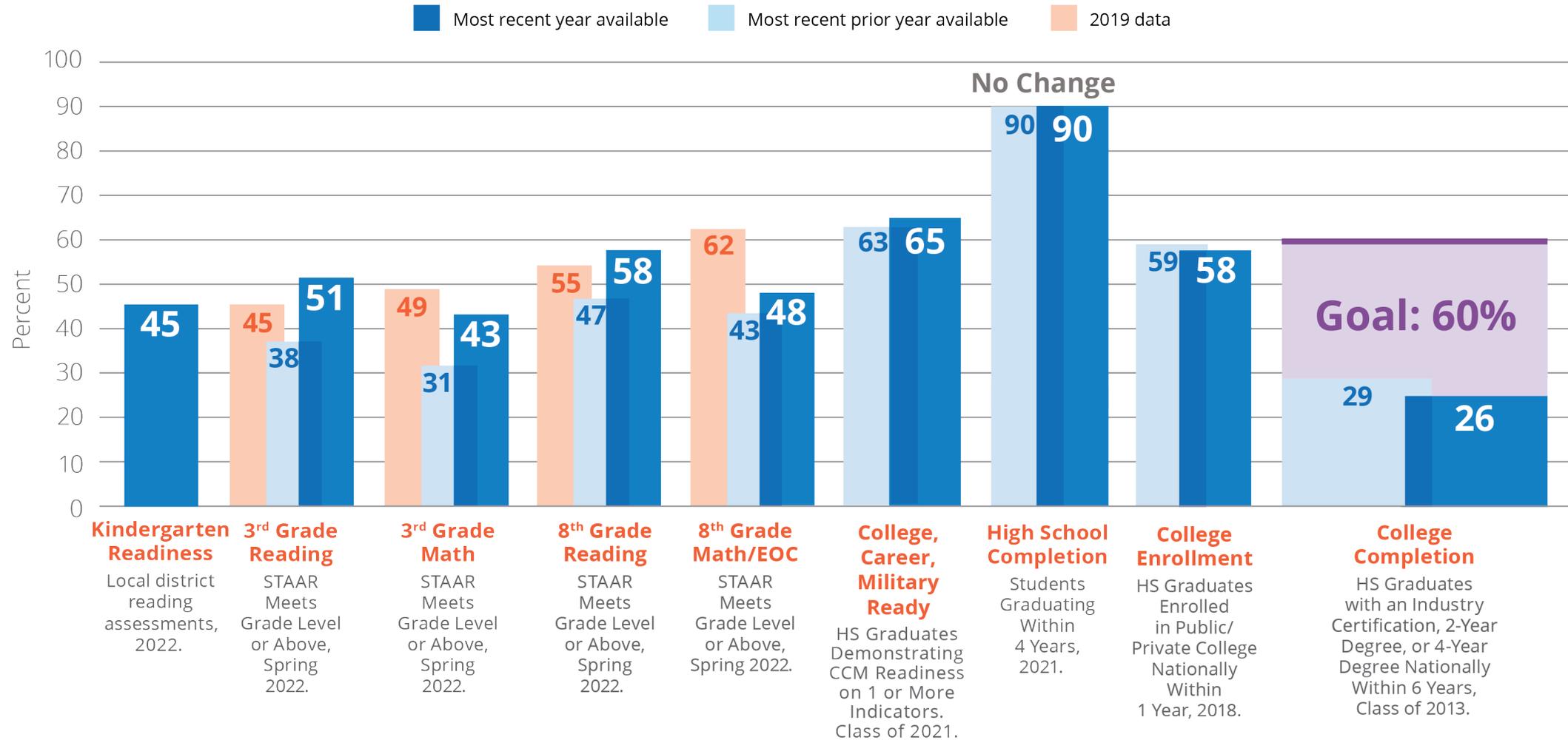




2022 Texas Assessment Conference

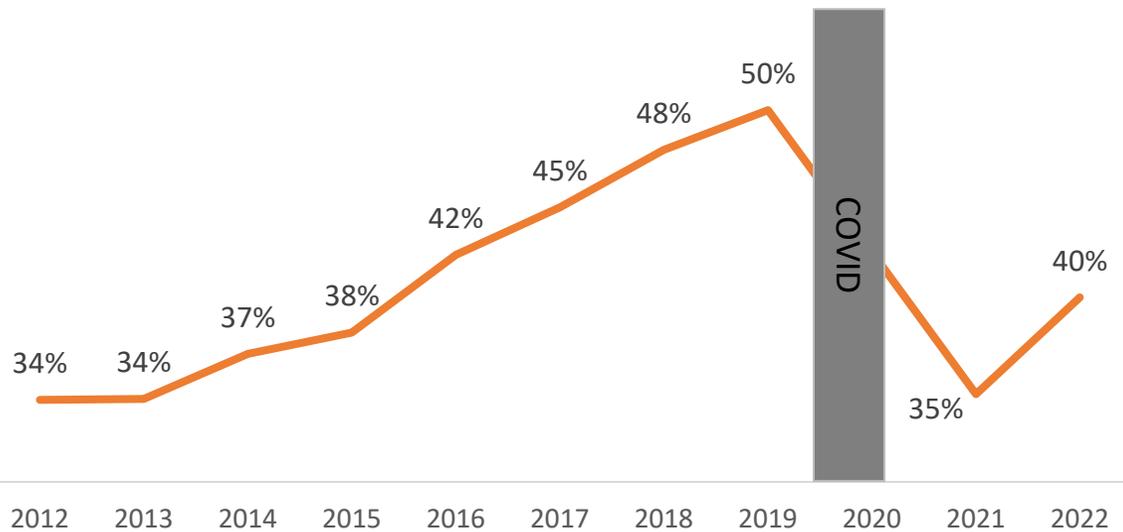
Our Outcomes

YEAR-OVER-YEAR STUDENT OUTCOMES

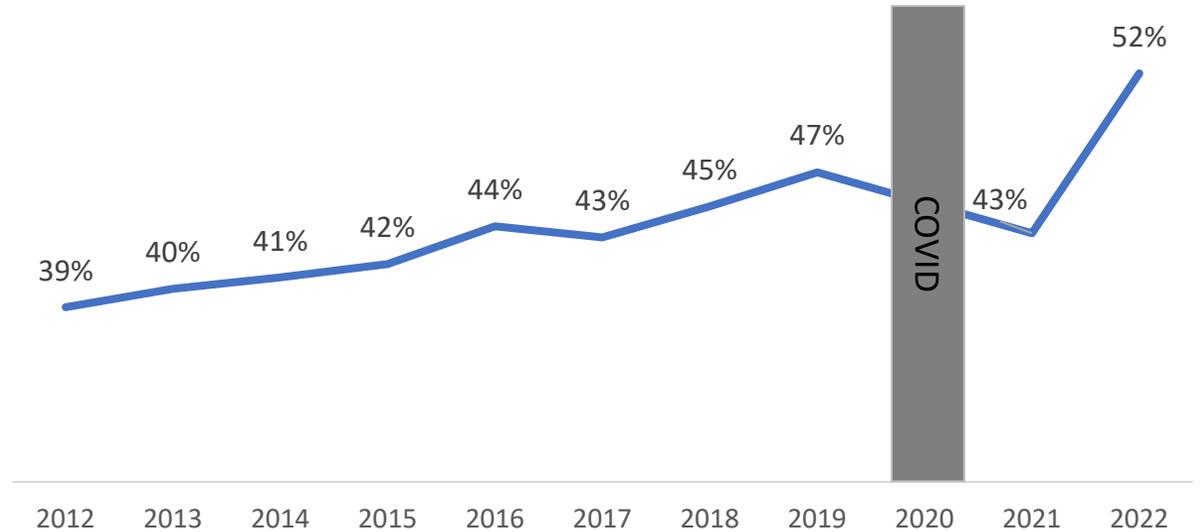


We are beginning to see some post-COVID success

Percentage of Students that Met Grade Level or Above in Math (Grades 3-8 and Algebra I)



