to be identified and emulated and for low-performing schools and teachers to be targeted for help. Professor Sanders has also suggested that this information may be used to protect students from being assigned to ineffective teachers two years in a row.<sup>3</sup>

The TVAAS system is particularly valuable because it allows for use of incomplete data. Thus, even though a student may have missed the test one year, the rest of his data can be compiled, something not possible with simpler statistical models. This allows the maximum possible data to be used in the analysis.

There are many other possibilities for this data. Dr. John E. Stone, professor at the College of Education at East Tennessee State University, has proposed that value-added assessment and state standards could be combined, calculating for schools how quickly students would need to progress to meet established standards.<sup>4</sup>

## Challenges to value-added assessment

Like any education measure, value-added assessment has had its fair share of criticism. Many of these criticisms either reflect misunderstanding of the process or a fundamental bias against any form of standardized testing.

Some claim that, since the model ignores socioeconomic differences between students, it does not fairly evaluate different results obtained by different teachers or schools. This criticism ignores reliable research demonstrating that socioeconomic differences *by themselves*

do not significantly limit students' ability to progress, although they may influence achievement levels.

Critics also suggest that value-added assessment will result in "teaching to the test"—the same criticism lodged against every proposed results-oriented evaluation. "Teaching to the test" is only a problem if the test does not measure the academic standards already established, and to which teachers should be teaching. The remedy, if this is a problem, is not to throw out evaluation, but to better align the test with learning objectives.

***"Teaching to the test" is only a problem if the test does not measure the academic standards already established.***

Familiar concerns are raised regarding the inability of standardized tests to measure all aspects of student learning. That value-added assessment may exacerbate this problem troubles critics. It is true that many kinds of achievement cannot be measured objectively, but this should not invalidate measuring that which can be objectively determined, nor should it diminish the worthiness of using value-added assessment as a tool to better "read" the progress of particular students.

Another concern is that value-added assessment may lead to a focus solely on whether students are progressing, without ensuring those students who start out further behind ever achieve necessary levels of proficiency. Again, this must be addressed by using other methods alongside value-added assessment. Standards can provide minimums that students must achieve while value-added assessment can make sure students at all levels are being challenged. As noted above, the combination of using value-added assessment against predetermined academic standards can provide useful information to schools regarding their ability to meet expectations for student learning.

## Findings

The key finding from value-added assessment may well be the objective confirmation that the single most important factor in how fast a child learns is not his or her past achievement, class size, socioeconomic status, or race, but the quality of the teacher in the classroom.<sup>5</sup> Although a student's life situation may statistically predict an academic starting place, it rarely limits the aca-

ademic ending place. The influence of a teacher dwarfs other factors in the learning equation.

A teacher's influence, for good or ill, has significant residual impact—results that can be measured through value-added assessment. Students who had top teachers for three years in a row scored as many as 50 percentile points higher in math than students who had poor teachers three years in a row.<sup>6</sup>

Value-added data has also been used to analyze how well students at different achievement levels are learning. The rates of academic gain of students at different ability levels can be compared to see whether any group is learning more slowly than others. The most common pattern in this case is that rates of academic gain decline with higher achieving students, indicating a lack of challenging coursework.

In part, the lower gains among higher achieving students may be due to a teacher's natural tendency to help students who need it most. However, the best teachers—those whose students showed the highest gains—also proved the most capable of reaching all levels of students.<sup>7</sup>

***Many inner-city schools with low aggregate test scores were actually doing an outstanding job of improving student learning.***

## Uses made

The most widespread and prominent use of value-added assessment has been in the state of Tennessee, where the Sanders model of analysis was mandated by the Education Improvement Act of 1992. Reports from TVAAS have been issued since 1994. School district and school reports are

made available to the general public. Teacher reports are made available only to teachers and their administrator. Initially reports were only collected for grades 2-8 using value-added assessment with norm-referenced data. As appropriate tests are developed for the high school level, the scope of the analysis is being expanded.

Tennessee school districts report that the data obtained through value-added assessment has given them objective verification of what some educators instinctively knew: *Many inner-city schools with low aggregate test scores were actually doing an outstanding job of improving student learning.* Teachers whose students did

best were described as hard workers, able to keep themselves and students on task; individuals who arrived early and stayed late; who demonstrated a good sense of humor—a good description of an excellent teacher.<sup>8</sup> This comparison of objective analysis gained through value-added assessment with more subjective, but valuable human experience and intuition provides a reliable context for expectations and accountability.

Not all Tennessee teachers have been enthusiastic about the system. Some ignore the data, while others have found it useful for improving instruction. Marsha Denton, a middle school social studies teacher, discovered that her seventh grade students were learning well, but her eighth grade students weren't showing as much progress. The feedback provided by value-added assessment allowed her to evaluate her teaching methods and modify her teaching style for the older students. Her eighth-graders' scores later went up.<sup>9</sup>

Colorado's Pueblo School District 60 has made use of value-added assessment for five years. They have used a norm-referenced test, but are in the process of incorporating data from the Colorado criterion-referenced assessment. According to John Brainard, Director of Assessment, the district has appreciated the ability to identify and follow best practices. He notes that the data pave the way for better communication between schools about student achievement. Educating principals and parents to understand what the data means has been crucial.<sup>10</sup>

While Tennessee is the only state to date to make broad use of value-added assessment, the statistical model has been used by more than 80 school districts around the nation. The Seattle School District has recently begun using the data received from value-added assessment to analyze school performance. Other districts, such as the Dallas Independent School District, have tried to achieve a similar effect using a simpler statistical analysis.

## Conclusion

Value-added assessment not only supports the operational understanding that all students can learn, it also provides a means to determine the pace at which students are learning.

While no one system of assessment will completely explain and evaluate all aspects of student learning, data obtained from value-added assessment can be a useful tool for teachers and administrators to improve student

learning. Education leaders in Washington state owe this assessment model a much closer look.

## Recommendations

- *School districts should appoint a committee to review value-added assessment.* The review should include evaluation of the model from those who use it and from those who have determined it to be invalid or unworkable. Final recommendations should be based on objective findings.
- *The legislature should commission a preliminary study on value added assessment.* The A+ Commission or State Board of Education could be charged with reviewing the data, analyzing whether or not it can be used effectively in our state, and determining what adjustments would have to be made in our current assessment system to use it. A final report should be given to the 2003 legislature.

## Endnotes

1. "Frequently Asked Questions," *SAS in Schools*, <<http://www.sasinschool.com/evaas/resources/faq/index.shtml>>.
2. William L. Sanders, "Value-Added Assessment," *The School Administrator Web Edition*, December 1998.
3. William L. Sanders, and June C. Rivers, *Cumulative and Residual Effects of Teachers on Future Academic Achievement*, (University of Tennessee Value-Added Research and Assessment Center, 1996).
4. John E. Stone, "What is Value-Added Assessment and Why Do We Need It?" *Policy Brief* (The Foundation Endowment, 1999), 8.
5. S. Paul Wright, Sandra P. Horn, and William L. Sanders, "Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation," *Journal of Personnel Evaluation in Education*, 11:63 (1997)
6. Sanders, "Value-Added Assessment."
7. Sanders, "Value-Added Assessment."
8. Samuel E. Bratton, Jr., "How We're Using Value-Added Assessment," *The School Administrator Web Edition*, December 1998.
9. David Hill, "He's Got Your Number," *Education Week*, May 2000.
10. John Brainard, Director of Assessment, telephone conversation with Karen Helland, EFF Research Analyst, 19 October 2001.