

 **Innovations in Assessment**

**Texas Education Agency**



BILL & MELINDA  
GATES *foundation*

WALTON FAMILY  
FOUNDATION

 **educationfirst**



## *Introduction*

The vast disruptions caused by the Covid-19 pandemic prompted states and school networks to pause standardized tests and ask hard questions about the role and influence of traditional, once-a-year exams. This amounted to an opportunity to reimagine and redefine K-12 assessment and accountability alongside the students, families, teachers, and school leaders most familiar with the possibilities and pitfalls of our current approach.

*The Innovations in Assessment and New Measure Grant Program, founded in 2020 by the Bill and Melinda Gates Foundation, Walton Family Foundation, and Education First, was developed to support this work.*

**The grant program supports new assessments** that measure learning across multiple domains and capture student readiness for college or career, while continuing to provide the detailed data that an equitable approach to accountability requires. The grant program also supports a new strategy to create assessments: one that uses human-centered design principles, to more broadly share planning and decision-making power and build on the expertise of the people most proximate to the problem.

Five grantees are developing new measures to assess student learning across multiple domains. In contrast to traditional, top-down program design, these grantees are engaged with this work at the grassroots level, with a particular focus on how institutional, systemic, or ideological biases influence the experiences of some individuals and groups. These grantees are using **human-centered equitable design**, which builds on the expertise and experiences of students, families, and educators (informed by [Christine Ortiz's methodology](#)<sup>1</sup>).

The experiences of these grantees can inform assessments that more fully meet the needs of students, families, teachers, and school leaders. While no one test will perfectly assess all that we need to know about student learning, a richer menu of options and base of knowledge rooted in the direct experiences of the people most proximate to the problem can support the high-quality measures of learning that all students and educators deserve. Such measures can capture the breadth of student achievement, with explicit connections to the broad range of knowledge and skills that nurture thriving adults across diverse cultures and communities.

### Equitable Design in Action



<sup>1</sup>Ortiz Guzman, Christine M., *Just Design, The equityXdesign Book*.

## Testing As We Go: Using Through-Year Assessments to Measure and Propel Student Learning

On July 1, 2022, standardized test scores were not exactly a top priority for students and teachers in Texas. In Houston, the state’s biggest district, the school year had ended three-and-a-half weeks earlier, and the next school year was still seven long weeks away. Local news reports were focused on how an impending heat wave might affect July 4th celebrations, not trends in academic performance.

But in Texas, as in just about every state in the nation, statewide results on annual math, reading, and science tests typically arrive during summer vacation. Texas students had taken standardized tests that past May, schools had received student and schoolwide scores in June, and the statewide report was available on that hot Friday afternoon in July—when most teachers and students were long gone from the classroom.

This type of schedule made sense when standardized test results were based on physical scans of No. 2 pencil marks on bubble sheets. But the broad availability of technology, connectivity, and adaptive, computer-based assessments has opened an innovative possibility, one that Texas

education leaders are actively working to develop: through-year assessments that can provide score reports multiple times during the school year. While annual standardized testing programs provide detailed data and information, those insights only are available after the year has ended. What if students, teachers, families, and school leaders could track progress toward grade-level learning goals—and use that knowledge to focus instruction where it was needed most—while school was still in session?

Approximately 130 Texas districts and charter schools will have that chance during the 2022-23 school year, thanks to the Texas Through-

year Assessment Pilot. As with current annual assessments, this new through-year assessment is aligned to the Texas Essential Knowledge and Skills (TEKS), a detailed set of grade-level learning goals. But the pilot test, which students would take three times during a school year, is designed to play a different role in classrooms and schools.

Student performance on the through-year tests will aim to produce a final score based on whether students showed that they achieved proficiency for grade-level material, just like traditional, once-a-year assessments. But the ongoing testing program also will give students a chance to demonstrate their learning



several times before the year has ended, as well as provide insights about student progress that teachers and families can use during the learning cycle.

Districts currently use a diverse set of locally selected tools to measure growth throughout the year, but those are not always being utilized within a balanced assessment system and take up a portion of the district’s budget. It also can serve as a standard, high-quality, interim measure of student progress across the state that is aligned to TEKS standards and statewide annual assessments.

