Texas Through-Year Assessment Pilot (TTAP)
Year 1 Pilot Report

2022–2023 School Year

Published August 1, 2024

Executive Summary

- House Bill (HB) 3906, 86th Texas Legislature, 2019, required the Texas Education Agency (TEA) to develop and pilot an innovative, through-year assessment model as a possible replacement for the State of Texas Assessment of Academic Readiness (STAAR®) summative tests. A through-year assessment model refers to a progress monitoring system that provides students multiple opportunities throughout the school year to demonstrate their mastery of standards and to contribute to their summative performance level reported at the end of the year.
- The Texas Through-year Assessment Pilot (TTAP) was designed in collaboration with Texas educators, administrators, students, and families. Hundreds of individuals were involved in the design process through focus groups, surveys, and advisory committees. Based on feedback, the initial pilot design features three, short testing opportunities—in fall, winter, and spring. To ensure that all local education agencies (LEAs) can retain their local curriculum and provide students a growth measure, each TTAP progress monitoring opportunity covers the full breadth of the curriculum. In other words, each TTAP testing opportunity is a proportionally shortened version of the STAAR summative assessment. TTAP uses a multi-stage adaptive design, allowing for shorter tests with greater accuracy to minimize the disruption to instructional time. For more details on the test design, see Figure 2.
- The pilot first launched during the 2022–2023 school year after delays due to COVID-19. Four test titles were offered: grade 6 math, grade 7 math, grade 5 science, and grade 8 social studies. Approximately 65,000 students participated in the pilot across 121 LEAs.
- Initial pilot feedback from year 1 participants and results from psychometric and statistical analyses are optimistic. However, at least 2–3 more years of data is needed to be collected for TEA to determine the technical feasibility of the pilot prior to the legislature passing and funding a bill that would require full implementation. In future years, TEA will need to:
    1. expand test titles to include grade 3 and reading language arts (RLA) to evaluate operational and assessment policy implications across all grades and content areas.
    2. continue exploring the feasibility of using a cumulative scoring model, which uses scores from all three testing opportunities to inform final scores. The current analyses remain inconclusive regarding the integration of scores from all three testing opportunities.
- The pilot revealed an additional path to implement through-year assessments. This alternative option includes a new, shorter summative assessment paired with an updated interim assessment model. Agency-provided trainings, support, and guidance to administer both products would reinforce local control of balanced assessment system components. This option allows for implementation as early as spring 2026, approximately three years faster than a fully developed through-year assessment model.

For more technical details on the pilot, see the 2022–23 TTAP Technical Report.

