PERFORMANCE ASSESSMENT SYSTEMS

EXAMPLES FROM FOUR STATES

ECP

Education Commonwealth Project

Performance Assessment Systems: Examples from Four States

Education Commonwealth Project

University of Massachusetts Lowell
Coburn Hall 222, 850 Broadway St.
Lowell, MA 01854
EdCommonwealth.org

Authors

Dan French, Director of Performance Assessment, danvfrench@gmail.com
Susan Lyons, Technical Adviser, susan@lyonsassessment.com
Sanford R. Student, Research Associate, sandy@lyonsassessment.com

About the Education Commonwealth Project

The Education Commonwealth Project (ECP) works to support assessment of student learning and school progress that is valid, democratic, and equitable. Pushing back against the overreliance on standardized testing, ECP offers free and open-source resources that all schools can use. And thanks to support from the Massachusetts State Legislature, ECP offers additional support for public schools and districts in Massachusetts.

Massachusetts

System Overview

The Massachusetts Consortium for Innovative Education Assessment (MCIEA) believes there are richer means of assessing student and school progress than high-stakes state standardized tests. The state's current assessment system fails to accurately reflect school quality and student learning for all students, particularly for students from historically marginalized groups, including students of color, low-income students, multilingual learners, and students with disabilities. MCIEA believes all students should have access to equitable and transparent education communities with authentic, fair, and responsive learning and assessment systems. Robust measures of accountability should highlight strengths and areas for growth of students and schools. Thus, MCIEA seeks to build a more robust system of multiple measures of student learning and school quality. The consortium is a partnership of public school districts and their local teacher unions from Attleboro, Boston, Lowell, Milford, Revere, Somerville, Wareham, and Winchester.

MCIEA Principles

- Community members identify what is most important to know about school quality
- Multiple measures provide a robust picture of student learning and school progress
- Local leaders, teachers, parents/guardians, and students use data from multiple measures to make decisions that meet the assets and needs of their schools and communities
- Students demonstrate what they know and can do through real-world application of teacher-designed, curriculum-embedded performance assessments
- Benchmarks based upon the full characteristics of a high-quality school establish a fair measure for school performance
- State support and resources, rather than high stakes tests and sanctions, build capacity of schools and districts, which leads to improvement

MCIEA is building a robust assessment system through two overlapping frameworks: School Quality Measures (SQM) and Quality Performance Assessment (QPA). Together, QPA and SQM make up a holistic form of student and school assessment that fosters a more accurate and nuanced understanding of strengths and areas for improvement for both students and schools.

Drawing on multiple data sources, SQM captures school strengths and areas needing improvement across School Culture, Community & Wellness, Resources, Teachers & Leadership, and Academic Learning. School communities examine this data, celebrate strengths, engage in data-based inquiry to uncover causes of identified gaps, and create and

implement plans to address them. SQM assesses school quality in a fair and comprehensive way that reflects the unique character of each school community.

To assess student learning, educators design standards-based, culturally responsive performance assessments, and submit them to a MCIEA Performance Assessment Task Bank for peer review and approval. Teachers administer these tasks for students to demonstrate what they know and can do in ways that are authentic, culturally responsive, and engaging. Looking to the future, if this performance assessment model were adopted statewide, teachers across schools and districts would be regularly brought together to engage in blind scoring of student work to ensure scoring reliability. High school graduation decisions would be made at the local level, using in part portfolios of student work. The state's role would be limited to provide districts with resource support and technical assistance. State assessments would be used for diagnostic purposes to provide a consistent data source on student learning to schools and educators, without the requirement of passing state tests to graduate high school.

Curriculum & Instruction

MCIEA embraces a belief that inquiry- and project-based instruction should go hand in hand with performance assessment, and that curriculum, instruction and assessment should be seamless, with performance tasks a culminating activity at the end of curriculum units.

Professional Development

In order to build teacher capacity in curriculum-embedded, standards-based performance assessment design, MCIEA brings together lead teams of teachers and an administrator from MCIEA schools to participate in a hands-on, four-day institute spread across summer and school year to learn performance assessment design, task validation (e.g., alignment, Universal Design, engagement, fairness) and reliable scoring of student work. The teacher lead teams are then supported to engage the entire faculty in this same performance assessment capacity building. Volunteer teachers are also trained and stipended to be peer reviewers of teacher-generated performance tasks that are submitted to the MCIEA Performance Assessment Task Bank, assessing tasks against a set of criteria for high quality performance tasks and giving feedback to teacher authors on how to strengthen the task prior to approval for uploading to the task bank.