The through-year assessment also can serve as a standard, high-quality,

.....

***What if students, teachers, families, and school leaders could track progress toward grade-level learning goals—and use that knowledge to focus instruction where it was needed most—while school was still in session?***

.....

interim measure of student progress across the state that is aligned to TEKS standards and statewide annual assessments. Districts currently use a diverse set of locally selected tools to measure growth throughout the year, but those are not always utilized within a balanced assessment system and take up a portion of district budget.

The pilot comes on the heels of—but is separate from—a comprehensive effort to redesign the State of Texas Assessment of Academic Readiness (STAAR®) program to better align with classroom instruction. The **STAAR Redesign** includes incorporating writing into all reading and language arts assessments, prioritizing cross-curricular content, adding new non-multiple-choice questions that are more like questions teachers ask in class, and moving to online assessments that provide a full suite of robust accommodations for students with specific learning needs. The pilot and these innovations are intended to ensure that statewide interim and summative testing not only measure, but also enhance student learning. This is a critical goal for Texas, which educates more than 10 percent of the nation’s K-12 students, including large numbers of students who are English-language learners or from families experiencing poverty.





## Key Facts

### Who:

Texas Education Agency

---

### What:

Developing a new through-year assessment pilot as a potential replacement for the current summative testing program. This pilot would test the feasibility of a model that would:

- Establish multiple opportunities for students to demonstrate their growth and learning throughout the school year, including students who move between districts.
  - Monitor student progress to show teachers where students have and have not yet achieved proficiency before the school year is over and provide actionable information that can immediately inform classroom instruction.
- 

### Why:

In 2019, the Texas Legislature unanimously passed House Bill 3906, which mandated a host of changes to annual standardized tests. The law required tests to include writing in all grade levels, cross-curricular content, and new non-multiple-choice questions, as well moving to online assessments that provide robust accommodations for students with specific learning needs. It also directed the state education agency to create and pilot an integrated through-year assessment that would monitor student progress, inform classroom instruction, and potentially replace once-a-year summative assessments.

---

### How:

Before getting involved with the Innovations in Assessment program, the Texas Education Agency developed two potential prototypes of the through-year assessment. The team then solicited input from focus groups, including by surveying district leaders and testing coordinators, chief academic officers, and teachers. After joining the program, the team focused on understanding the needs of the people most proximate to the problem and convened 15 focus groups of teachers, parents, and students to identify their priorities for new through-year tests. Those insights inspired several changes to the prototype that will be piloted in 2022-23, such as how to make score reports more informative and how to fairly measure growth among transient students who move from one district to another. The pilot will iterate over several years. The agency and project team will work to improve the testing experience and analyze data to ensure the assessment maintains the validity and reliability of the current annual tests, while also meeting the needs of students, teachers, families, and school and district leaders for more frequent and timely information.

## Rethinking Summative Testing

The path to the Texas Through-year Assessment Pilot included research institutions, state commissions, and the state legislature. A pair of research studies found the current test was appropriately rigorous, valid, and aligned to the TEKS learning standards. In addition, the [Texas Commission on Next Generation Assessments and Accountability](#) recommended implementing a computer-adaptive system of multiple integrated tests administered throughout the school year. Then came the 2019 legislation, House Bill 3906, which aimed to improve the summative

.....

***Through these programs, the state education agency aims to create a cohesive assessment system that allows students to track their own learning and feel more ownership of the growth process throughout the year.***

.....

STAAR® assessment program to better support high-quality classroom instruction. In addition to a redesigned STAAR®, the Texas Education Agency spurred the development of integrated formative assessments. The goal was to create and pilot a formative

assessment to inform teaching decisions and improve instructional supports, as well as potentially replace the current summative assessment. The [Texas Through-year Assessment Pilot](#) and [Texas Formative Assessment Resource](#) were established as a result.

Through these programs, the state education agency aims to create a cohesive assessment system that allows students to track their own learning and feel more ownership of the growth process throughout the year. The program also provides teachers with a TEKS-aligned tool to identify students who may need intervention during the course of the school year, with minimal interruption of instructional time.