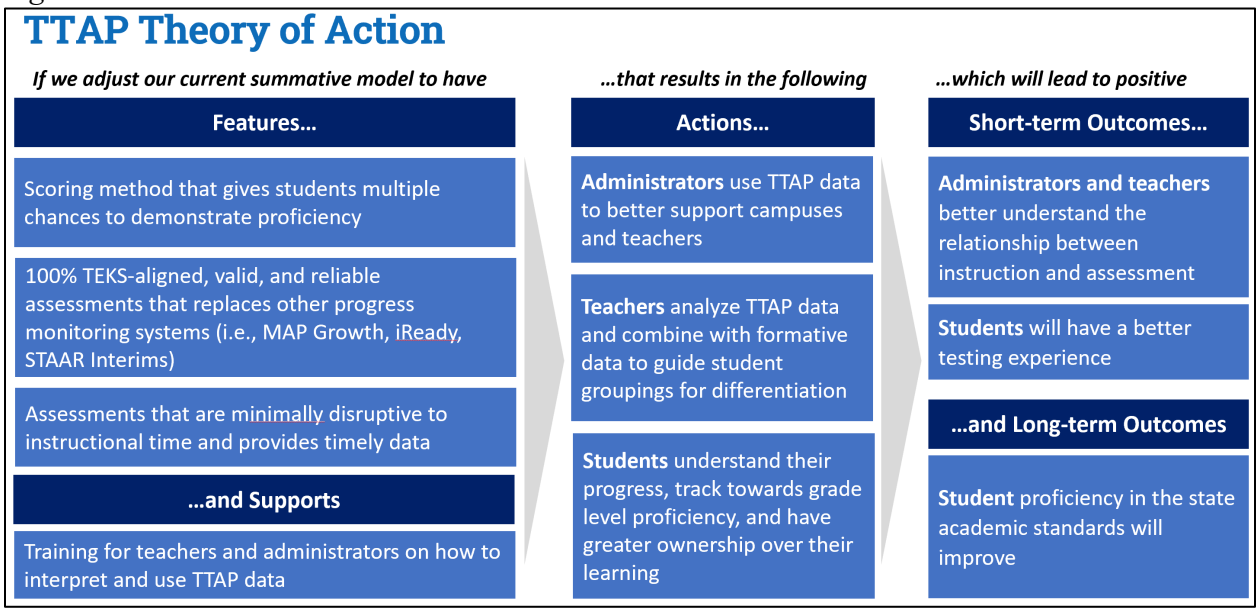
Year 1 TTAP Update

Pilot Introduction and Context

HB 3906, passed in 2019, required TEA to develop an integrated formative assessment pilot. The purpose of the pilot is to assess the feasibility of a through-year assessment model applied to assessments required under Texas Education Code (TEC) Section 28.006 or 39.023. The pilot is optional for LEAs, and it does not affect the participating LEA's obligation to administer STAAR. The establishment of the pilot in statute was subsequent to a formal report of recommendations made by the Texas Commission on Next Generation Assessments and Accountability[1] in 2016. The report recommended a system of multiple through-year assessments, stating, "to ensure the individualized Texas specific computer-adaptive assessment system provides useful, real-time feedback to educators, parents, and students, multiple, shorter assessments—as opposed to lengthy one-time assessments—could be used to inform individual student learning and growth."

In response to HB 3906, the Texas Through-Year Assessment Pilot (TTAP) was introduced as an innovative, through-year assessment model launched in fall 2022. Positioned as a potential alternative to the STAAR summative assessments, TTAP operates as a progress monitoring system, offering students multiple shorter opportunities throughout the academic year to demonstrate their mastery of the curriculum standards. TTAP also contributes to the prediction of a student's summative performance level reported at the end of the school year.

Figure 1.



The through-year assessment model pilot will continue to undergo evaluation over several years to assess its benefits while ensuring that its design maintains the rigorous level of validity and reliability that STAAR currently meets. The pilot seeks to establish a scoring methodology that is comparable in performance classification to STAAR (i.e., suitable for both progress

[1] https://tea.texas.gov/system/files/TCNGAA-Report_Final_2016-08-30.pdf

monitoring purposes as well as state accountability). Stakeholders have also articulated the hope that the new assessment system takes into account student performance from all three testing opportunities.

Pilot Design and Stakeholder Involvement

TTAP was initiated with the goal of creating an assessment system that could enhance the student testing experience and provide stakeholders with more real-time information to provide appropriate student supports and thereby improving student academic outcomes. The initial TTAP design was developed through collaboration with Texas educators, administrators, students, and families. This included surveying over 1,500 superintendents and testing coordinators, multi-year convenings with an educator advisory committee, and conducting numerous focus groups with over 250 teachers, parent organization groups, and students. TEA prioritized a wide distribution of voices in the engagement process, considering urbanicity of campuses, demographics, and geography within Texas. Regular consultations with a technical advisory committee at crucial milestones, in partnership with assessment vendors, further ensured alignment with stakeholder values and technical feasibility, resulting in a model taken up by participating LEAs (see Figure 2 and Figure 3).

Figure 2.

