



Date: November 16, 2017

Prepared By: Tim Bedeaux

Purpose: Examine the Innovative Assessment Pilot, the requirements of ESSA, reforms other states have made, and the work of innovative assessment groups in New Mexico.

Witness: Ian Esquibel, Contractor, Learning Alliance New Mexico, Tony Monfiletto, Director, New Mexico Center for School Leadership, Tori Stephens-Shauger, Co-Founder, Executive Director, and Principal, ACE Leadership High School, and Natalie Olague, teacher, Valle Vista Elementary School.

Expected Outcome: Understand the policy context and implications of using innovative assessments in New Mexico.

Senate Joint Memorial 1: Innovative Assessment Pilot

Background

The 2017 Legislature passed Senate Joint Memorial 1 (SJM1): Innovative Assessment Pilot. See **Attachment 1, SJM1: Innovative Assessment Pilot**. Citing increased flexibility under the federal Every Student Succeeds Act (ESSA), the memorial asked the Public Education Department (PED) and LESC to convene a diverse work group of teachers, researchers, community leaders, union members, and representatives from industry and higher education to continue work on developing innovative assessments. ESSA's Innovative Assessment Pilot presents a unique opportunity to pilot an assessment framework with local student success and educator input in mind.

Every Student Succeeds Act and Assessments

ESSA balances holding schools accountable for student outcomes using valid and reliable systems, while also supporting the needs of every student through local innovation and location-specific interventions. While requiring states to meet several criteria for statewide assessments, ESSA also offers incentives to state education agencies (SEAs) to pursue the development and implementation of innovative assessments.

Requirements of ESSA

To maintain compliance with Part A of Title I of ESSA, states must submit plans showing state assessments meet certain criteria to be used for federal accountability.

General statewide assessments must:

- **Be the same assessment for all students;**
- **Include English language arts (ELA) and math tests,** administered annually in third through eighth grade, and once in high school;
- **Include periodic science testing,** once in each of the grade ranges third through fifth, sixth through ninth, and 10th through 12th;
- **Be valid and reliable,** and consistent with nationally recognized professional and technical standards;
- **Assess all students,** with a participation rate of at least 95 percent;

New Mexico's ESSA state plan outlines how the use of the Partnership for the Assessment of Readiness for College and Careers (PARCC) assessment meets these guidelines for general education students. The PARCC ELA assessment is administered in third grade through 11th grade, and the math assessment in third grade through eighth grade with several subject-specific exams at the high school level. In 2017, 27 percent of students were proficient in PARCC ELA and 19 percent in PARCC math.

All students must be assessed using the PARCC, except for students with disabilities who take the New Mexico Alternative Performance Assessment (NMAPA), and English learners (ELs), who are assessed with ACCESS for ELs.

- **Accommodate all students**, including those with disabilities and English learners;
- **Provide individual student reports**, including descriptive and diagnostic achievement data; and
- **Allow results to be disaggregated** by local education agencies and schools.

Opportunities within ESSA

The Innovative Assessment Pilot is an option afforded by ESSA, and quickly became a topic of discussion for policymakers and stakeholders nationwide. The pilot program promises funds to up to seven states or consortia of states to experiment with innovative metrics for measuring student success, like performance-based or competency-based assessments. In addition to meeting the general assessment requirements of ESSA, innovative assessments must meet additional guidelines.

State proposals for innovative assessments must demonstrate:

- Evidence of consultation with experts and support from stakeholders, including teachers;
- Alignment, quality, and fairness of the test;
- Proof that standards and assessment requirements of ESSA are met;
- Demographic-related information from participating school districts;
- Descriptions of state rationale;
- Performance metrics for the test to ensure validity and reliability; and
- Availability of technology, expertise, funding, professional development, and other resources for implementation.

Innovative assessments can include a variety of assessment practices, as long as the assessments generate results that are “valid and reliable.” The U.S. Department of Education (USDE) has counseled states interested in piloting innovative assessments to start small, using innovative assessments in a small number of school districts or schools. Federal regulations for innovative assessments authored by the Obama administration remain intact, and require states to establish the innovative assessment’s validity, reliability, and comparability with the statewide assessment.

