

**LEGISLATIVE EDUCATION STUDY COMMITTEE
BILL ANALYSIS**

Bill Number: SB 497

52nd Legislature, 1st Session, 2015

Tracking Number: .199707.2

Short Title: Quantifiable Data in Teacher Evaluations

Sponsor(s): Senator John M. Sapien

Analyst: Heidi L. Macdonald

Date: March 3, 2015

Bill Summary:

SB 497 amends the *School Personnel Act* to require that quantifiable data account for at least 30 percent of a teacher's overall evaluation.

Fiscal Impact:

SB 497 does not contain an appropriation.

Substantive Issues:

Testimony presented to the Legislative Education Study Committee (LESC) suggests that an examination of the issues addressed by SB 497 may be in order. To illustrate, during the 2014 interim the LESC heard testimony on the evaluation of teachers from the Coalition for Excellence in Science and Math Education (CESE).¹

Addressing the state's teacher and principal evaluation system, the CESE testimony identified several issues with the requirement that 50 percent of a teacher's evaluation be based on his or her students' performance growth, among them:

- student growth can be volatile with no distinct trend;
- an average teacher's ranking is determined primarily by the previous two teachers' performance;
- standardized test score data indicate certain trends through grade levels that are likely not attributable to teacher performance, for instance a consistent upward trend between sixth, seventh, and eighth grade; and
- regarding the use of the value-added model (VAM), the American Statistical Association (ASA) suggests that, aside from teachers accounting for only between 1.0 percent and 14 percent of the variability in test scores, using VAM scores to rank teachers can have unintended consequences that reduce quality.

¹ The CESE describes itself as a nonprofit, non-partisan 501(c)(3) charitable corporation, whose members include national laboratory personnel and retirees, industrial scientists, educators, parents, college professors, and others. CESE has analyzed New Mexico public education data and policy issues for more than 15 years, with a primary focus on helping improve New Mexico schools using data unique to the state.

The CESE testimony concluded with the statement that the best way to evaluate any professional is through rigorous observation, not, in the case of teachers, through reliance on a VAM measure of student performance growth.

Additionally, during the 2014 interim, the LESC heard testimony from school districts and charter schools (see “Background,” below) on the difficulties the VAM model presented during the first year of the teacher and principal evaluation systems implementation.

The Public Education Department (PED) analysis contends that the current effectiveness evaluation system contains data that is 100 percent quantifiable.

Technical Issues:

The term “quantifiable data” is not defined within the amended section. The sponsor may wish to consider defining the term.

Background:

Teacher and School Leader Effectiveness Evaluations

Adopted in August 2012 and amended in September 2013, the PED rule, *Teacher and School Leader Effectiveness*, implements an evaluation program for public school teachers and administrators called the Effectiveness Evaluation System (EES), sometimes also called the NMTEACH Effectiveness Evaluation System. Under this system, districts have the option of using the plan developed by PED or submitting a custom plan to PED for department approval.

In general, 50 percent of a teacher’s evaluation is based on student achievement measures, whether derived from the state standards-based assessments or some other student assessment. Details vary, however, depending upon whether a teacher is a member of Group A, Group B, or Group C:

- Group A teachers teach subjects tested by the standards-based assessments in those grades in which the assessments are administered;
- Group B teachers teach either non-tested subjects or tested subjects in grades in which the standards-based assessments are not administered; and
- Group C teachers teach in grades K-2.

Briefly, the rule requires that:

- school districts use a department-adopted student achievement growth measure or, with department permission, use a combination of PED-approved growth measures and, for non-tested subjects or grades, a PED-approved alternative measure;
- whenever possible, the performance rating include three years or more of student achievement growth data; and
- if a school district has not implemented appropriate course assessments or adopted a comparable measure, student achievement growth be measured by:
 - the growth achievement of the classroom teacher’s students on state assessments;
 - the school’s A through F letter grade for courses in which enrolled students do not take the state assessment, provided that a school district may assign instructional team

- student achievement growth to classroom teachers in lieu of using the school grade growth calculation; or
- state-developed end-of-course examinations or other PED-recommended options.

Upon request by the school district, the rule allows the rating for teachers who are assigned to courses not associated with state assessments to include achievement growth that is demonstrated on state assessments as a percentage of the overall evaluation. In addition, student achievement growth is measured through VAM, which, according to PED, accounts for the individual student's background by using three years' worth of data.² Those years of data produce a teacher's overall value-added score (VAS).

For the remainder of a teacher's evaluation:

- 25 percent is based on teaching observations by one of two types of observers – either “approved” or “certified” – using the NMTEACH rubric or protocol; and
- 25 percent is based on “multiple measures,” which vary, again, according to the group to which the teacher belongs.

