

Fiscal impact reports (FIRs) are prepared by the Legislative Finance Committee (LFC) for standing finance committees of the NM Legislature. The LFC does not assume responsibility for the accuracy of these reports if they are used for other purposes.

Current and previously issued FIRs are available on the NM Legislative Website (www.nmlegis.gov) and may also be obtained from the LFC in Suite 101 of the State Capitol Building North.

FISCAL IMPACT REPORT

SPONSOR Sapient ORIGINAL DATE 2/20/17
 LAST UPDATED _____ HB _____

SHORT TITLE School Assessments & Teacher Evaluations SB 470

ANALYST Liu/Serna Marmol

ESTIMATED ADDITIONAL OPERATING BUDGET IMPACT (dollars in thousands)

	FY17	FY18	FY19	3 Year Total Cost	Recurring or Nonrecurring	Fund Affected
Total		See Fiscal Implications				

(Parenthesis () Indicate Expenditure Decreases)

Conflicts with HB125, HB241, HB350, SB34, SB40
 Relates to HB124, HB158, HB163, HB248

SOURCES OF INFORMATION

LFC Files
 Legislative Education Study Committee (LESC) Files

No Responses Received From
 Public Education Department (PED)

SUMMARY

Synopsis of Bill

Senate Bill 470 amends the Public School Code, substituting PED-required end-of-course examinations with short-cycle assessments for all grades and establishing teacher evaluation system criteria that do not:

- use student test scores as a criterion, but reward teacher flexibility and creativity in meeting the needs of all students;
- account for teachers' use of personal or sick leave, but support professional growth using analysis and application of relevant data; or
- use numbers alone to evaluate teachers, but use observations with appropriate and timely feedback.

FISCAL IMPLICATIONS

The bill does not make an appropriation. Costs of developing short-cycle assessments and adhering to a new teacher evaluation system framework would be borne by school districts and

PED. It is likely most educators have short-cycle assessments based on a course curriculum aligned with state-adopted standards, so additional costs to develop new assessments is anticipated to be minimal for school districts if no standardized framework is sought.

SIGNIFICANT ISSUES

The bill defines “short-cycle assessment” as an evaluation method based on a course curriculum aligned with state-adopted standards. Short-cycle assessments given at the end of the year can provide a teacher with information about the extent of student mastery on the material. Provisions of the bill substitute any PED-required end-of-course examinations with short-cycle assessments for all grades. The bill does not specify the frequency of short-cycle assessments or the entity responsible for developing or approving the administration of these assessments. See Technical Issues for more information.

Provisions of 6.19.7.8 NMAC require short-cycle diagnostic assessments to be administered each year for ninth and tenth grades students in the fall and at least two other times during the school year for reading, language arts, and math. Short-cycle diagnostic assessments are defined as “a formative measure that is regularly used to assess student performance over a short time period.” School districts, charter schools, or any state agency wishing to develop or adopt a short-cycle diagnostic type assessment as an alternative to the PED-approved statewide short-cycle diagnostic type assessment must receive PED approval before administration. Results of these short-cycle diagnostic assessments must be reported in written form to students, parents, and school administrators no later than four weeks from the date of administration of the assessments.

Provisions of 6.69.8.9 NMAC require school districts to measure the achievement gains of their students in all subjects and grade levels by administering a student assessment for each course offered (aside from subjects and grade levels required for the state student achievement testing programs). Student assessments may include:

- statewide assessments currently administered in mathematics and reading;
- other standardized assessments approved by PED, including nationally recognized standardized assessments;
- industry certification examinations; and
- PED-approved school district-developed or selected end-of-course assessments.

School districts may also develop a PED-approved assessment that measures student achievement growth for classroom teachers who do not teach in a standards-based assessment grade or subject.

PERFORMANCE IMPLICATIONS

Provisions of 6.69.8.9 NMAC indicate measures of student achievement growth will be measured by appropriate assessments of courses for classroom teachers or:

- the growth in achievement of the classroom teacher’s student on state assessments;
- the school’s A through F letter grade pursuant to 6.19.8 NMAC 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.

Provisions of the bill substituting PED-required end of course examinations with short-cycle assessments may affect performance measures relating to teacher and school ratings.

ADMINISTRATIVE IMPLICATIONS

School districts will need to develop short-cycle assessments for all grades to replace any PED-required end-of-course examinations. PED will need to develop a new teacher evaluation framework that follows the criteria set forth within the bill.

The current NMTEACH teacher evaluation framework implemented by PED is comprised of four categories: improved student achievement; classroom observations; planning, preparation and professionalism, and surveys and attendance. According to the PED website, each category is currently weighted according to the amount of student achievement data available for the teacher.

	Student Achievement	Classroom Observation Creating an Environment for Learning and Teaching for Learning	Planning and Preparation and Professionalism	Teacher Attendance and/or Surveys
Step 1: Teachers who have no student achievement in the last 3 years	0%	50%	40%	10%
Step 2: Teachers with 1-2 years of student achievement data (STAM) who teach courses related to STAM	25%	40%	25%	10%
Step 3: Teachers with 3 years of student achievement data who teach courses related to STAM	50%	25%	15%	10%

Provisions of the bill prohibit PED from adopting criteria that use student test scores, account for teacher use of personal or sick leave, and use numbers alone to evaluate teachers. As such, the student achievement and teacher attendance components of the current framework would need to be removed or replaced.