Design Question	Considerations	TTAP Design Choice
Curricular Alignment—*to what degree should each assessment be aligned to state curriculum standards?*	• **Full scope assessments** allow LEAs to maintain local scope and sequence and provide within-year growth measures, but students will be tested on content that they have not yet learned in the fall and winter. • **Curricular-aligned assessments** test students only on the portion of the curriculum that has been taught by that point in the school year but require LEAs to adopt a statewide scope and sequence and makes it less likely that student performance from earlier assessments can be included in a student's final score.	TTAP assessments are **full scope** to avoid requiring all LEAs to adopt a specific curricular sequence within the school year and to allow for within-year growth measures.
Computer Adaptivity—*how individualized should assessments be?*	• **Static, or linear, designs** are easiest to understand and require the fewest items, which are currently released on an annual basis. However, it is the longest test design, and the items are not individualized based on the student's demonstrated ability. • **Multi-stage computer adaptive designs** maximize efficiency and transparency for students at all levels but require a larger	TTAP assessments use a **multi-stage adaptive model.** This model permits the release of items but allows for a shorter test than the static design and requires less field-testing and less upfront cost than the

Figure 2.

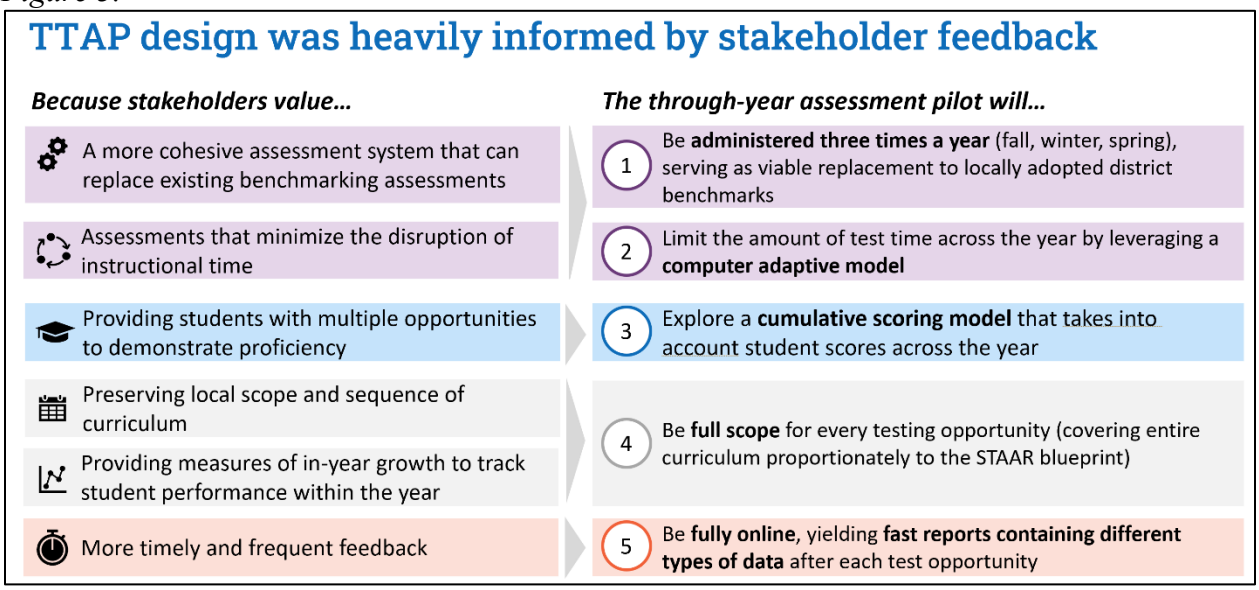
	annual production of items and are cost-intensive to accommodate students with special needs. • **Item-level computer adaptive designs** are the most individualized at the student level and provide greater measurement precision for the lowest and highest student abilities (which could include above and below grade items). But this model is the least transparent as only a sample of items can be released annually and is the costliest to start up due to its need for the largest item bank. This design would not accommodate students with certain special needs and for RLA, adaptivity is somewhat limited as RLA items are bundled around passages.	item-level adaptive design. TTAP studies indicate item-level tests are minimally beneficial in terms of psychometric measurement and test length compared to the multi-stage adaptive model.
Test Items—*what types of items should Opportunity 1 and 2 assessments include?*	• **Assessments that use only multiple-choice items** allow for immediate scoring and reporting but do not use items as aligned to the questions teachers ask in classrooms. • **Assessments that include all non-multiple-choice items** (except constructed-response items) also allow for immediate scoring and reporting and include more items that align with classroom practices. • **Assessments that include all non-multiple-choice items** (including constructed-response items) are most aligned to the types of questions that teachers ask in classrooms but would increase the burden on LEAs due to additional field-test requirements and reduced instructional time. Results will be less immediate after each testing opportunity.	TTAP Opportunity 1 and 2 **assessments have non-multiple-choice items but exclude constructed-response items**. This design allows the assessment to align with classroom practices and provides timely scoring and reporting, which teachers and students need in the fall and winter.
Test Length—*what is an appropriate length for assessments given the information*	• **Tests that fit within a class period** would work well in middle and high schools and preserve more instructional time. But this design makes it difficult to include constructed-response items and puts the reliability of assessments at risk.	TTAP assessments are currently designed to be **completed in one sitting but could take longer than a class period** (depending on the