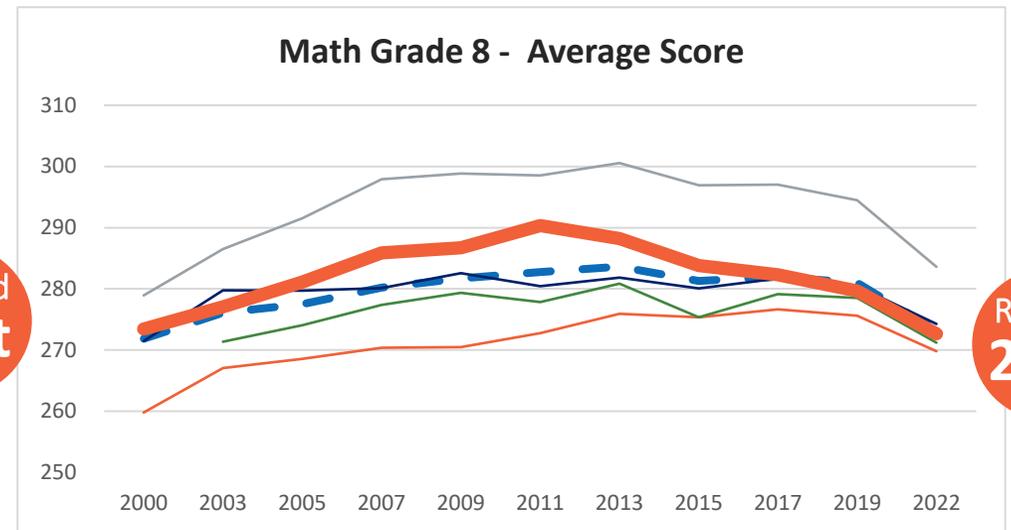
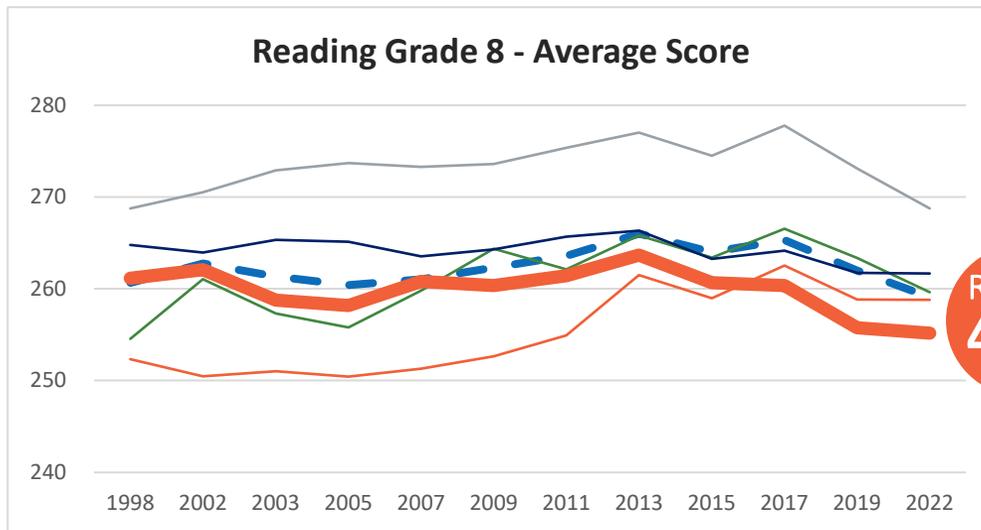
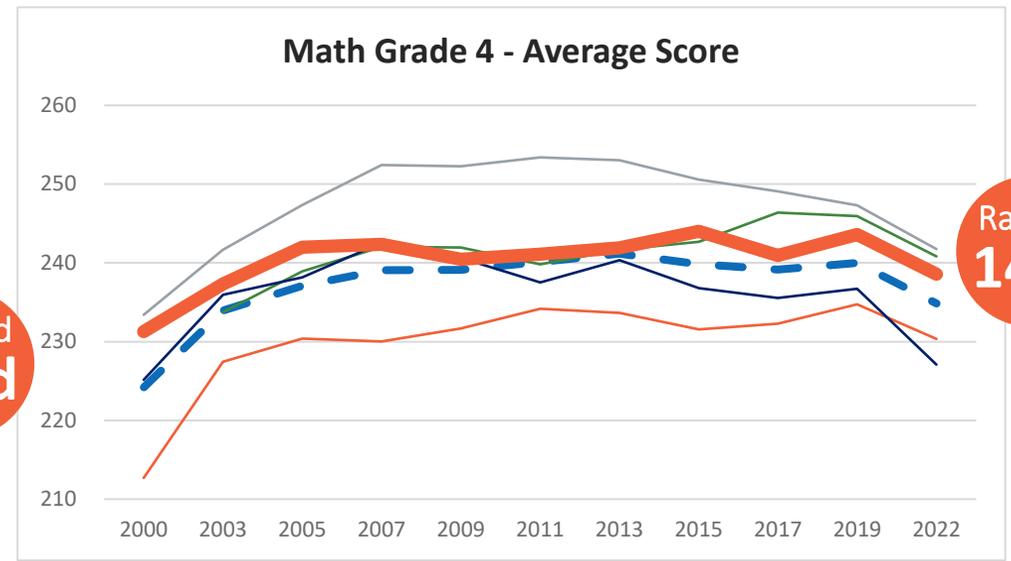
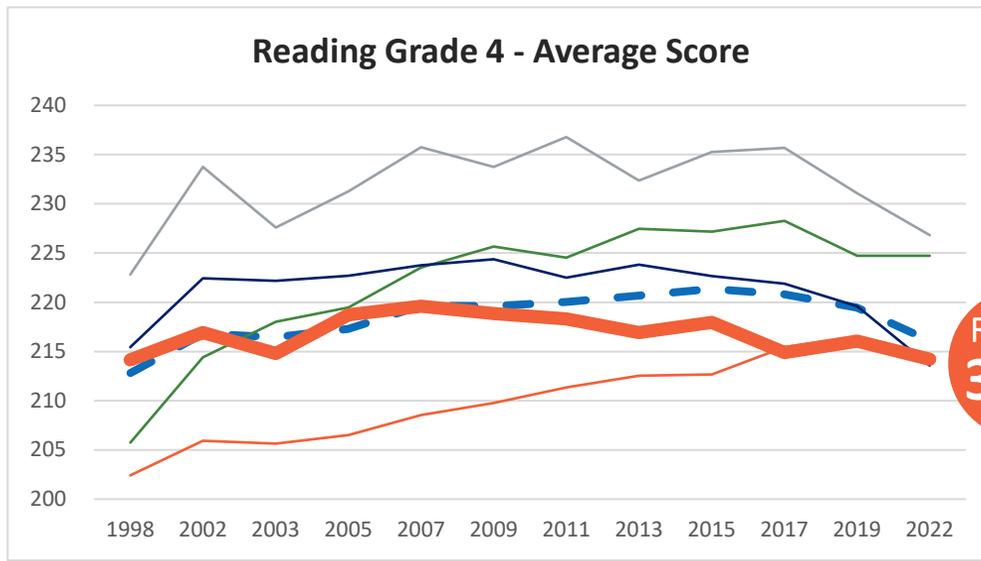
Percentage of Students that Met Grade Level or Above in Reading Language Arts (Grades 3-8, English I & II)



2022 NAEP All Students

KEY:

	USA
	TX
	CA
	FL
	MA
	NY



	Economically Disadvantaged		White Students		African American Students		Hispanic Students				
	4 th Grade	8 th Grade		4 th Grade	8 th Grade		4 th Grade	8 th Grade			
Reading	17 th	22 nd	Reading	15 th	40 th	Reading	6 th	6 th	Reading	21 st	26 th
Math	6 th	9 th	Math	2 nd	24 th	Math	1 st	1 st	Math	4 th	7 th

“If you don’t measure it, you don’t know what to improve”

Fitness



Weighing yourself regularly helps with hitting weight loss targets.

University of Pittsburgh, University of California, San Francisco School of Medicine.