CONFLICT, RELATIONSHIP

This bill conflicts with HB125, which establishes a council to develop and recommend a new educator evaluation system; HB241, which would restrict types and amount of leave that can be considered in teacher evaluations; HB350, which changes teacher and administrator evaluation component weights; SB34, which establishes a temporary educator evaluation system and a council to develop and recommend a new evaluation system; and SB40, which adjusts factors and reporting requirements in the educator evaluation system and establishes a work group to study and recommend changes to the evaluation system.

This bill relates to HB124 which codifies the professional development dossier as the method for advancement within the three-tier licensure system; HB158, which establishes a teacher evaluation pilot project; HB163, which eliminates the inclusion of test scores of students with eight or more unexcused absences from use in teacher evaluations; and HB248, which prohibits the use of test scores, value-added methodology, school employees’ utilization of leave, and numerical teacher ratings as components of a teacher’s annual performance evaluation.

TECHNICAL ISSUES

The bill does not specify the frequency of short-cycle assessments or the entity responsible for developing or approving the use of these assessments. The Aspen Institute Education and Society Program; Achieve, Inc.; and National Center for the Improvement of Educational Assessment provided a policy brief in 2007 that provides some helpful definitions for summative, formative, and interim assessments, which may include an example of “short-cycle assessment” as intended by the sponsor:

- *Summative assessments* are generally given one time at the end of a specified period, such as a semester or school year, to evaluate student performance against a defined set of content standards. These assessments are typically administered statewide as part of an accountability program or used to inform policy.
- *Formative assessments* are used by teachers and students during instruction to provide ongoing feedback so teachers can adjust instruction to fit student needs. It is typically used by the teacher to diagnose student understanding, identify gaps in knowledge, adjust teaching strategy or approach, and evaluate student progress and learning. The assessments are generally embedded within a learning activity and linked directly to the current unit of instruction. Assessments are small-scale (usually a few seconds or minutes) and short-cycle (conducted multiple times within a lesson or unit of instruction). Tasks presented may vary from student to student depending on the teacher’s assessment. Providing corrective feedback, modifying instruction to improve the student’s understanding, or indicating areas of further instruction are essential aspects of a classroom formative assessment. Aggregating formative assessment information may not provide information that is useful beyond that specific classroom.
- *Interim assessments* may include medium-scale, medium cycle assessments that fall between formative and summative assessments. These assessments evaluate students’ knowledge and skills relative to a specific set of academic goals, typically within a limited time frame, and are designed to inform decisions at the classroom, school, or district level. Thus, they may be given at the classroom level to provide information for the teacher, but unlike true formative assessments, the results of interim assessments can be meaningfully aggregated and reported at a broader level. As such, the timing of the administration is likely to be controlled by the school or district, which makes these assessments less instructionally relevant than formative assessments. These assessments may serve a variety of purposes, including predicting a student’s ability to succeed on a large-scale summative assessment, evaluating a particular educational program or pedagogy, or diagnosing gaps in a student’s learning. Many assessments currently in use that are labeled “benchmark,” “formative,” “diagnostic,” or “predictive” fall within this definition.

OTHER SUBSTANTIVE ISSUES

The teacher evaluation process is being challenged in two lawsuits. The first lawsuit, brought by the American Federation of Teachers New Mexico, the Albuquerque Teachers Federation, and other plaintiffs, argues the state’s teacher evaluation system is unfair and could put teachers at risk of being punished or fired. The other lawsuit, brought by the National Education Association of New Mexico, claims the evaluation system unlawfully takes control of teacher evaluations and supervision away from local school districts.

In December 2015, state District Judge David Thomson granted a preliminary injunction preventing consequential decisions against teachers using the state’s teacher evaluation data until the state developed a reliable, fair, and uniform system. PED announced in January 2016 plans to simplify the evaluation system and make it more uniform across the state by reducing the number of tests included in calculating teachers’ scores, ending the use of student achievement data over a year old, removing a measure that evaluated teachers on students they had never taught, and releasing evaluation results in the fall rather than the spring. The American Federation of Teachers New Mexico case has been scheduled for a hearing on October 23, 2017.

A 2013 Measures of Effective Teaching (MET) Project report, funded by the Bill and Melinda Gates Foundation, noted:

“What counts most gets the most attention. When combining measures into a single index, we have found that approaches that allocate between 33 percent and 50 percent of the weight to student achievement measures are sufficient to indicate meaningful differences among teachers. Moreover, balanced weights avoid the risks posed by too narrow a focus on one measure. Overweighting any single measure (whether student achievement, observations, or surveys) invites manipulation and detracts attention and effort away from improvement on the other measures.”

Among the significant lessons learned through the MET project and the work of its partners:

- Student perception surveys and classroom observations can provide meaningful feedback to teachers. They also can help inform professional development decisions.
 - Students learn better in classrooms of teachers with better observation scores, better student survey results, and prior success in raising student test scores.
- Rigorous training and certification of observers and observation of multiple lessons by different observers are necessary to improve reliability and accuracy of evaluations.
- There is great potential in using video for teacher feedback and for the training and assessment of observers.
- Each measure adds something of value.
 - Classroom observations provide rich feedback on practice.
 - Student perception surveys provide a reliable indicator of the learning environment and give voice to the intended beneficiaries of instruction.
 - Student learning gains (adjusted to account for differences among students) can help identify groups of teachers who are helping students learn more.

Enactment of the Every Student Succeeds Act in 2015 removed state requirements to set up teacher evaluation systems based in significant part on students’ test scores, a key requirement of the U.S. Department of Education’s state-waiver system in connection with the No Child Left Behind Act.