Figure 2.

stakeholders want?	Additionally, it would not be possible to cover the full scope of the curriculum standards within the test, thus prohibiting a growth measure and the potential for creating a cumulative student score. Less information for stakeholders is available with a shorter test. • **Tests that take longer than a class period** would end up being more reliable and allow for a full scope test, which enables local curricular autonomy, growth measures, and the potential to include performance on earlier assessments in a student's final score. It also allows for more non-multiple-choice item types.	campus). In addition to the considerations listed, time to set up an administration takes an extra 15–20 minutes on top of the actual time to test. Regardless of the test length, typical test security measures (e.g., use of hall monitors, removal of instructional material displays that may aid or are a direct source of answers) must be implemented to maintain the security of the test content.

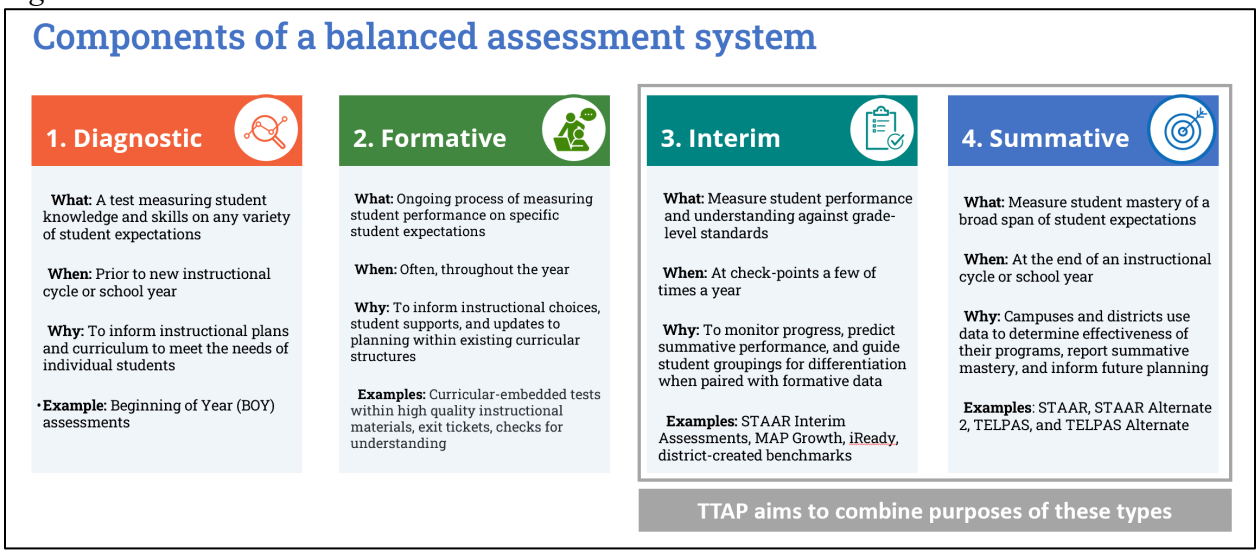
TTAP was developed to adhere to the same rigorous standards of STAAR, with every question undergoing Texas educator reviews as well as field testing. To ensure that all participating LEAs are able to maintain their local curriculum, each TTAP progress monitoring opportunity covers the full scope of the curriculum. As a result, the test blueprints for TTAP represent proportionally shortened versions of the STAAR summative assessments. Each testing opportunity uses a multi-stage adaptive design, enabling shorter tests with enhanced accuracy to minimize disruptions to instructional time. Each testing opportunity involves two stages with routing cut scores that determine student progression from Stage 1 to Stage 2 based on performance. Performance in the previous testing opportunity informs the starting point for the student in the next testing opportunity.