Health



Food journals can dramatically reduce the progression of type 2 diabetes.

American Heart Association

Business



Firms who conduct routine budget audits have increased profitability.

Harvard Business Review

Students Are Helped In School & In Life

Monitoring performance with summative assessment results using school ratings has been shown to have long term benefits for students:

“

Our analysis reveals that pressure on schools to avoid a low performance rating led low-scoring students to score significantly higher on a high-stakes math exam in 10th grade. These students were also more likely to accumulate significantly more math credits and to graduate from high school on time. **Later in life, they were more likely to attend and graduate from a four-year college, and they had higher earnings at age 25.**

”

The biggest risks come if the system allows certain students to be exempted from accountability. **The system design matters.**

Source: <https://www.educationnext.org/when-does-accountability-work-texas-system/>



Different types of assessment serve different purposes

1. Diagnostic



What: A test measuring student knowledge and skills on any variety of student expectations

When: Prior to new instructional cycle or school year

Why: To inform instructional plans and curriculum to meet the needs of individual students

Example: Beginning of Year (BOY) assessments

2. Formative



What: Ongoing process of measuring student performance on specific student expectations

When: Often, throughout the year

Why: To inform instructional choices, student supports, and updates to planning within existing curricular structures

Example: Curricular-embedded tests administered via TFAR, and unit assessments included within high quality instructional materials

3. Interim



What: Measure student performance and understanding against grade-level standards

When: At check-points a few of times a year

Why: To monitor progress, predict summative performance, and identify students for intervention

Example: STAAR Interim Assessments

4. Summative

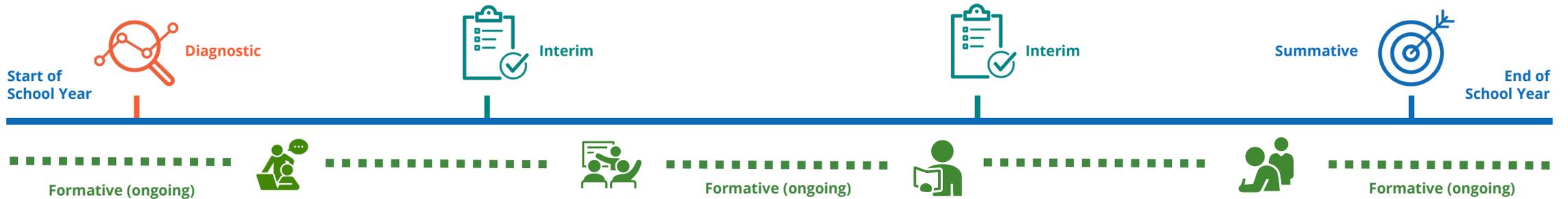


What: Measure student mastery of a broad span of student expectations

When: At the end of an instructional cycle or school year

Why: Campuses and districts use data to determine effectiveness of their programs, report summative mastery, and inform future planning

Example: STAAR, STAAR Alternate 2, TELPAS, and TELPAS Alternate



Testing doesn't just evaluate learning; it can improve it

The Testing Effect



“

One of the most striking research findings is the **power of active retrieval—testing—to strengthen memory**, and that the more effortful the retrieval, the stronger the benefit.

”

The act of retrieving learning from memory has two profound benefits.

- 1** It tells you what you know and don't know, and therefore where to **focus further study** to improve the areas where you're weak.
- 2** Recalling what you have learned **causes your brain to reconsolidate the memory**, which strengthens its connections to what you already know **and makes it easier for you to recall in the future.**

Make it Stick: The Science of Successful Learning





STAAR Redesign

The STAAR redesign is based on improving alignment to the classroom experience

In effective classrooms, teachers are...

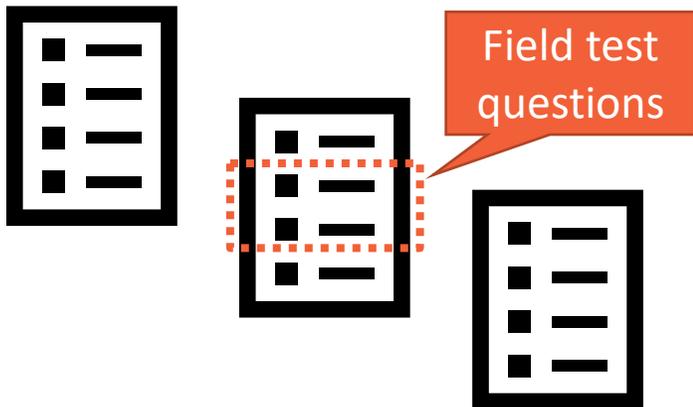
The STAAR redesign will...

-  Coherently building students' **background knowledge and vocabulary** in all subject areas   Prioritize **cross-curricular passages** in RLA that reference topics that students have learned about in other classes.
-  Asking students to **write about what they read using evidence from text**   Include **writing in all RLA tests**, reflecting our updated TEKS, and having **students write text-based responses**
-  Providing **various open-ended formats** for students to respond to questions...   Add **new, non-multiple-choice questions** that are more like questions teachers ask in class
-  Supporting the learning needs of all students by providing **appropriate accommodations**   Move to online assessments that provide a full suite of **robust accommodations** for students with specific learning needs
-   Moving to **online assessments** supports all the changes above and provides faster test results to support accelerated learning.