Establishing comparability can be done in a number of ways. While it is possible that pilot states will be required to simultaneously use both the innovative assessment and the statewide assessment for the duration of the implementation timeframe in pilot school districts, but the details of this provision remain unclear. States may opt to give both the state assessment and the innovative assessment to all students to establish comparability, but they may also do so by assessing only a small demographically representative sample of students taking both assessments. Other options also exist in regulation, like using shared “items and performance tasks” on both innovative and statewide assessments, or using an alternative method for demonstrating comparability that is “equally rigorous and statistically valid.” Determinations are made in states’ innovative assessment pilot applications, presenting an opportunity for states to use other methods of comparability.

States must scale the innovative assessment to statewide implementation over a period of five years. After statewide implementation, states must submit their innovative assessments for federal peer review. The review will determine whether the assessment meets the requirements of Part A of Title I of ESSA. States can request a two-year extension and an additional one-year waiver period to submit evidence for peer review, but the USDE secretary can withdraw a state’s innovative assessment flexibility if the state cannot demonstrate validity, scalability, or comparability of the innovative assessment system.

Since the change in the federal administration, USDE has remained largely silent on the program. According to Education Week, a spokesman from the current



administration has stated that the agency is considering what it will need to do to open the pilot and begin distributing funding in the 2018-2019 school year.

Grants for Enhanced Assessment Instruments

Separate from the Innovative Assessments Pilot, USDE offers grants to states for the development of Enhanced Assessment Instruments. The grant has been offered nearly every year since 2002, and is designed to help states improve the quality, validity, and reliability of assessments, measure student achievement using multiple measures, chart student progress over time, and develop comprehensive assessment instruments, including performance-based or technology-based instruments. Two states received awards in 2016, and five states in 2015. Applicants include SEAs and education research centers, often with multiple stakeholders submitting a joint application to undertake a large assessment endeavor.

Enhanced Assessment Instruments Grants in 2015 and 2016

| Year | State | Award Amount | Abstract |
|------|------------|--------------|--|
| 2016 | Maryland | \$3,843,805 | Create a learning map model for Next Generation Science Standards (NGSS), develop assessments, and design a dashboard for reporting student results. |
| 2016 | Nebraska | \$3,987,395 | Develop large-scale science assessments that can be shared across states and districts. |
| 2015 | Arizona | \$1,977,086 | Develop an English proficiency exam to track growth of English learners with cognitive disabilities. |
| 2015 | California | \$2,690,672 | Index results of assessments to include measurements of career readiness. |
| 2015 | Kansas | \$5,816,159 | Use learning maps to link curriculum, instruction, and assessment. |
| 2015 | Michigan | \$4,341,835 | Develop set of technology-interactive assessment tools aligned with NGSS. |
| 2015 | Minnesota | \$2,961,888 | Use classroom progress data to evaluate individual accessibility needs of students. |

Michigan and Maryland were both awarded funds in 2015 and 2016, respectively, to develop assessments aligned with the NGSS.

New Mexico plans to adopt similar standards, but has not yet identified a funding source for the development of an aligned assessment. The Legislature may want to consider options to obtain funds from an enhanced assessment grant.

Source: USDE

Innovative Assessments in Other States

New Hampshire

Widely regarded as a one of the states taking huge steps in innovative assessment development, New Hampshire designed and piloted a performance-based assessment model. The Performance Assessment of Competency Education (PACE) emphasizes local control over assessments, relying almost entirely on decisions made by local educators. During the 2015-2016 school year, New Hampshire piloted PACE in eight of its 166 school districts.

Educators develop assessments and submit them for validation by the New Hampshire Department of Education (NHDOE). Educators in PACE schools give assessments multiple times per week in the form of homework or quizzes. Schools participating in PACE also collaborate to develop a “common complex performance

PACE establishes comparability among schools' and school districts' results not by showing comparable results, but by holding school districts to comparable expectations. For example, a student who scores at a level 3 (meeting expectations) on PACE should score at that level in every other school district.

NHDOE shared results of a technical workshop demonstrating this point, in which randomly assigned pairs of teachers were responsible for evaluating over 400 different student papers. Teachers were able to reach a consensus score on all but three papers, validating the quality of the rubrics.

**Estimated One-Year Cost
of PACE Pilot
(Thousands)**

| | |
|----------------------------|-----------------|
| NHDOE Personnel | \$ 250 |
| Technical Consultants | \$ 1,600 |
| Grants to School Districts | \$ 300 |
| Communications, Publicity | \$ 150 |
| Total | \$ 2,300 |

| | |
|---------------------------------|-----------|
| Students Served (8 districts) | 13,596 |
| Per Pupil Expenditure (dollars) | \$ 169.17 |

Source: CADRE, NHDOE

task” to ensure that teachers’ evaluations of student performance are comparable to other teachers’ evaluations.