Technical Quality

Teachers receive training and protocols on how to design a quality performance task, including task alignment (validity and fairness) and scoring of student work (inter-rater reliability).

Impact

MCIEA has yet to measure impact on student learning, graduation, and college-going. A survey found that MCIEA teacher leaders agreed they had improved their skills and knowledge to create performance assessments that adhere to all five key elements of a quality performance assessment (validity, reliability, data analysis, fairness, and student voice and choice), and that teacher leaders are successfully embedding this work in their practice. A second survey found that almost two-thirds of all MCIEA teachers agreed their skills and knowledge improved in all five areas of the Performance Assessment Literacy Scale. In a separate small pilot study, teachers felt that tracking student learning progress using teacher-developed, multi-grade, and standards-based learning progressions was a useful practice and provided added support for their conversations about student learning during parent/guardian conferences. Rather than relying on external instruments to provide growth determinations, teachers reported they were glad to have multiple examples of student work aligned to a content-based learning progression so they could talk specifically about student progress.

What Can We Learn from MCIEA?

- Teacher unions having a seat at the decision-making table, as well as superintendents communicating to educators that the district is adopting performance assessments, ensures there will be more buy-in from teachers in engaging in performance assessment work.
- Training lead teacher teams from participating schools who participate in building faculty-wide performance assessment capacity can be an effective means of school-wide adoption of performance assessment practices.

Colorado

The Colorado Department of Education (CDE) is currently engaged in three ambitious initiatives related to promoting the use of high-quality performance assessment. One of these initiatives introduces performance assessment as one of several ways that students can demonstrate college and career readiness in order to graduate from high school.

System Overview

As of 2021, Colorado provides its students with a menu of eleven options for high school graduation. One of these options is a Collaboratively-developed, Standards-based Performance Assessment (CSPA), which is broadly defined as the creation of a complex product or presentation that involves application of Colorado's Essential Skills for Postsecondary and Workforce Readiness. A main goal of providing performance assessment as a menu option is to provide students with a "content-agnostic, student-centered" assessment option that "values local context," making CSPA–if implemented well–a truly equitable method of assessing readiness to graduate high school for Colorado's students. To ensure that performance assessments meet expectations for rigor and equity, Colorado has developed design elements that describe performance assessments in terms of being ready to use or in need of revision across six distinct criteria. These design elements are meant to guide local educators toward the creation of high-quality, relevant performance assessments that allow students to demonstrate both specific academic competencies and the non-academic Essential Skills linked above.

Implementation of CSPA as a graduation option in schools is in its early stages. A series of reports conducted by the Center for Assessment, Design, Research and Evaluation (CADRE) at the University of Colorado Boulder summarizes findings from the first two years of implementation, which included both statewide professional learning about performance assessment in a cohort-based model and the development of pilot performance assessments in two high schools. Currently, CDE continues to work with CADRE and is now conducting exploratory analyses of the technical properties of students' scores on CSPAs as implemented in a rural district in southern Colorado.

Curriculum & Instruction

Colorado's performance assessment program is designed to be content-agnostic, which means that it is curriculum-agnostic as well. At the same time, the use of performance assessment was expected to lead to improvements to teaching and learning, only some of which were realized during the initial pilot. It became clear that performance assessment alone cannot compensate for poor formative assessment practices, and that students being introduced to performance assessment after an academic lifetime of more familiar assessment formats need

support from educators in order to be able to engage with complex tasks. Going forward, recommendations from CADRE's initial reports are being used to inform expansion of the performance assessment pilot.

Professional Development

One lesson on professional development that can be learned from Colorado is the importance of support at the local level for educators implementing performance assessment. As part of the first two years of its performance assessment pilot, Colorado ran statewide professional learning communities for educators looking to implement performance assessments in their local contexts. CADRE's research into implementation determined that although the statewide PLCs promoted a shared vision for performance assessment in Colorado, it was local support–that is, district-run PLCs in addition to statewide PLCs–that ultimately led to the highest quality implementation of performance assessment. Putting structures in place for educators to collaborate with their school and district colleagues appears to be crucial to successful implementation.

Technical Quality

Colorado's [design elements](#) describe performance assessments in terms of being ready to use or in need of revision and provide guidelines on what makes for a high-quality performance assessment. CADRE is continuing its analysis work and is currently investigating statistical properties of scores on one district's performance assessments.

Impact

Student surveys from the first two years of the pilot underscored the importance of support for both educators and students when transitioning to a performance-based approach to assessment. Work is now ongoing to understand the impact of the pilot on student graduation trends in a district where implementation was strongest. Results of these analyses are not yet available.

