

2023 State Assessment Conference

Focus 1: Assessment Peer Review in the Context of Supporting Multiple Approaches to Student Assessment
Example Program Designs

The descriptions below are examples of program features being considered or implemented in assessments that fit under the “multiple approaches” umbrella. The descriptions are for illustrative purposes. Actual program features may vary and may change over time.

	Texas (TTAP)	New York Science Assessments	DLM Instructionally Embedded Model
General description of design	The TTAP system has three, short testing opportunities – one in the fall, winter, and spring. To ensure that all districts can retain their local curriculum, each TTAP progress monitoring opportunity will cover the full scope of the curriculum. Each opportunity will use a multi-stage adaptive design, allowing for shorter tests with greater accuracy to minimize the disruption to instructional time.	Written tests with performance tasks (accounting for 15% of score) Although not currently planned to be used for accountability purposes under ESSA, NY’s Performance-Based Learning and Assessment Networks (PLAN) Program is exploring the potential for New York’s educational assessment strategy to be reimaged in a way that purposefully fosters high-quality instructional opportunities, provides authentic measures of deeper learning, and better prepares students for college and the workplace.	Through-year model in ELA and mathematics with two test windows. Teachers choose what and when to assess, from a blueprint that has flexible options and constraints to ensure breadth of content coverage. Students cover the full blueprint in the fall and spring. DLM instructionally embedded assessments have been operational since 2014-2015. The system has met all peer review requirements.
Purpose	The Texas Through-year Assessment Pilot (TTAP) will explore whether Texas’s current summative assessment can be replaced with a cohesive progress monitoring system.	The science performance items allow students to demonstrate specific knowledge and skills articulated in the learning standards through hands-on laboratory experiences.	The purpose of the system is to improve academic experiences and outcomes for students with the most significant cognitive disabilities by setting high and actionable academic expectations and providing appropriate and effective supports to educators.

	Texas (TTAP)	New York Science Assessments	DLM Instructionally Embedded Model
Intended score uses	<p>Results from TTAP are intended to be used to support interpretations about what students know and are able to do on grade level content.</p> <p>Results provide information to allow educators to identify students who are in need of intervention and administrators to better support campuses and teachers.</p>	<p>The results of the NY Science Tests are intended to provide a measure of the extent to which individual students achieve the New York State Science Learning Standards for their course/grade level. In addition, the results are aggregated in order to determine whether schools, districts, and the State meet the required progress objectives specified in the New York State accountability system.</p>	<p>Results from the DLM alternate assessment are intended to support interpretations about what students know and are able to do and to support inferences about student achievement in the given subject.</p> <p>Results provide information that can guide instructional decisions as well as information for use with state accountability programs.</p>
What is assessed	<p>The same grade levels and subjects that were piloted during year 1 will be offered during year 2. Additional grade levels and subjects will be added to the pilot in following years.</p> <ul style="list-style-type: none"> ● Grade 5 science ● Grade 6 math ● Grade 7 math ● Grade 8 social studies 	<p>Science learning standards that are not able to be authentically addressed through tasks on a written test.</p>	<p>Extended academic content standards (Essential Elements, or EEs) linked to college and career readiness standards in English language arts and mathematics. The EEs specify grade-level learning targets, and the DLM learning maps clarify how students can reach those learning targets. Small collections of nodes are called linkage levels, which provide access to the EEs at different levels of complexity. Assessment items align to nodes.</p>
When students are assessed	<p>There are three opportunities each year. Opportunity one is given in the fall. Opportunity two is given in the winter. Opportunity three is given in the spring, currently about a month before our STAAR administrations.</p>	<p>Once at the end of a course/grade level.</p>	<p>There are two instructionally embedded testing windows: fall and spring. Each window is about 15 weeks long. Teachers choose when to test on each EE after instruction.</p>

	Texas (TTAP)	New York Science Assessments	DLM Instructionally Embedded Model
How students are assessed	TTAP is administered online via a multi-stage adaptive approach.	Performance tasks are administered shortly prior to the written test as a series of “stations” where students participate in a hand-on activity and record their observations/findings/conclusions	DLM assessments are administered in small groups of items called testlets. Some testlets are administered 1:1 by a teacher, others are computer-administered. Testlets are short, instructionally relevant measures of student skills and understandings. Students take a series of testlets (one per EE) to achieve blueprint coverage.
Item pools, test development implications	TTAP items are developed similarly to the STAAR test. TTAP uses the same content standards and item specifications as STAAR. However, TTAP items are maintained in a separate bank and are developed as part of a U.S.I Department of Education 2019 Competitive Grant for State Assessments awarded to the Texas Education Agency.	Tasks are developed once new learning standards are adopted. The written test is continually refreshed while the performance tasks are maintained from year-to-year with occasional revision or refresh. Test security is less of a concern due to the emphasis on performance (there are also variations within the tasks that prevent memorization).	The DLM Instructionally Embedded assessments has a testlet pool with a current (2023-24) total of 3,157 operational testlets, including 1,928 operational ELA testlets and 1,229 operational mathematics testlets. Updates are made to the operational pool each year to refresh and deepen the available pools of content.
Summative scoring and reporting	Still working on final format for summative scoring and reporting	Performance tasks make up 15% of the total test score, combined with a written test. Scale scores and performance levels are reported to parents and results (i.e., PLs) are aggregated for accountability	Summative scoring is based on the number of linkage levels mastered (derived from diagnostic classification modeling) across all assessed content standards. Performance levels describe overall achievement in each subject based on the total skills mastered, with students achieving At Target or Advanced demonstrating proficiency. The Learning Profile shows which Essential Elements and linkage levels the student mastered.

	Texas (TTAP)	New York Science Assessments	DLM Instructionally Embedded Model
Non-summative scoring and reporting	Still working on final format for non-summative scoring and reporting	None explicit, but often used for teacher improvement, program evaluation, and occasionally classroom grades.	Teachers can generate score reports on-demand during each assessment window. These reports provide similar information as the summative Learning Profile but only for the linkage levels tested during that window.
Other information	We currently do not have a date by which we plan to go operational with TTAP.		