Graduation Requirements and Measures

A Review of Performance Assessment Implementation in Select States for the New York State Education Department

Region 2 Comprehensive Center

August 2021

Jenna Howard Terrell
Senior Research Associate

Sarah Barzee
Director, R2CC Comprehensive Center

Executive

Network (PLAN) Pilot team last year. In consultation with NYSED, the R2CC team then narrowed the scan to seven states to conduct deeper case studies with state and consortium leaders. The case studies consisted of interviews and focus groups with state education agency (SEA) staff or consortium leadership to understand implementation processes, successes, areas for improvement, and lessons learned. This report outlines the general themes that emerged from the case studies, followed by state profiles (which detail more information about the implementation of performance-based assessments) for each of the states and the consortium included in the data collection.

General Case Study Findings

Despite the case study states taking different approaches to implementing performance-based assessments, a common theme shared by the states was that performance-based assessments are used to provide more opportunities for students to demonstrate mastery of skills and content. Case study states integrate performance-based assessments as a complement or an alternative to traditional, multiple

Introduction

In 2015, the enactment of the Every Student Succeeds Act (ESSA) created a renewed interest in performance-based assessments across the country after a focus on standardized, multiple-choice assessments under the No Child Left Behind (NCLB) Act (Blinet al., 2020; Parn & Darling-Hammond, 2015). Many states, districts, and institutes of higher education recognized that using only standardized, multiple-choice tests did not demonstrate the full range of students' skills and knowledge (Guha et al., 2018). Performance assessments are one type of assessment that states and districts are exploring and incorporating into their assessment systems following NCLB under the Innovative Assessment Demonstration Authority (IADA).[1]

Performance assessments describe a continuum of activities to demonstrate mastery and proficiency, ranging from constructing a response to creating a product to performing an activity (Darling-Hammond, 2017). Performance assessments can include tasks such as essays, authentic demonstration of learning (often found in career and technical education), portfolios, exhibitions, and capstone projects (where several performance tasks culminate into a product). Although there are some nuances in the literature about what a performance-based assessment must entail (Bland & Gareis, 2018), for the purposes of this report to NYSED, we conceptualized performance-based assessments according to the definition at the New York State Education Department (NYSED).

New York State Education Departments definition of
performance-based assessment

A performance-based assessment requires students to demonstrate or apply their knowledge, skills, and strategies by creating a response or product or doing a task. Students responses or performances are typically judged against standards or criteria in a checklist or rubric focusing on the stages of skill development and what a student can do

---

[1] We note that many states incorporated performance-based assessment prior to NCLB, though the focus shifted to standardized assessments under NCLB.

Since the enactment of ESSA, researchers have focused on descriptive studies of performance-based assessments or case studies within individual states. For example, Stosich and others (2018) identified four strategies used by 12 states, including supporting classroom instruction, graduation requirements, school accountability, and federal accountability. At the time of publication, most states were incorporating performance-based assessments as part of their classroom instruction or school accountability (Stosich et al., 2018). Other researchers have studied districts or consortia to understand if and how performance-based assessments support deeper student learning (Bland & Gareis, 2018; Evans, 2019; Guha et al., 2018; Kim, 2005; Maier et al., 2020; Marion & Leather, 2015). Each of these studies has demonstrated that performance-based assessments are a promising practice for measuring higher-order thinking, depth of knowledge, and college and career readiness.

Yet, the policy context is changing as states and districts are rethinking how to implement performance-based assessments to measure mastery of content and skills after NCLB. As a result, many studies published in the last five years are already outdated. This report contributes to the growing research on performance-based assessment, with a focus on implementation beyond smaller-scale initiatives. We highlight specific state-level policies that include performance-based assessments as part of a more robust assessment system, including the challenges and successes across the states, and lessons learned. Our primary goal is not to advocate for performance-based assessments but to support the Commission in their information gathering as they rethink NYSED graduation requirements.

Methods

This report addresses the following questions:

> How do the identified states use and integrate performance assessments as part of a measurement system for students to demonstrate college and career readiness?

> What can New York (and other states) learn about performance assessment implementation (i.e., challenges, successes) from the identified states?