What can we learn from Colorado?
- Performance assessment may be a viable alternative to large-scale standardized testing for demonstrating college and career readiness (what Colorado calls "postsecondary and workforce readiness").
- Colorado's design elements demonstrate methods for assessing the technical quality of performance assessments, and could easily be adapted to skillsets beyond Colorado's PWR skills.
- Local professional learning communities appear far more effective in supporting implementation of performance assessment than statewide or otherwise dispersed PLCs on their own.

New York

System Overview

In agreement with the New York State Education Department, 38 [New York Performance Standards Consortium](#) ("Consortium") public high schools, mostly within New York City but also Rochester and Ithaca, are released from administering four of the five required Regents exams (schools are required to still administer the English language arts exam). In their place, Consortium schools administer a set of common performance assessment tasks commonly known as Performance-Based Assessment Tasks, or PBATs, which all students must pass in order to graduate. The Consortium website states the consortium "is the only network of schools in the nation that has put into practice [a state-approved] authentic assessment option other than standardized testing."

In addition to passing courses, there are four cross-school tasks every Consortium high school student must complete and pass to graduate. Students may complete them at any time in 11th and 12th grade. The tasks are broad and flexible, enabling teacher latitude to embed them in the curriculum as well as for students to devise their own questions. Tasks require both written and oral presentations. Written accommodations exist for multilingual learners and students with disabilities. Each task has required common rubrics for written work and for oral presentations, to be used for scoring. Each PBAT written paper is scored by the assigning teacher and another teacher of the same subject. Once a student's written work has been rated as "competent" or higher by both reviewers, the student orally presents their work before at least two evaluators, a teacher of the same subject and an adult knowledgeable in the field.

- *English language arts: Analytic Essay* – Literary analysis using a text of appropriate complexity.
- *Social studies: Research Paper* - Analytical paper, with a thesis, argument, and accurate evidence.
- *Science: Experiment* – Designed and implemented extended science project or original experiment.
- *Math: Applied Math* – Solution of math problem that requires application of higher level math.

Consortium schools may also create and administer additional graduation performance tasks in other disciplines, such as the arts or world languages. All schools use a mix of local and Consortium performance assessment tasks to determine student readiness to graduate.

Curriculum & Instruction

According to the Consortium website (2022), "In the Consortium's performance assessment system, assessment tasks *grow out of* the work of the classroom. They are *not imposed on* curriculum, a process that inevitably leads to teaching-to-the-test. Tasks become possibilities for assessment only after students and teachers have studied the material, discussed and debated it, and subjected it to their questions and writing. All content areas are inquiry and literacy-based, focus on extensive reading, writing, discussion, and critical thinking, and build towards the graduation-level performance-based assessment tasks (PBATs)....[A] series of [school-based] interim assessments, roundtables, classroom argumentation based on content and evidence, creative and first-person writing, and hands-on projects all prepare students for their final PBATs."

Professional Development

The Consortium conducts professional development for teachers on designing quality performance tasks and scoring student work, and encourages schools to send teams to observe assessment practices in other Consortium schools.

Technical Quality

Biannually, about 200 teachers across Consortium schools convene to engage in a moderation study, or "blind" rescoring of student work to establish inter-rater reliability. Teachers also examine the tasks to ensure they elicit complex thinking from students. Using Consortium protocols, schools also conduct their own within-school inter-rater reliability sessions among teachers.

Every five years, a Performance Assessment Review Board, an external body of educators and community leaders, monitors the performance-based assessment system and conducts sampling of student work. A PAR team visits each school to ensure it has the required components in place: active learning, formative and summative data, strategies for corrective action, multiple ways to exhibit learning, graduation performance tasks, external evaluators, and professional development.

Impact

According to the Consortium website (2022), Consortium high schools serve a higher percentage of Latinx students, multilingual learners, students with disabilities, students in temporary housing, and low-income students than aggregate New York City high schools, and incoming students have lower average math and English language arts test scores (2018-2019 school year). Yet, Consortium schools have higher aggregate and disaggregated graduation rates.

	Consortium	Citywide
4-Year Graduation Rate for 9-12 High Schools*	84.0%	77.3%
4-Year Grad Rate — Black*	84.0%	73.7%
4-Year Grad Rate — Hispanic*	80.0%	72.0%
4-Year Grad Rate — English Language Learners	75.0%	69.0%
4-Year Grade Rate — Students with Disabilities	73.0%	68.0%

Source: These results were calculated using publicly available NYC DOE data (Accessed November 2019 and January 2020): 2018-2019 School Quality Guide Citywide Data; High School Citywide Results; Transfer High School Citywide Results; * New York City Graduation Rates Class of 2019 (2015 Cohort)

Note: Consortium graduation data based on 2 or more years of enrollment in a Consortium member school.

