Executive Summary

- House Bill (HB)

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

Figure 2.

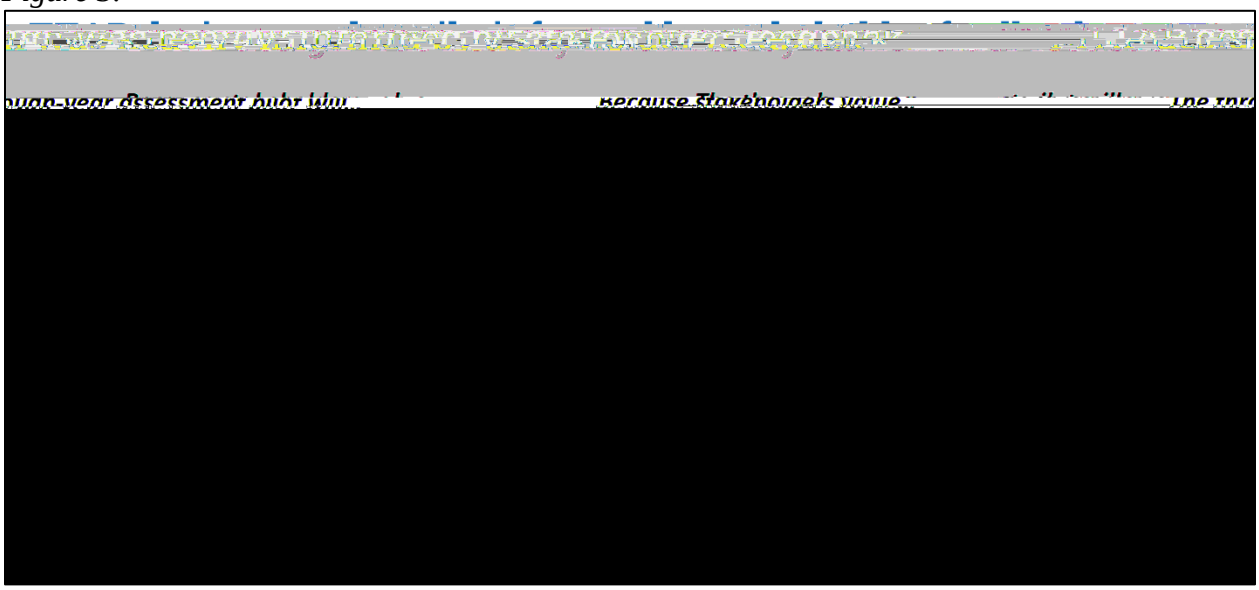
	annual production of items and are cost-intensive to accommodate students with special needs. • **Item-level computer adaptive designs** are the most individualized at the student level and provide greater measurement precision for the lowest and highest student abilities (which could include above and below grade items). But this model is the least transparent (	

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.

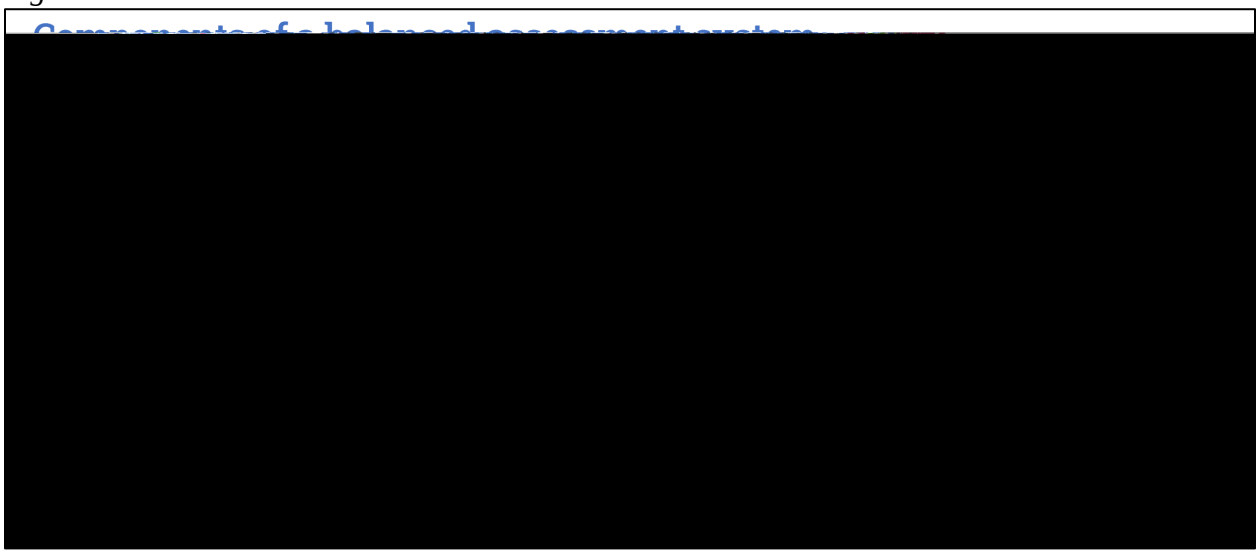
TTAP was developed to adhere to the same rigorous standards of STAAR, with every question undergoing Texas educator reviews as well as field testing. To ensure that all participating LEAs are able to maintain their local curriculum, each TTAP progress monitoring opportunity covers the full scope of the 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

- **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