The study began with a landscape scan of 12 states identified by NYSED through the PLAN Pilot exploratory phase. These states are in two "tiers," the first including states with established, statewide requirements or options for performance assessments (Colorado, Oregon, Rhode Island, New Hampshire, and Virginia) and the second including states with emerging performance assessment options or frameworks (New Mexico, Kentucky, Vermont, Massachusetts, Connecticut, Ohio, and Washington). The landscape scan included a review of publicly available websites, documents, reports, and policies to understand the purpose and history of performance assessments in each state, specific policies in support of performance assessments, and how the assessments were implemented statewide. After completing the landscape scan, the R2CC team, in consultation with NYSED, selected seven states (Colorado, Massachusetts, New Hampshire, Oregon, Rhode Island, Virginia, and Vermont) in order to conduct a more in-depth data collection through focus groups and interviews with state department agency staff members. The interviews and focus groups lasted no more than one hour and were conducted by the R2CC team using a semi-structured interview protocol (see Appendix B).

---

[2] Note that the in-depth case study included the Massachusetts Consortium for Innovative Educational Assessment (MCIEA), which is a consortium of districts implementing performance assessments with support from the MCIEA.

Case Study Findings

How do the identified states use and integrate performance assessments as part of a measurement system for students to demonstrate college and career readiness?

The states take different approaches, but all use performance assessments to provide more opportunities for students to demonstrate mastery of skills and content

Each state representative who participated in the state interviews expressed that their goal for performance-based assessments was to provide districts, schools, and students with additional ways to demonstrate learning. Some also mentioned that the goal of performance assessments was to provide a more meaningful way to assess deeper learning compared with traditional, multiple-choice tests.[3] States like Rhode Island, New Hampshire, and Colorado implemented performance assessments because of new legislation requiring proficiency- or competency-based education, updated graduation requirements, or essential skills. These representatives said that the new policy left a gap in assessing this new approach to learning and instruction, they sought an alternative that provided more flexibility to demonstrate proficiency and mastery of skills. It is worth mentioning that not all of the states included in the case studies were using performance assessments for graduation specifically, but the state representatives still highlighted that the skills and content measured by performance assessments in their state were capturing the key skills they wanted their high school graduates to leave with to prepare them for college and/or the workforce.

Another commonality shared by the case study states is that performance assessments are developed and selected locally. Each state emphasized that the power of performance assessments is that they are developed locally and linked to classroom instruction. As one state representative noted, "What we don't want is for these to become the state assessment [that] replace the local work, because the power is in the teachers really thinking about the students and what they want their students to demonstrate." For a quick guide as to how each state in the study uses performance-based assessments, see the column titled "Purpose" in Table 1 at the end of this section

---

[3] This was often mentioned in contrast to standardized tests administered statewide or as a requirement for graduation

"We're trying to create opportunities so stu itsoen2(o)1(p)ß(il(s)1( )v1(p))-1(e)2( )-2

Other states don't have a formal policy or initiative, but encourage the use of performance assessments for classroom instruction. Vermont and Rhode Island, for instance, have a proficiency-based education system and encourage performance assessments that are embedded in the classroom to demonstrate mastery for graduation, though there is not currently a formal policy, statute, or initiative to support statewide implementation.

> *"We're not running away from standardized measures, we would just like them to be used sensibly and placed properly in our overall system [to measure learning] for kids."*
>
> — *SEA Representative*
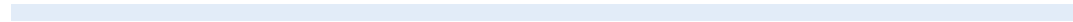
The information synthesized through the scan and case studies demonstrated that states are

Some states (such as Colorado) emphasized the importance of a statewide approach to implementation that included teachers in developing quality criteria and determining "non-negotiables" for performance assessment design, which they called a key factor in the development of high-quality performance assessments.