The redesign does not mean the test will be harder

While individual items can be easier or harder in a given year, the mix of item difficulty is balanced across years by using field test results

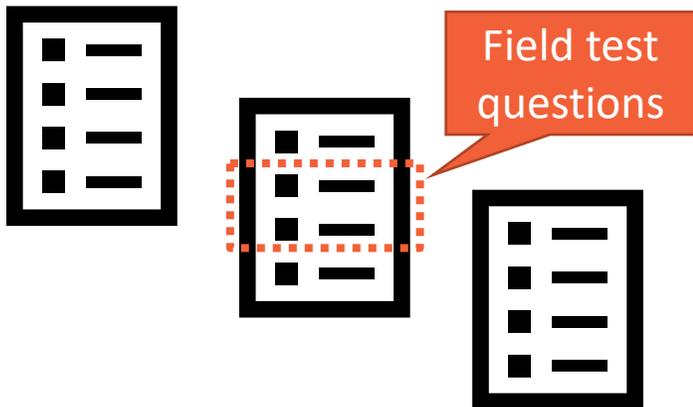
On each STAAR test, a small number of questions do not count towards the student's score. These are **field test questions**.



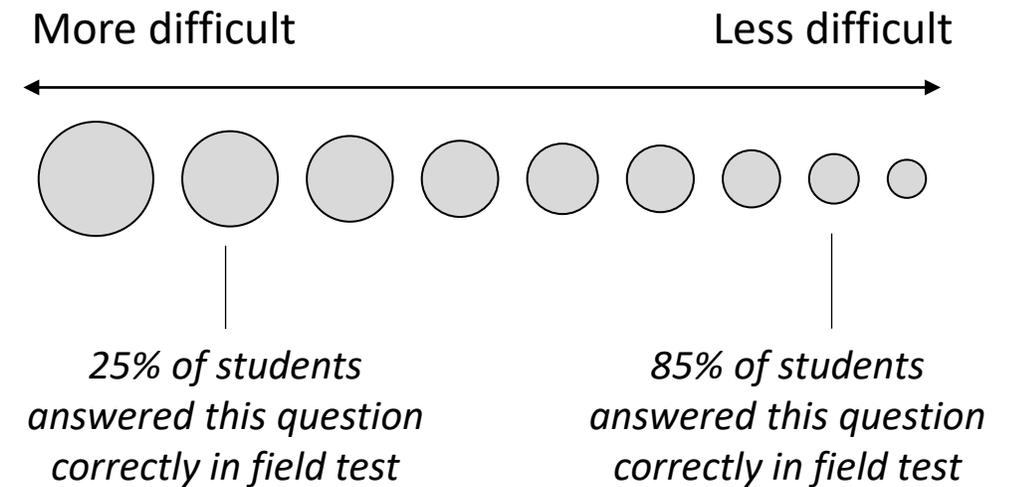
The redesign does not mean the test will be harder

While individual items can be easier or harder in a given year, the mix of item difficulty is balanced across years by using field test results

On each STAAR test, a small number of questions do not count towards the student's score. These are **field test questions**.



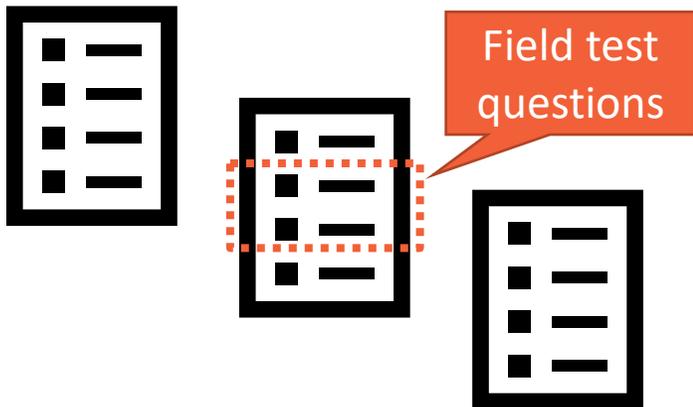
Through field testing, we determine how hard a question is (e.g., 80% of students got the question right).



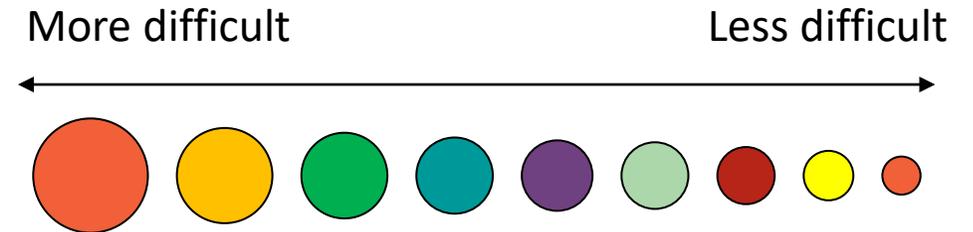
The redesign does not mean the test will be harder

While individual items can be easier or harder in a given year, the mix of item difficulty is balanced across years by using field test results

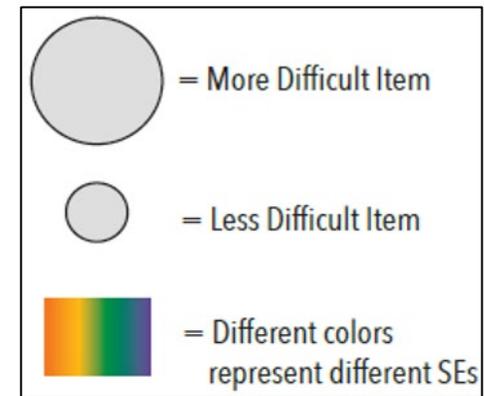
On each STAAR test, a small number of questions do not count towards the student's score. These are **field test questions**.