As well, In their report _Assessing College Readiness Through Authentic Student Work_, Michelle Fine and Karyna Pryiomka found that Consortium "students begin high school with more marginal academic records but graduate, enter college, persist in college, gain credits, and sustain higher GPAs than their peers….Black males, in particular, benefit from a Consortium education when compared to Black males educated in traditional high school settings: They are noticeably more likely to persist in college and to receive higher grades."

What Can We Learn From NYPSC?

- Through a concerted, multi-year advocacy effort, the Consortium was able to successfully receive a long-term waiver from the New York State Department of Education to forego the state Regents exams, except for English language arts, as a high school graduation requirement. It's possible a similar effort could be launched by MCIEA and ECP districts for their high schools to be released from passing the MCIEA tests.
- The Consortium has teamed with education researchers to document enrollment, graduation, college-going, and college persistence rates of Consortium students, providing evidence of their success when compared to aggregate NYC public high school students.
- The Consortium performance assessment system provides a robust system of professional development in performance assessment design, administration, and scoring, and embraces an instructional philosophy of inquiry-based learning, a key component of high quality performance assessment systems.

Michigan

System Overview

[Michigan Arts Education Instruction & Assessment (MAEIA)](#) is a non-profit organization that partners with the state department of education and the Michigan Assessment Consortium to promote the use of performance-based assessment for arts educators in Michigan. MAEIA offers a catalog of freely-available performance assessments developed to assess student learning in K-12 arts. The catalog includes more than 360 assessments in the disciplines of dance, music, theater and the visual arts.

Curriculum & Instruction

The performance assessments were developed by Michigan educators and are designed to be adapted to fit local contexts and curricula across the state. MAEIA encourages educators to use these performance tasks as summative assessments in their classrooms.

Professional Development

The organization offers support to schools and educators by offering professional learning resources that include personalized consultations and access to online courses and learning modules. Additionally, MAEIA offers guidance and resources for how performance assessment can be used as common assessments in districts, used to evaluate program quality, and as evidence for demonstrating educator effectiveness for use in educator evaluation.

Technical Quality

The evidence supporting the technical quality of the MAEIA assessments is strong. The organization provides transparent access to the detailed assessment specifications that were developed to ensure the assessments are aligned to the full breadth and complexity of the content standards. Each of the performance tasks was developed to align to these specifications, has undergone an item review process, was field tested in schools and collaboratively scored. All of the assessments come with administration procedures, a defined set of student products to be evaluated, and a set of teacher scoring rubrics.

Professional Learning—Engagement with Educators

Customized professional learning sessions and presentations
690 PARTICIPANTS

Arts Lounge online forum
148 PARTICIPANTS

EduPaths self-paced online courses
148 PARTICIPANTS

Website

Website visits
21,295

Website users
17,075

Total pageviews
43,507

HIGH ENGAGEMENT WEBPAGES
Performance Assessments
25,798

MAEIA homepage
7,197

Resources and performance assessments accessed
4,168

Communication

Subscribers to **MAEIA's newletter**, *MAEIA News and Updates*
2,920

Facebook total number of daily pageviews
207

Blog post views
4,419

Twitter impressions
3,064

Impact

While impact on student learning is not something that MAEIA is able to measure, the impact of MAEIA can be quantified by its impressive reach to arts educators in Michigan and beyond. Pictured to the right is a snapshot of their engagement data from 2021. Source: MAEIA 2021 Annual Report.

What can we learn from MAEIA?

MAEIA provides a model for supporting the use of performance assessment in Massachusetts districts in many ways. One of the notable strengths of MAEIA is the collaborative scoring platform it provides to support the inter-rater reliability of scoring student artifacts and performance. The Michigan Collaborative Scoring System (MI-CSS) is a technology-based solution for scoring student work. The goal of MI-CSS is to provide independent teacher scoring of student written work and performances at a lower cost than centrally-managed scoring at the state level. The online system allows educators to upload student work into MI-CSS in a variety of formats such as written work, video clips, and/or audio clips. Once uploaded, teachers score their own students' work, using the rubrics embedded within the software. Teacher-scored work is then anonymously and independently scored by other participating educators who teach the same courses in different schools. In cases where the two independent scores are not the same, resolution scoring is offered by a program leader.

Education Commonwealth Project
University of Massachusetts Lowell
Coburn Hall 222, 850 Broadway St.
Lowell, MA 01854
EdCommonwealth.org