Of the states reviewed by R2CC, four (Colorado, Massachusetts, New Hampshire, and Virginia) had established systems for ongoing professional learning and provided resources through the state education agency or an intermediary. Colorado, Massachusetts, and New Hampshire use professional learning communities and intermediaries to continue to provide support to

State	

State	Initiative	Policy	Rollout	Purpose (Stosich et al., 2018)	Subject/ grade

> *"You can't just have these big lofty goals around what you want for your graduates and then [not align those goals to] the opportunities you provide them for learning and for demonstration. So that's where we're seeing a lot of shifts being made. People are saying, 'These are the skills we want our kids to walk out of our buildings with?' In order to get there, we have to change things. And that includes looking at curriculum, looking at the way students are engaged, looking at opportunities for performance assessment, or competency-based learning, or whatever that pathway is."*
>
> — *SEA Representative*

Collaboration and Involvement With Teachers and Administrators Is Crucial

Interviewees also mentioned that collaboration among teachers, schools, and districts is crucial for the successful implementation of performance-based assessments. First, involving teachers in decision-making processes and providing them with resources, support, and time to create high-

Furthermore, at least three states highlighted that teachers who actively engaged in the design processes acted as "ambassadors," continuing to lead and support other teachers within their districts and in other districts. In Rhode Island, a former principal noted that teachers in their school would

State interviewees also highlighted that creating a clear rubric and providing the rubric to students can reduce this grading bias. At the state level, the development of standard practices and processes to norm and score tasks and assessments can also provide clear guidance to address equity concerns around scoring. Finally, states mentioned the ongoing capacity building and the use of an intermediary to drive professional learning for performance assessments as a way to alleviate unequal quality of implementation. They stressed that having a process to know and understand where implementation was successful and where it needed improvement, as well as the capacity and resources to address it, was critical to ensuring equity in access to high-quality performance assessments.

Some interviewees also called out the importance of multiple pathways and diverse demonstrations of knowledge in order to avoid a gatekeeper mentality. They also emphasized the importance of connecting with communities historically marginalized by assessment practices and incorporating community perspectives and ways of knowing into the assessment process as a way to address the first equity concern. Other states noted that special attention should be given to addressing the needs of diverse student populations, including language learners and students with disabilities. Performance assessments should be designed to be inclusive and provide opportunities for all students to demonstrate their knowledge and skills. For example, Colorado and Oregon emphasize the importance of considering language learners by allowing performance assessments for some subjects (e.g., English language arts in Oregon) in their home language.

"One of the things to think about any time we're doing large-scale assessments that have significant consequences, such as graduation, is making sure we are connecting with communities who have been traditionally and historically harmed by the assessment process—being able to honor and affirm different identities and ways of knowing and then being able to have that surface in the performance assessment and the tool that's being used to measure it . . . Because without those pieces and without that conversation, we will continue to perpetuate harm."

— SEA Representative

Appendix A: State Profiles

Colorado Tier 1

History

In 2019, the Colorado Department of Education launched a pilot initiative with five schools across four districts to develop and implement collaboratively developed, standards-based performance assessments (Diaz-Billeo & Pierre-Louis, 2021). The catalyst for this work was the new Graduation Guideline Menu of Options, which included performance-based assessments as one option for students to demonstrate postsecondary and work readiness. During the pilot, participants identified the essential skills that aligned with Colorado Academic Standards and would be assessed by the performance-based assessments. Additionally, participants developed a common rubric to assess and score quality (Diaz-Billelo et al., 2021). After the two-year pilot, the work shifted to developing professional learning communities to develop and norm performance-based assessments across the state. This work has continued with educators from 38 school districts.

Relevant Policy/Legislation

Colorado does not have a formal policy or statutory requirement to develop or implement performance-based assessments. These assessments are included as an option for students to demonstrate postsecondary and workforce readiness. As a local-control state, Colorado has high school graduation requirements set by local school boards, but boards can select from a menu of options developed by the Colorado Department of Education, which includes collaboratively developed, standards-based performance assessments.

Implementation

Implementation of the collaboratively developed, standards-based performance assessments began with five schools across four districts. During the first pilot phase, educators identified which skills should be emphasized in building out performance-based assessments for graduation (Diaz-Billelo et al., 2021). In the second phase, the Colorado Department of Education sponsored a statewide professional learning community that focused on educators developing high-quality examples of performance assessments to assess students' demonstration of Colorado's Essential Skills. The tools developed under the pilot included a statewide scoring rubric. Colorado educators can now use an online platform that integrates a

—

assessments in their classrooms. New Hampshire allowed for a varying degree of district participation based on participation levels. Level 1 districts fully implemented PACE as part of their accountability systems by collaboratively developing assessments, working with external experts to ensure high-quality rubrics and scoring, then field-testing, implementing, and refining those assessments each year. Other levels participated in different components of PACE, such as professional development, but didn't have to commit to the accountability system. The shift from PACE to NH PLACE moved the emphasis from an accountability system to one that focuses on performance learning in the classroom, more professional development and collaboration among teachers and district and school leaders, and more student agency and choice in the classroom (which then emerges in the assessments).