Figure 3.



TTAP aims to fill the roles of both interim and summative assessments. It is important to note that TTAP does not fulfill the purpose of locally administered formative assessments in which curriculum standard-specific data help indicate student misconceptions and inform specific instructional choices. Rather, the first two testing opportunities serve the purpose of an interim assessment in which teachers are informed of student proficiency and progress toward end-of-year learning goals (see Figure 4). By combining data from both locally administered formative assessments and TTAP, educators should be able to gain a full picture of a student's learning progress and its relation to end-of-year expectations. Participants are required to forgo other types of benchmark testing while administering TTAP to preserve valuable instructional time during the school year.

Figure 4.



Pilot Year 1 Logistics and Execution

In the first year of the pilot (2022–2023 school year), 121 LEAs participated (see appendix). The progress monitoring system incorporated three distinct, short testing opportunities held during the fall (Opportunity 1), winter (Opportunity 2), and spring (Opportunity 3).

In the first year of TTAP, TEA offered four test titles—grade 6 math, grade 7 math, grade 5 science, and grade 8 social studies—administered online. The grade 5 science test included a Spanish version for eligible students. Limitations of the first-year pilot included content and language supports, paper, braille, and American sign language versions of the assessments not being available. LEAs were advised to administer locally determined interim assessment alternatives to students who require special versions of assessments. Opportunities 1 and 2 were designed to be completed in one sitting with a testing time (outside of work required to set up) of approximately 40–75 minutes. Opportunity 3 was projected to take 80–120 minutes to complete. When scheduling, testing personnel also needed to consider the time needed to set up the testing environment and provide test instructions. In total, there were 64,738 students who were administered at least one TTAP assessment, and 53,984 students who partook in all three testing opportunities. The student sample collected was largely representative of the state population and was sufficient for data analyses purposes.

Pilot Year 1 Results and Feedback

Several special studies were conducted in year one of the TTAP pilot contributing to the ongoing refinements. A summary and implications of each study are as follows:

- In comparing the psychometric properties of TTAP Opportunity 3 and STAAR, the study showed positive preliminary evidence that the two assessments provide comparable interpretations of student ability if administered within the same testing window. TTAP Opportunity 3 achieved a similar level of reliability and classification accuracy as STAAR.
- The study on TTAP cumulative scoring methods reviewed eight different approaches to producing a summative score using actual and simulated TTAP scores from year 1. These approaches included Opportunity 3 only, maximum score, composite score, and a help-but-not-hurt (HNH) model. The Opportunity 3 only method was the most straightforward approach to generating a valid cumulative score as it only required Opportunity 3. This cumulative scoring approach was similar to the path that a few other states (e.g., Nebraska, Georgia) implementing a through-year model have taken where the third opportunity's starting point is informed by the second testing opportunity's performance. However, a variety of stakeholders have indicated interest in the HNH model as it incentivizes a student to demonstrate strong Opportunity 1 and 2 performance but does not negatively impact the student's cumulative score if he or she demonstrates a different growth pattern over the course of the school year. TEA will continue to explore the feasibility of the HNH model by developing a robust research plan for future years of the pilot to meet the continually evolving needs of this unique model in the face of federal assessment and accountability requirements. To implement a HNH model in the future, it is likely that tests for both Opportunity 1 and 2 will need to be longer for greater consistency with the Opportunity 3 tests.