The Summative Report form, which summarizes a teacher's progress through the EES to the end of the year, provides basic information about the final score a teacher received in each evaluation category – student achievement, observation, and multiple measures – the individual teacher's overall score, and the median score for comparable group and grade-level teachers. Two sections are completed by the evaluator: one to identify strengths and areas needing improvement and another for identifying next steps. The form also allows the application of “graduated considerations,” which are a means of adjusting the weight of student achievement for those teachers with fewer than three years of student achievement scores.³

2014 Interim

Altogether over the course of five interim meetings, the LESC heard testimony from 27 school districts, two charter schools, and two special state-supported schools on the implementation of the teacher and principal evaluation systems. Testimony from school districts and charter schools noted a lack of clarity with regard to a number of aspects of the evaluation system, among them:

- understanding and then explaining the concepts of the VAM and the VAS;
- the application of graduated considerations;
- the reassignment of student achievement data points to teachers' students from the previous year; and
- districts' lack of access to specific calculations and procedures used to populate data in the summative evaluation reports, limiting their ability to explain and substantiate ratings.

² VAM uses statistical models to predict student test performance, controlling for potential variables that could affect performance such as student, teacher, or school characteristics. The difference between the predicted and actual scores, if any, is assumed to be due to the performance of the teacher, rather than to the student's natural ability or socioeconomic circumstances.

³ PED explains that, if a teacher has three years of VAS, then improved student achievement counts for the full 50 percent of the evaluation score; if the teacher has fewer than three years of VAS, then improved student achievement counts to a lesser degree and observations and multiple measures count for more. Depending upon the number of student assessments used and the number of years of a teacher's VAS, any one of nearly 40 calculations may be applied to determine the points for improved student achievement.

During the November meeting, the Secretary of Public Education provided the PED response to the issues and concerns that districts and charter schools had raised. This testimony began with a description of the impact upon students of effective and ineffective teachers. It also explained how the NMTEACH protocol was designed ultimately to improve student outcomes through certain initiatives targeted at teachers, and it illustrated the differences in teacher ratings under the current evaluation system versus the previous evaluation system.

The Secretary then enumerated several areas for improvement in the implementation of the evaluation system that will rely on PED/district partnerships, among them:

- “incomplete” or “inaccurate” data;
- understanding the VAS; and
- increased weighting of NMTEACH observations.

The Secretary proposed a number of solutions to these issues, among them:

- establishing a NMTEACH liaison for each district and charter school, as well as providing ongoing training on roster verification, VAS, and the summative reports; and
- partnering with Las Cruces Public Schools and Hobbs Municipal Schools on training modules to help school personnel better understand VAS and VAM.

Finally, in response to a number of questions from committee members, the Secretary made the following additional points, among others:

- the components of principal evaluations are based on school growth measures, how well the principal implemented observations in a timely fashion, and the highly objective uniform statewide standard of evaluation (HOUSSE) competencies; and
- if a teacher has a discrepancy in the observation aspect of the evaluation or in VAM data, there will be a second review of that particular summative report.

Limitations Using the Value-Added Model

An April 8, 2014 statement released by the ASA highlights the limitations of using VAMs in the context of assessing educator performance. According to ASA, VAM is used to estimate effects of individual teachers or schools on student achievement while accounting for differences in student background. VAMs are increasingly promoted or mandated as a component in high-stakes decisions such as evaluating and ranking teachers.

Additionally, ASA makes the following recommendations regarding the use of VAMs:

- the ASA endorses wise use of data, statistical models, and designed experiments for improving the quality of education;
- VAMs are complex statistical models, and high-level statistical expertise is needed to develop the models and interpret their results;
- estimates from VAMs should always be accompanied by measures of precision and a discussion of the assumptions and possible limitations of the model, and these limitations are particularly relevant if VAMs are used for high-stakes purposes;
- VAMs are generally based on standardized test scores, and do not directly measure potential teacher contributions toward other student outcomes;

- VAMs typically measure correlation, not causation: positive or negative effects attributed to a teacher may actually be caused by other factors that are not captured in the model;
- under some conditions, VAM scores and rankings can change substantially when a different model or test is used, and a thorough analysis should be undertaken to evaluate the sensitivity of estimates to different models; and
- VAMs should be viewed within the context of quality improvement, which distinguishes aspects of quality that can be attributed to the system from those that can be attributed to individual teachers, teacher preparation programs, or schools.

Committee Referrals:

SEC/SPAC

Related Bills:

SB 91 *Teacher Licensure Levels & Advancement*

SB 138 *Repeal A-B-C-D-F School Rating Act*

SB 202a *Public Education Data Advisory Council*

SB 205 *Delay Use of Certain Test in Teacher Evals*

SB 378 *Teacher & Admin Differential Performance*

SB 558 *Use of Leave & Teacher Evaluations*

SB 562 *Teacher Evaluation Use of Data*

FL/HB 76a *Teacher Licensure Levels & Advancement*

CS/CS/HB 144 *Teacher & School Leader Effectiveness Act*

HB 156a *Innovations in Teaching Act*