Definition

Performance assessments are multi-step assignments with clear criteria, expectations, and processes which measure how well a student transfers and applies knowledge and complex skills to create or refine an original product and/or solution. This can include portfolios, exhibitions, student-led committees, or other performance tasks.

Relevant Subjects/Grades

PACE districts combined the statewide assessment system, common performance tasks, and other course-specific or local performance tasks. The statewide tests were administered in grade 3 English language arts (ELA); grade 4 math; grade 8 ELA and math; and grade 11 ELA, math, and science. The move to NH PLACE allows teachers in any grade-subject combination to participate in professional development and task design each year.

Scoring Process

Under PACE, common and local tasks were scored using teacher-developed rubrics, which describe student work and evidence at different competencies (Becker et al., 2017). The scoring process for common tasks involved teachers field-testing the tasks, revising them, and then scoring student work. The rubrics were then revised to ensure inaccuracies or vagaries were addressed. Scoring also involved a generalizability analysis by the Center for Assessment, who conducted cross-district comparability analyses that were critical for accountability purposes (Becker et al., 2017; Evans & Lyons, 2017).

Supports Provided

The New Hampshire Learning Initiative is the intermediary primarily responsible for the professional learning and facilitation of performance task development.

Relevant Subjects/Grades

Local performance assessments are administered in grades 3 through 8 and at least once in high school. Performance assessments are also an option for high school students to demonstrate proficiency in the Essential Skills. Students may opt to use a local assessment option or a Work Sample (see Definition, above, for distinction) in lieu of Oregon's Statewide Assessment or another standardized assessment (Oregon Department of Education, 2015). This applies to the following Essential Skills:

> Read and comprehend a variety of text
>
> Write clearly and accurately
>
> Apply mathematics in a variety of settings

Other Essential Skills may also be assessed using a local performance assessment, such as thinking critically, using technology, civic and community engagement, and global literacy (Oregon Department of Education, 2016).

Scoring Process

Work Samples used to assess proficiency in Essential Skills must use the state-developed scoring guide. Oregon Department of Education provides options for scoring responses to questions (e.g.,

completed. Under the current Proficiency-Based Graduation Requirements, students are awarded a diploma based on proficiency of coursework in six content areas, which may include performance-based assessments (such as a senior project, capstone, portfolio, or exhibition) (Rhode Island Department of Education, n.d.). In 2017, the Rhode Island Learning Champions developed cross-curricular performance indicators, scoring criteria, and student anchor work that could be shared with educators across the state (Rhode Island Department of Education, n.d.).

Relevant Policy/Legislation

The Council on Elementary and Secondary Education approved the Diploma System in 2003 and the Proficiency-Based Graduation Requirements in 2016. These policies converged under

Champions work included scoring criteria for districts to use when implementing performance assessments, though it is not clear how many districts are using the criteria and with what fidelity (Rhode Island Department of Education, n.d.)

Supports Provided

The Rhode Island Department of Education offers resources developed under the Rhode Island Learning Champions initiative. The state agency does not currently offer other professional learning to districts for developing, administering, or scoring performance assessments. Districts tend to collaborate with each other by sharing performance assessment resources and processes

Virginia: Tier 1

History

In 2014, the Virginia Assembly enacted legislation that eliminated five standardized assessments that were required by students in the state. Rather than administering the Standards of Learning (SOL) test, as was traditionally required in those subjects and grade levels, the state required that local school districts include local alternative assessments, which may include performance assessments (Virginia Department of Education, n.d.) In 2019, the Virginia Department of Education also encouraged each school district to develop Balanced Assessment Plans outlining the type of assessment used for all course content and subjects, with an eye to examining the breadth of assessments required for all students and identifying those that do not serve instructional purposes, are redundant, or might be replaced by new assessments that can more accurately measure content and skills (Virginia Department of Education, n.d.)