PACE employs rigorous quality control to ensure assessments are high-quality. Teachers involved in assessment development are given professional development on assessment development, and are given time to collaborate in creating performance tasks. Content area experts vet assessments with and facilitate the development of the common performance assessment. Additionally, the National Center for the Improvement of Educational Assessment (NCIEA), a national nonprofit assessment group, conducts technical reviews of the PACE common performance tasks. Final approval comes from NHDOE for all common performance tasks.

NHDOE estimated the cost of employing the PACE model at the current eight-district scale at \$169.17 per student for 13,596 students. A majority of the cost was for technical consultants, costs which NHDOE notes will recur and may rise as the pilot continues. Grants to school districts included stipends to teachers for summer work and the cost of local assessment work in each district.

Colorado

An August 2017 report by the Center for Assessment, Design, Research and Evaluation (CADRE) at the University of Colorado Boulder highlights work that Colorado has initiated in taking stock of the state’s local assessment plans. The report examines three school districts’ local assessment plans in detail.

Common, Integrated Interim and Summative Assessment. The Cherry Creek School District (CCSD) employs the ACT Aspire system as the primary indicator of student performance. Fourth grade through ninth grade students take ACT-aligned preparatory tests, 10th grade students take a pre-ACT test, and 11th grade students take the ACT. The ACT Aspire system offers interim and summative assessments in ELA, math, and science, and, the district argues, meets the validity and reliability requirements of ESSA.

Student-Centered Accountability Project (S-CAP). S-CAP is comprised of nine rural school districts that focus on “whole child development.” S-CAP districts must use the Colorado state assessment and two measures of student achievement, like commercial standardized test scores or graduation rates, to track student performance across schools. However, S-CAP schools also use a school quality review tool to evaluate student learning dispositions, professional culture, and allocation of resources.

Combination of Common and Local Assessments. Westminster Public Schools (WPS) participates in local assessments tied to curricular goals and competencies, an interim assessment from the Scantron Performance suite of assessments to compare performance across all schools, and the Colorado state summative assessment. WPS

The nine Colorado S-CAP districts are Buena Vista, Buffalo, Huerfano, Julesburg, Kit Carson, Mancos, Monte Vista, South Routt, and La Veta.



requested flexibility to administer the state assessment at different times throughout the year. Noting that students learn at different paces, WPS assesses students when they have demonstrated competency and the readiness to progress to the next grade.

Recent Work of New Mexico Groups

Learning Alliance New Mexico

On October 11, 2017, a group of education stakeholders from across the state met at the National Hispanic Cultural Center in Albuquerque to discuss alternative assessments. The discussion, facilitated by Learning Alliance New Mexico, centered on rethinking the role that assessments play in improving education. During introductions, attendees were asked to name either a positive or negative experience related to assessments. Negative responses very quickly outnumbered the positive, with themes emerging. Parents recalled the stress “testing season” put on their children. A former student shared a story about having a panic attack during standardized testing. Teachers recounted the pre-test stress and post-test lull, and the shared concerns with having teacher evaluation results be based on student proficiency. Stakeholders expressed general distrust with the PARCC exam, stating proficiency on an exam does not necessarily equate to college or career readiness. Instead, individuals offered ideas like individualized performance-based or competency-based exams, including discussions of using student portfolios to determine college readiness.

Mission: Graduate and the New Mexico Center for School Leadership

A September 2016 joint report by Mission: Graduate and the New Mexico Center for School Leadership found strong support for differentiated ways to measure student achievement. The report, titled “What We Want to See in Schools,” follows a variety of Albuquerque business and academic stakeholders on a school innovation tour through eight Albuquerque-area charter schools. During post-tour interviews, the tour participants supported performance-based alternatives to traditional testing. Tour participants cited the need for students to show problem solving and critical thinking skills to work on a problem that may not have a single answer. Participants noted a disconnect between standardized testing and actual workforce performance, with local employers explaining they would rather have employees who are capable of performing a task than those who can demonstrate that task on a test.