Through field testing, we determine how hard a question is (e.g., 80% of students got the question right).



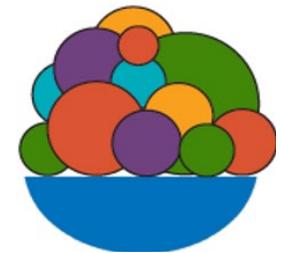
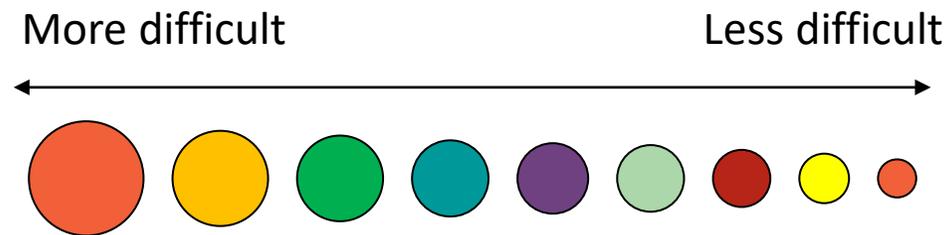
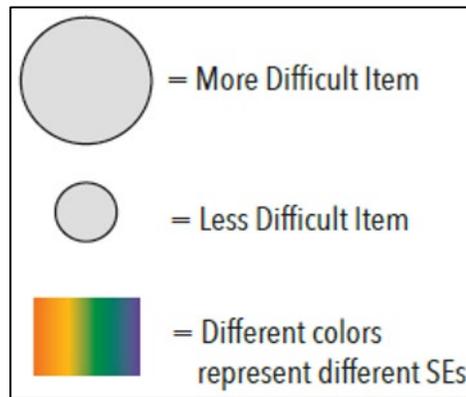
Questions also represent a variety of different student expectations



The redesign does not mean the test will be harder

While individual items can be easier or harder in a given year, the mix of item difficulty is balanced across years by using field test results

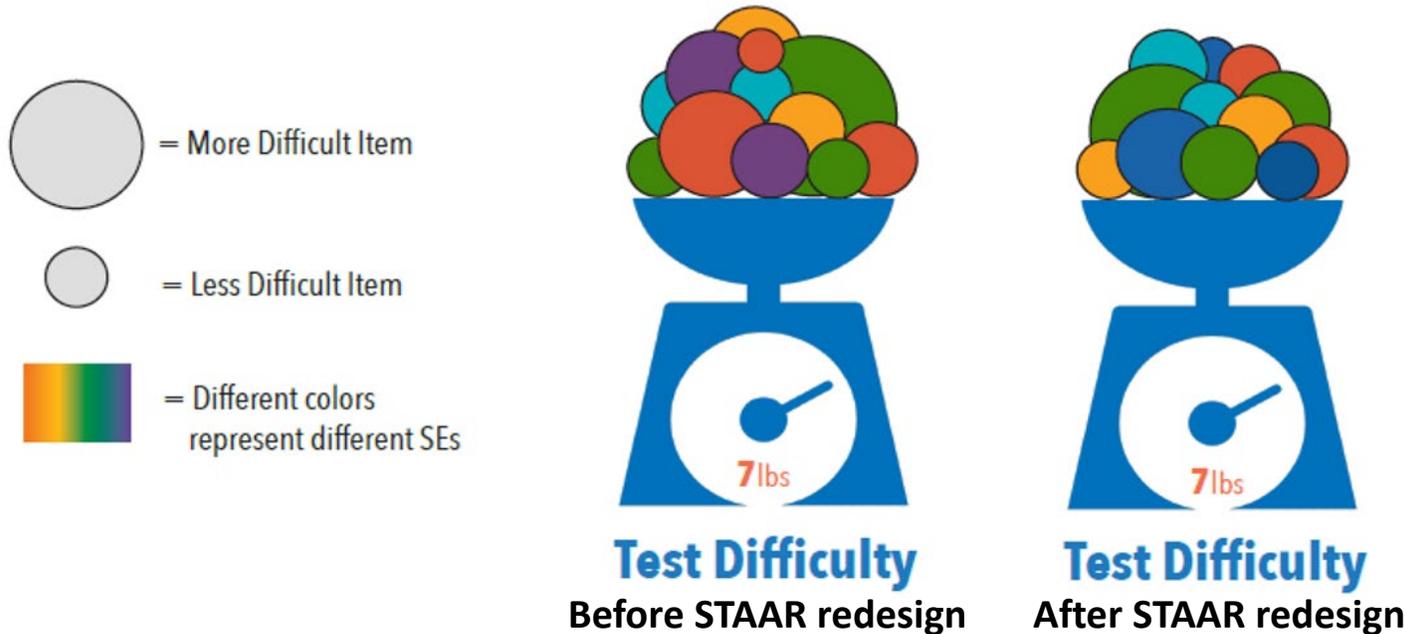
After questions have been field tested, they can be used to build STAAR tests



STAAR test

The redesign does not mean the test will be harder

While individual items can be easier or harder in a given year, the mix of item difficulty is balanced across years by using field test results