- An initial study on item-level computer adaptive tests (CAT) showed that there were only small improvements to be gained when adopting this model compared to a multi-stage adaptive model. This initial study used both the existing TTAP item pool as well as a more robust 600–700 item pool. The improvements in measurement precision and ability recovery increased near the tails of the ability distribution for the lowest and highest scoring students. However, most differences were small to negligible, especially for the majority of students who are near the middle of the ability distribution. Because TEA annually releases the spring STAAR assessments, a CAT model would require substantial start-up costs to build a usable item bank and pose a significant operational burden compared to the current multi-stage adaptive model.

TEA collected feedback from year 1 participants, including teachers, administrators, and students, throughout the school year. This collection was done through surveys for teachers and administrators, optional student surveys after each test administration, an advisory committee that meets three times a year, and LEA site visits. Initial feedback from educators showed satisfaction with operations and execution of the pilot and some favorability towards the TTAP model as a summative replacement. For example, in the year-end survey, 80 percent of educators agreed that the communication, information, and training provided were effective and helpful in supporting a successful administration. A majority, 60 percent, of educators also believed that the TTAP model is a better alternative to STAAR. Students also shared a positive-neutral experience as 75 percent of students surveyed at the end of the year preferred TTAP to STAAR, and 56 percent of students agreed that the TTAP model allowed them to better show what they have learned and know. While feedback from participants was mostly positive, participants also raised questions around the operational and policy implications of adopting a three-times-a-year replacement to the STAAR. Given the nature of a pilot, some sentiments cannot be fully captured until if and when this design replaces the current summative. There will be additional operational lifts for LEAs that are not captured at this point in time, such as increased number of test administrations and security requirements.

TEA received constructive feedback from districts on the through-year summative model. Feedback received centered on the logistics of administering multiple high-security tests, potential impacts on student instructional time, and the need to educate students about a new, through-year testing model. Additionally, educators had a diverse array of feedback on TTAP score reports and the utility they serve in understanding and monitoring the progress of students throughout the year, as Opportunity 1 and 2 tests were meant to serve as interim checkpoints for teachers rather than a standalone formative tool for determining instructional next steps for students. A majority of stakeholders (i.e., teachers, district and campus administrators, district and campus testing coordinators) believed that TTAP has the potential to replace their LEA's long-term need for other interim assessment products (e.g., MAP Growth, STAAR Interim Assessments, district-created benchmarks).

Additional Considerations Beyond the 2022–23 School Year

Since the 2022–23 pilot launched, the agency has been clarifying the scope of the pilot to determine what is the best fit for various student groups. A few things to note include:

- **Limiting the TTAP model to exclude STAAR Alternate 2:** Students with the most significant cognitive disabilities cannot be accurately measured in a through-year format. Student ability and demonstrated performance on STAAR Alternative 2 do not vary sufficiently to support the model. STAAR Alternate 2 is administered on paper in a one-on-one setting and is a teacher-led assessment.
- **Limiting the TTAP model to assessments for grades 3–8:** Assessments for high school courses, which can be completed in a single semester, and re-testers overcomplicate the testing process and could overburden students needing to graduate.
- **Potentially limiting the TTAP model to certain content areas:** The agency is working to measure and observe the utility of a TTAP model for science and social studies. Factors include the structure of the curriculum standards, potential for longitudinal data, potential for progression throughout the year, and overall testing burden.

Moreover, a new finding from the pilot revealed an additional path to implement through-year assessments. As noted in the Pilot Year 1 Results and Feedback section, TTAP's Opportunity 3 has shown to be just as reliable as the STAAR summative assessment, despite being approximately 15 percent shorter than the current summative assessment. The alternative option includes a new, shorter summative assessment paired with an updated interim assessment model.