During post-tour interviews, business stakeholders offered methods of performance assessments they use in their own workplaces. Stakeholders introduced the prospect of applying cyclical performance reviews or self-evaluation in a school setting. They believed qualitative assessments would help students “feel more ownership” of their education, and that constant feedback was important for student development.

Still, interviewees recognized quantitative challenges with individualized assessments. Even though personalized feedback is important on a student level, the system can lose the ability to benchmark student performance on a large scale. The use of common measures is important to compare students statewide and understand which schools are performing well, and which schools need improvement.

1 A JOINT MEMORIAL

2 REQUESTING THE PUBLIC EDUCATION DEPARTMENT AND THE
3 LEGISLATIVE EDUCATION STUDY COMMITTEE TO CONVENE A WORKING
4 GROUP TO DEVELOP RECOMMENDATIONS FOR A NEW STUDENT ASSESSMENT
5 POLICY ALIGNED WITH THE ALTERNATIVE ASSESSMENT MODELS ALLOWED
6 BY THE FEDERAL EVERY STUDENT SUCCEEDS ACT.

7
8 WHEREAS, the recently enacted federal Every Student
9 Succeeds Act empowers states to develop their own
10 accountability systems to measure student performance and
11 encourages states to begin pilot projects that use a variety
12 of ways to measure student success; and

13 WHEREAS, for states wishing to develop alternative,
14 evidence-based assessment systems, the Every Student Succeeds
15 Act creates an enhanced assessment instruments program, which
16 will provide funding to seven states to develop pilot
17 programs of innovative assessment programs that use a variety
18 of indicators to show how a student is performing; and

19 WHEREAS, applications for the pilot programs were
20 accepted through September 22, 2016, and it is anticipated
21 that additional applications will be accepted in the future;
22 and

23 WHEREAS, the Every Student Succeeds Act provides states
24 with the flexibility to adopt high-quality academic
25 assessments in mathematics, reading, language arts or

SJM 1

1 science, which may allow students to demonstrate proficiency
2 in the form of portfolios, projects or extended performance
3 tasks and other locally designed formative assessments that,
4 along with standardized tests, create a summative score for a
5 student; and

6 WHEREAS, the Every Student Succeeds Act permits local
7 educational agencies to choose a locally selected assessment
8 in lieu of the state assessment, but only if the local
9 educational agency "selects a nationally recognized high
10 school academic assessment that has been approved for use by
11 the state"; and

12 WHEREAS, the Every Student Succeeds Act requires each
13 state to add at least one non-academic indicator of school
14 quality or student success as part of its statewide
15 accountability system, opening the door to the possibility of
16 using factors such as student engagement, post-secondary
17 readiness, school climate and safety, career readiness and
18 other metrics to measure student success; and

19 WHEREAS, at this time, New Mexico is not yet prepared to
20 take advantage of the flexibility provided in the Every
21 Student Succeeds Act and the associated pilot program funding
22 opportunities; and

23 WHEREAS, studies by learning alliance New Mexico and the
24 New Mexico center for school leadership show that there is a
25 demand throughout the state for alternative assessment models SJM 1

1 that use a variety of approaches to measure a student's
2 progress; and

3 WHEREAS, while the New Mexico center for school
4 leadership is working with stakeholders to develop
5 alternatives to the current assessment process, that project
6 is local to Albuquerque, and a broader statewide effort,
7 including participation from the diverse communities in the
8 state, has not been made; and

9 WHEREAS, statewide studies, research and analyses are
10 necessary for New Mexico to be ready to propose a pilot
11 program and take advantage of the new flexible assessment
12 possibilities provided by the Every Student Succeeds Act;

13 NOW, THEREFORE, BE IT RESOLVED BY THE LEGISLATURE OF THE
14 STATE OF NEW MEXICO that the public education department and
15 the legislative education study committee be requested to
16 convene a working group composed of educational
17 practitioners; academic and pedagogical researchers;
18 community, tribal and civic leaders; union members; industry
19 representatives; higher education representatives; and others
20 who have an interest in developing an alternative assessment
21 model to build upon the work that is already taking place in
22 the state and to develop recommendations for a new assessment
23 policy; and

24 BE IT FURTHER RESOLVED that the working group be
25 requested to report its recommendations and a proposed new

SJM 1

1 assessment policy to the legislative education study
2 committee by November 2017; and

3 BE IT FURTHER RESOLVED that copies of this memorial be
4 transmitted to the secretary of public education and the
5 director of the legislative education study committee. SJM 1

6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25