Texas educators are key to designing and building high quality assessments

Classroom teachers, instructional coaches, campus and district content specialists, and campus administrators can serve in a variety of ways:

Assessment Design and Standard-Setting:

- **Subject-area advisory groups** – groups of educators are convened to provide feedback on subject-area-specific assessment design topics
- **STAAR redesign focus groups** – groups of educators are convened to provide input on implementation of the components of the STAAR redesign
- **Standard-setting meetings** – groups of educators are convened to provide recommendations on cut scores for performance standards

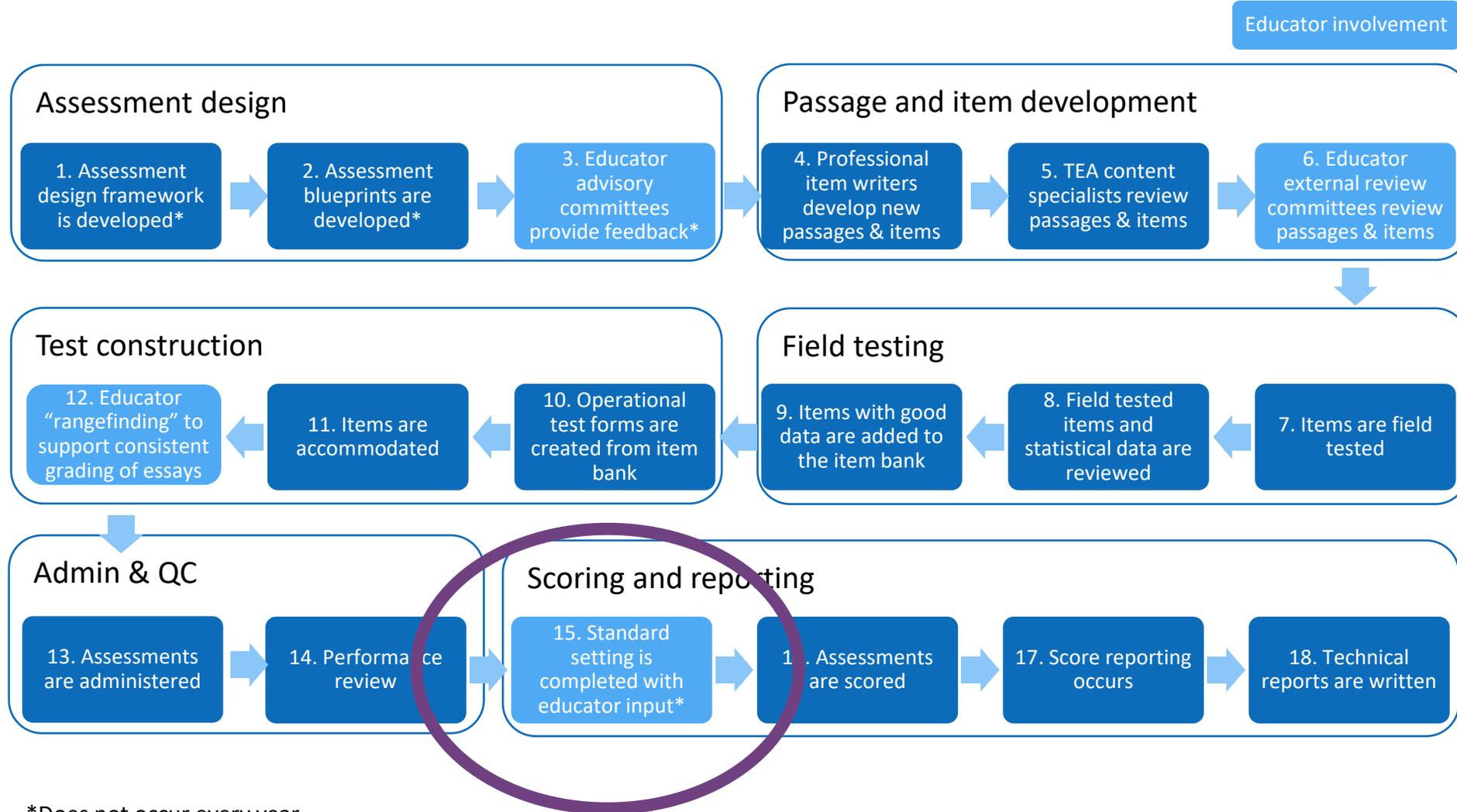
Passage and Item Development and Test Construction:

- **Educator passage review** – each potential passage for the RLA test is reviewed and approved by a committee of Texas educators
- **Educator item review** – each potential question for a state test is reviewed and approved by a committee of Texas educators
- **Constructed response rangefinding** – educators are convened to set the scoring boundaries for student essays based on the rubric

Visit the [Texas Assessment Learning Management System](#) to apply



Creating High-Quality Assessments is a Rigorous Process



Standard Setting

Standard setting is the process for defining what it means to be on grade level. The process includes scaling and linking studies and teacher committees.