- The **new, shorter summative assessment** would include:
    - a modified blueprint that reduces test length by 11–30 percent,
    - approximately 30 minutes reduction in test time for each test,
    - the elimination of reporting categories,
    - a concordance study to bridge between the difficulty of the new test and the current STAAR to ensure parity, and
    - continued adherence to STAAR redesign principles (e.g., inclusion of extended constructed responses at every grade level in RLA)
- The **updated interim assessment model** would include:
    - improvements to the interim assessment design and
    - training and guidance for the interim assessment model so they may serve a through-year, progress monitoring purpose.
- **Agency-provided trainings and guidance** to support both products would reinforce local control of balanced assessment system components. Based on the changes listed above, this option allows for implementation as early as spring 2026, approximately three years faster than TTAP.

Conclusion and Next Steps

The initial year of TTAP shows promise, but substantiating its viability as a through-year summative replacement requires more years of data collection and some additional modifications based on feedback. Data from the first year indicates that the multi-stage adaptive design coupled with a proportionally reduced test blueprint effectively optimizes reliability and classification accuracy, rivaling that of STAAR. While participants expressed optimism, both educators and students grappled with the operational and instructional changes posed by a three times a year full-scope test program, along with uncertainties regarding the utility of score reports throughout the school year. These findings indicate TTAP's potential as an alternative to STAAR, while

emphasizing the ongoing need for refinement and studies to establish a suitable scoring methodology for state accountability. In its pilot phase, the through-year assessment serves as an optional, supplementary tool for LEAs to enhance their understanding of student proficiency throughout the school year.

TEA intends to further explore TTAP's potential over the next 3–-4 years seeking additional evidence on its viability as a through-year summative replacement. A comprehensive evaluation across grades and content areas assessed is essential in validating TTAP as a reliable replacement for STAAR. TEA needs to (see Figure 5):

- expand the pilot to encompass at least one grade 3 test title and
- incorporate at least one RLA test title.

Figure 5.

	SY22-23	SY23-24	SY24-25	SY25-26
Phase 1: Initial pilot titles	*Grade 5 Science, Grade 6 Math, Grade 7 Math, Grade 8 Social Studies*			
Phase 2: Add grade 3 titles			*+ Grade 3 Math, Grade 8 Math*	
Phase 3: Add RLA titles				*+ RLA (TBD)*

Additionally, TEA will address psychometric and logistical challenges associated with administering a high-security state assessment program (similar to STAAR) three times annually. With accumulated data spanning multiple years, TEA aims to:

- assess the stability of findings longitudinally and
- explore the feasibility of generating a cumulative score that aligns with the pilot's objectives.

Lastly, the alternative option to implementing through-year assessments through new, shorter summative assessments and pairing it with an updated interim assessment model is a new consideration for the legislature.

TEA will produce an addendum summarizing the 2023–2024 school year and an additional legislative report after the 2025–2026 school year. For detailed information about TTAP, refer to the TTAP webpage at https://tea.texas.gov/student-assessment/assessment-initiatives/texas-through-year-assessment-pilot.

Appendix

2022–2023 TTAP Participating LEAs

ACADEMY OF DALLAS – 057810
ALICE ISD – 125901
ALVORD ISD – 249901
AMARILLO ISD – 188901
ANAHUAC ISD – 036901
ARCHER CITY ISD – 005901
ATHENS ISD – 107901
AUBREY ISD – 061907
BANQUETE ISD – 178913
BELLEVUE ISD – 039904
BETTY M CONDRA SCHOOL FOR EDUCATION INNOVATION – 152806
BEXAR COUNTY ACADEMY – 015809
BLOOMINGTON ISD – 235901
BOERNE ISD – 130901
BORGER ISD – 117901
BRIDGEWAY PREPARATORY ACADEMY – 057851
BRONTE ISD – 041901
BURKBURNETT ISD – 243901
BURLESON ISD – 126902
BURNHAM WOOD CHARTER SCHOOL DISTRICT – 071801
CALLISBURG ISD – 049905
CARRIZO SPRINGS CISD – 064903
CENTRAL ISD – 003907
CISCO ISD – 067902
COMFORT ISD – 130902
CORRIGAN-CAMDEN ISD – 187904
COTULLA ISD – 142901
CROCKETT COUNTY CONSOLIDATED CSD – 053001
CROSS ROADS ISD – 107904
CROWLEY ISD – 220912
DAMON ISD – 020910
DAYTON ISD – 146902
DE LEON ISD – 047902
DELL CITY ISD – 115903
DEW ISD – 081906
DORAL ACADEMY OF TEXAS – 105804
DR M L GARZA-GONZALEZ CHARTER SCHOOL – 178801
DUMAS ISD – 171901
DUNCANVILLE ISD – 057907