## Co-Designing With Communities

### *Starting with Stakeholders, not Test Designers*

The team began its work with surveys and conversations with the people most proximate to the problem: teachers, school and district leaders, families, and students. The goal was to better understand their concerns and what they valued. First came a “listening session” with the state-commissioned Educator Advisory Committee of teachers and administrators. The project team also surveyed superintendents, met with assessment experts from across the country, and convened 15 focus groups with members from urban, suburban, and rural



.....

***Early conversations with people proximate to the problem established core features of the through-year assessment.”***

.....

communities, including members of TeachPlus Texas and the Texas PTA. The team then shifted away from a committee-based approach to

gathering feedback and conducted in-depth interviews with individual students and teachers, who were selected informally through personal or peer network connections. Despite limitations on outreach due to the ongoing Covid-19 pandemic, the team interviewed about 40 teachers, 50 parents, and 25 students, to capture a broad range of perspectives and priorities.

The priorities of these stakeholders were clear: parents and teachers wanted a way to track student progress throughout the school year that did not increase testing time, and they wanted to get test results quickly. Teachers and district leaders made clear that through-year tests could not lock schools

into a particular scope or sequence of instruction. Allowing districts to retain local control over curriculum was critical. In addition, the tests could not penalize students who move schools or districts during the year or English-language learners, whose proficiency increases throughout the year.

### ***A Level Playing Field and Local Control***

Those early conversations established core features of the through-year assessment. The test would be given no more than three times over the school year. It would be computer-adaptive, to quickly identify a student’s level of proficiency through a smaller

number of questions that could be completed in a single class period. And it would be “full scope,” or cover the material for the entire year, rather than rely on an assumed sequence of instruction and test specific material at different times in the school year. A “full scope” assessment was important to accommodate differences in pacing and sequence among districts in a large and diverse state.

This model also makes it possible to measure student growth over the course of a school year, which stakeholders identified as another major priority. Rather than assessing whether a student did or did not meet a static benchmark, a growth model can show gains over time. Measuring growth allows for acknowledgement of a student’s progress even if overall proficiency is not yet where it needs to be. Every

.....

***A “full scope” assessment was important to accommodate differences in pacing and sequence among districts in a large and diverse state.***

.....

student grows in different ways, and growth measures enable teachers to celebrate progress with their students.

***Teachers Influence Technical Details***

Crafting tests that are relevant, actionable, reliable, and can be used to track and predict student performance is technically complex. The project team has nonetheless prioritized the feedback of the Educator Advisory Committee throughout the testing and refinement process leading up to the 2022-23 pilot and called on test

designers to creatively meet educators’ needs. This process has influenced the pilot in several ways. For example, score reports will eventually include classroom-level predictive analytics, to gauge a student’s likely future performance on the STAAR®. Test item statistics will be released annually. And student growth measures will, over time, be expressed in categorical terms rather than in percentiles.

Most notably, Educator Advisory Committee feedback has influenced how tests will be scored. The project team explored several possibilities, such as using a student’s single highest score, a weighted average across all three tests, or the score on the final test. Based on educator feedback, the pilot will use a “help not hurt” weighted score model, which bases a student’s final score on either the final test or, if the student earns a higher score on one of the earlier test days, an adjusted average of all three test scores.



## Looking Ahead

As the pilot begins, the project team will be focused on ensuring score reports are both reliable and relevant. Educators have clearly identified offering a growth measure as a key priority for the through-year assessment. But many growth assessments have low measures of reliability. The Texas team is starting with a simple gain score for each student in the first year of the pilot, to be followed by more sophisticated measures of growth in later years.

Relevance is also a complex goal. The project team’s theory of action holds that through-year assessments that

are integrated into the school year and offer timely results can help

.....

***As the pilot begins, the project team will be focused on ensuring score reports are both reliable and relevant.***

.....

identify gaps in student learning while teachers and families still have time to help students address them.

To unlock that potential, student score reports must be straightforward to understand, clearly aligned to academic standards, and easily mapped to diverse approaches to curriculum and instruction. In the pilot stage and beyond, co-designing with stakeholders will continue.

In its first year, the pilot will serve not as an endpoint but rather a launch point, making way for the program to evolve in future years. In the pilot and beyond, co-designing with stakeholders will continue.