Relevant Policy/Legislation

Legislation adopted by the 2014 General Assembly amended § 22.1

Implementation

Virginia only requires local alternative assessments for primary grades and does not require performance assessments for high school graduation or for federal or state accountability purposes. However, the Virginia Department of Education allows districts to provide an option for a local performance assessment to verify high school English credits (one credit in reading and one credit in writing). The Virginia Department of Education requires that local school boards certify the instruction and assessments for the required grades and content areas. In 2019, the Department began encouraging districts to complete a Balanced Assessment Plan. Additionally, the Department conducts annual desk reviews of a sample of districts each year to provide accountability and technical assistance to districts. When legislation was originally enacted, the change from standardized testing to a local alternative assessment was immediate and statewide. Since then, the Department has provided more clarity on the guidelines and built more capacity by bringing in external groups to help districts with their Balanced Assessment Plans and provide districts with a review tool to understand how to measure the quality of their performance assessments.

Definition

According to the Virginia Department of Education Assessment Literacy Glossary, performance assessment "generally requires students to perform a task or create a product and is scored using a rubric or set of criteria. In completing the task, students apply acquired knowledge and skills. This type of assessment often includes a written component" (Virginia Department of Education, 2019)

Relevant Subjects/Grades

Local school districts are required to implement local alternative assessments, which may include performance assessments in grade 3 science, grade 3 history, grade 5 writing, U.S. history to 1865, and U.S. history from 35 wr 35506 -0.116774 l 0 -0.ent-2(t)1(o)-3(f 6 l W* n BT J ET C

provided on the Department's website. Virginia Department of Education does not provide a bank of performance assessments but encourages districts to share performance assessments with each other.

Massachusetts: Tier 2

History

The Massachusetts Consortium for Innovative Education Assessment (MCIEA) was formed in 2016 with six districts to build a new model of assessment and schoolwide accountability system that "offers a more dynamic picture of student learning and school quality than a single standardized test" (MCIEA, n.d.). MCIEA provides intermediary support to build the capacity of school district administrators, building-level administrators, and teachers to create high-quality, performance assessments that are embedded into the curriculum and generated by teachers in the classroom. Since its inception, MCIEA has grown from six to eight districts, with the governing board consisting of superintendents or their designees, as well as teacher union presidents, ensuring teacher involvement in decision-making. MCIEA is funded in part by the Commonwealth of Massachusetts and partners with the University of Massachusetts Amherst.

Relevant Policy/Legislation

In 1993, Massachusetts enacted the Education Reform Act, which resulted in the development and administration of the Massachusetts Comprehensive Assessment System (MCAS), a set of standardized tests in ELA, math, and science. MCIEA proposes to move away from standardized tests toward a more robust measure of student learning, such as performance assessments.

Implementation

MCIEA is not a state-driven effort but, rather, operates as a grassroots partnership with voluntary support for partner districts. MCIEA's goal is to work with schools, districts, and communities to "create a humanistic accountability system that is grounded in educational equity" (Beyond Test Scores Project, n.d.). Within partner districts, curriculum-embedded performance assessments are the primary measure of student learning. Districts undergo a year-long institute with coaching, cross-district support, and dedicated resources for the development and implementation of performance assessments. MCIEA also launched a Performance Assessment Task Bank, which allows consortium teachers to submit and access a range of performance tasks created by teachers within the consortium. Some districts have MCIEA district-led teams comprising instructional coaches or other instructional leaders to support and oversee the work.

Definition

Performance assessment consists of

> an extended task in which students have opportunities for sense-making and problem solving and/or original thinking in the context of a phenomenon or unresolved question;

> a method of capturing student work that is open-ended and generative, designed to represent not only a solution, but also the student thinking that underlies that solution; and

> evaluation criteria that describe how different aspects of student's work can be connected to substantive conclusions about what they know and can do.

Relevant Subjects/Grades

MCIEA advocates for performance assessments at all grade levels and subjects. Consortium interviewees noted that the lower grades are the most difficult to implement because most of the curriculum is skills-based.