EAST BERNARD ISD – 241902
EL PASO LEADERSHIP ACADEMY – 071810
FALLS CITY ISD – 128904
FLORESVILLE ISD – 247901
FORESTBURG ISD – 169910
FRUITVALE ISD – 234909
GODLEY ISD – 126911
GOLD BURG ISD – 169906
GUNTER ISD – 091917
HAMILTON ISD – 097902
HENRIETTA ISD – 039902
HUFFMAN ISD – 101925
INTERNATIONAL LEADERSHIP OF TEXAS (ILTEXAS) – 057848
JACKSBORO ISD – 119902
JIM NED CISD – 221911
JOURDANTON ISD – 007902
KENEDY ISD – 128902
LAMAR CISD – 079901
LEXINGTON ISD – 144902
LIVINGSTON ISD – 187907
LYFORD CISD – 245902
MCDADE ISD – 011905
MCGREGOR ISD – 161909
MIDWAY ISD – 161903
MILES ISD – 200902
MULLIN ISD – 167902
MUNDAY CISD – 138903
NEWTON ISD – 176902
NORTH HOPKINS ISD – 112906
NORTHSIDE ISD – 244905
O'DONNELL ISD – 153903
ODYSSEY ACADEMY INC – 084802
PFLUGERVILLE ISD – 227904
PIONEER TECHNOLOGY & ARTS ACADEMY – 057850
PLEASANT GROVE ISD – 019912
RALLS ISD – 054903
RED OAK ISD – 070911
REVE PREPARATORY CHARTER SCHOOL – 101876
RIO HONDO ISD – 031911
RIVER ROAD ISD – 188902
ROBINSON ISD – 161922
ROGERS ISD – 014907
ROOSEVELT ISD – 152908

ROUND ROCK ISD – 246909
S AND S CISD – 091914
SABINE PASS ISD – 123913
SAN ANTONIO PREPARATORY SCHOOLS – 015840
SAN BENITO CISD – 031912
SAN FELIPE-DEL RIO CISD – 233901
SAN MARCOS CISD – 105902
SAN PERLITA ISD – 245904
SHELBYVILLE ISD – 210903
SMITHVILLE ISD – 011904
SPEARMAN ISD – 098904
SPRING ISD – 101919
STAFFORD MSD – 079910
STERLING CITY ISD – 216901
STRAWN ISD – 182905
TARKINGTON ISD – 146907
TEXARKANA ISD – 019907
TEXAS CITY ISD – 084906
TOLAR ISD – 111903
TORNILLO ISD – 071908
TRINITY ISD – 228903
TULOSO-MIDWAY ISD – 178912
TYLER ISD – 212905
UNION GROVE ISD – 230908
VALERE PUBLIC SCHOOLS – 227824
VALLEY VIEW ISD – 049903
VAN VLECK ISD – 158906
VIDOR ISD – 181907
VISTA DEL FUTURO CHARTER SCHOOL – 071809
WAELDER ISD – 089905
WAXAHACHIE FAITH FAMILY ACADEMY – 070801
WEBB CISD – 240904
WELLINGTON ISD – 044902
WELLMAN-UNION CISD – 223904
WESTWOOD ISD – 001908
WHARTON ISD – 241904
WINDTHORST ISD – 005904
WORTHAM ISD – 081905
YSLETA ISD – 071905