Scoring Process

MCIEA works with districts during the year-long institutes to develop scoring processes that generate high inter-rater reliability and promote high-quality performance assesssments

Supports Provided

cross content areas (e.g., effective communication, creative and practical problem-solving) were created to support the implementation of both Act 77 and the Vermont Education Quality Standards. The Vermont Agency of Education has und

a natural component of project-based learning, allow learners to be assessed in engaging and authentic ways. In 2021, the Agencm 612.06 791.943226 l 612* n BT 20943226 l 612* n BT 20943242D5

Appendix B: Case Study State Interview Protocol

[Foundational question] Is there a common definition of performance assessments in your state?

1. Describe the history of performance-based assessments in your state, including the initial rationale for implementing performance assessment and how that may have changed over time.

   a. How long have performance assessments been used in your state? Have there been any interruptions?

   b. Describe relevant policy or legislation related to performance assessments in your state, districts, and networks.

   c. Can you describe what the early goals were for performance assessment, and how those goals have changed or evolved over time?

2. What role do performance assessments play in your state assessment system?

   a. (Probe the categories from Stosich and others [2018] minus federal accountability)

      i. Is it used for state (as opposed to federal) school or educator accountability? (This category includes replacing state tests with performance tasks and/or allowing performance assessments in nontested grades/subjects.)

      ii. For graduation?

      iii. For classroom purposes?

   b. What types of performance assessments or tasks are permitted or required?

   c. How are performance assessments developed? Assessed for quality?

   d. Describe how rubrics are created for assessing student work under performance assessment

      i. (Probe: state/local/educator)

      ii. How did you develop the rubrics? What guided that work or decision-making process?

e. To what extent are the needs of special student populations (e.g., ELs, SWDs) supported and addressed with respect to the use of performance assessments? If supported and addressed, what are the strategies or supports used? Describe.

3. Describe how performance assessments are aligned with curriculum, instruction, and learning standards.

   a. To what extent were there changes in learning standards and/or curriculum in order to ensure performance assessments were embedded? If changes were made, describe the process. Were learning standards, curriculum, and/or instruction changed first?

   b. To what extent are instruction and performance assessment mutually reinforcing?

4. What resources and supports are provided to districts and schools for implementing performance assessments and/or implementing new standards to support performance assessments?

   a. Describe how teachers and school administrators can receive professional learning or support regarding performance assessment or instructional approaches

   b. What do you see as the "key shifts" for teachers, school leaders, and district leaders in moving to instruction aligned to performance assessments, and what strategies best support them in making those shifts, including providing opportunities for collaboration?

   c. Were shifts needed in school operations or schedules in order to support performance assessment?

   d. Describe resources and supports provided to engage families regarding performance assessment (e.g., changes in score reporting, assessment literacy info).

5. What are three main lessons learned that you can share for other states embarking on this journey?

   a. What challenges has your state faced in implementing performance assessments?

   b. What would you have done differently?

   c. What successes do you see or have you seen? Is there evidence that implementing performance assessments is supporting your state's and network's goals?

Diaz-Bilello, E., & Pierre-Louis, M. (2021). *The Colorado performance-based assessment pilot: Background and context*. Center for Assessment, Design, Research and Evaluation, University of Colorado at Boulder. https://www.cde.state.co.us/postsecondary/pacasestudypaper1

Evans, C. M. (2019). Effects of New Hampshire's innovative assessment and accountability system on student achievement outcomes after three years. *Education Policy Analysis Archives*, *27*(10). https://doi.org/10.14507/epaa.27.4014

Evans, C. M., & Lyons, S. (2017). Comparability in balanced assessment systems for state accountability. *Educational Measurement: Issues and Practice*, *36*(3), 24–34. https://doi.org/10.1111/emp.12152

Fitzsimmons, P. (2020). *Strengthening local assessment systems for personalized, proficiency-based education: Strategies and tools for professional learning*. Aurora Institute. https://aurora-institute.org/wp-content/uploads/strengthening-local-assessment-systems-for-personalized-proficiency-based-education.pdf

High Sch m791.943226 l W* n BT .243137 .32549 .372549 rg /F0 12 Tf 72 456.663226 Td [(H)-2(i)-1(g