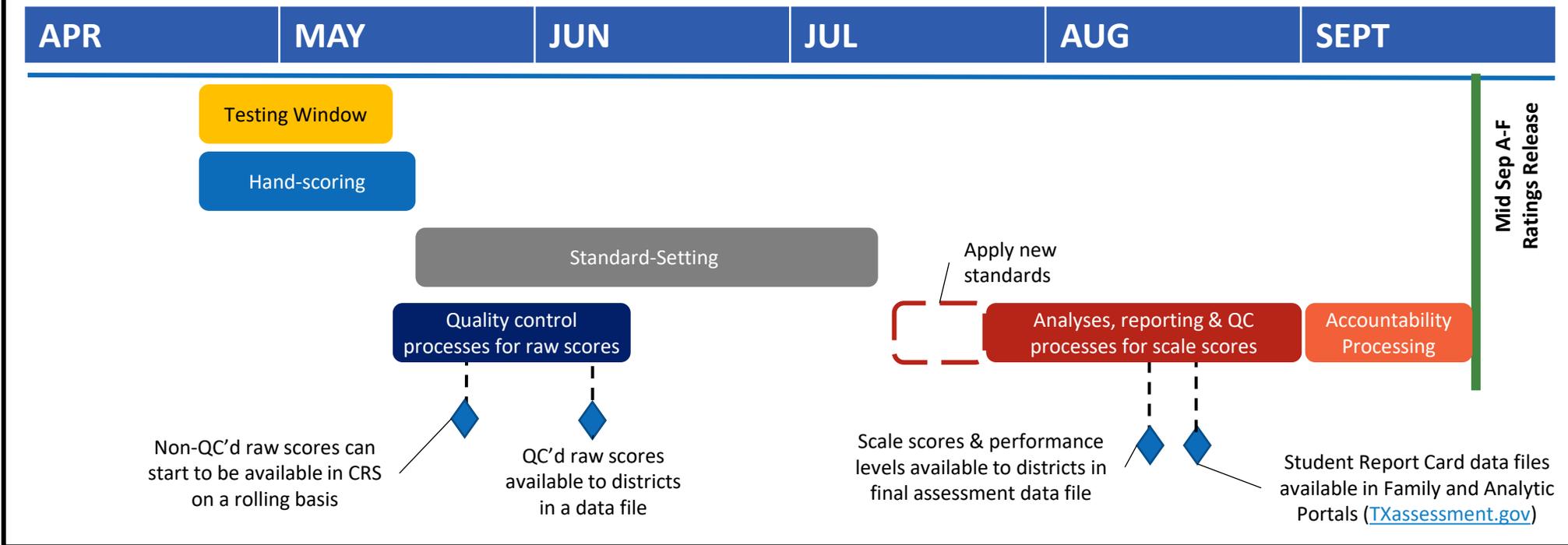
Standard-setting is conducted when the assessment changes significantly (e.g., when TEKS are revised), and periodically in-between significant changes to validate the continued accuracy of results.

*Does not occur every year



In future years, results will be faster. But in the 2023 redesign year, results for 3-8 will be later than usual.

Tentative 2023 STAAR 3-8 Scoring and Reporting Internal Process Timeline



A TAA on 10/20 describes the specific timeline, and how to plan for possible summer support under HB 4545.

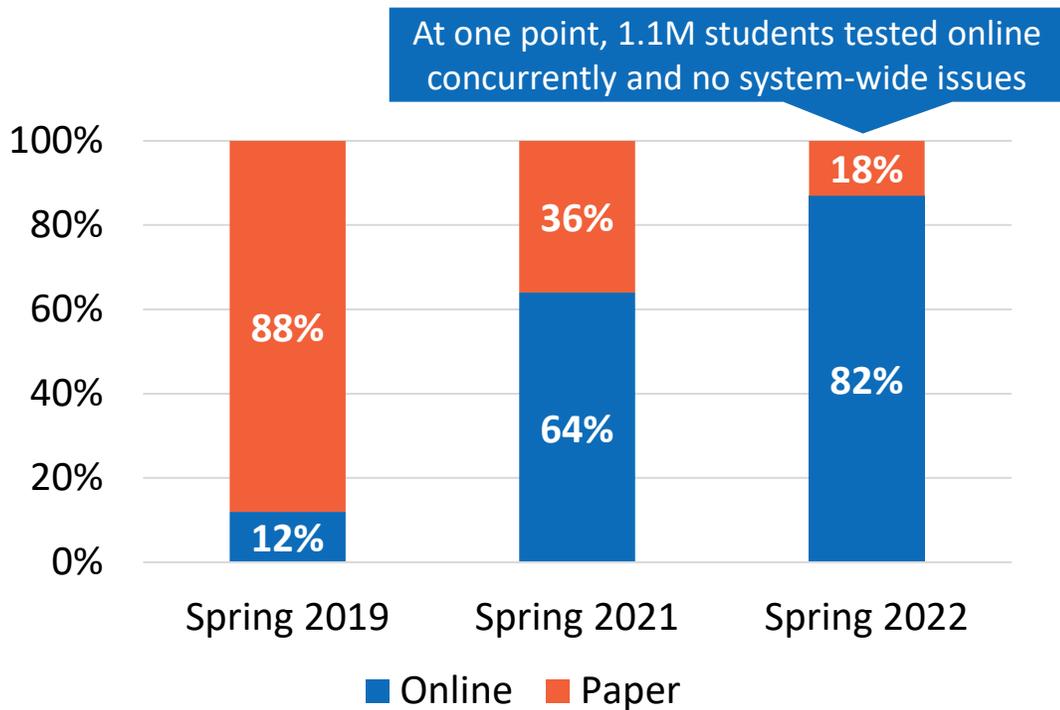
TEA process improvements have led to faster results. For example, in 2015, when new math TEKS were implemented and standard-setting was conducted for all Math STAAR tests, results weren't available to districts and families until October.

All dates are tentative



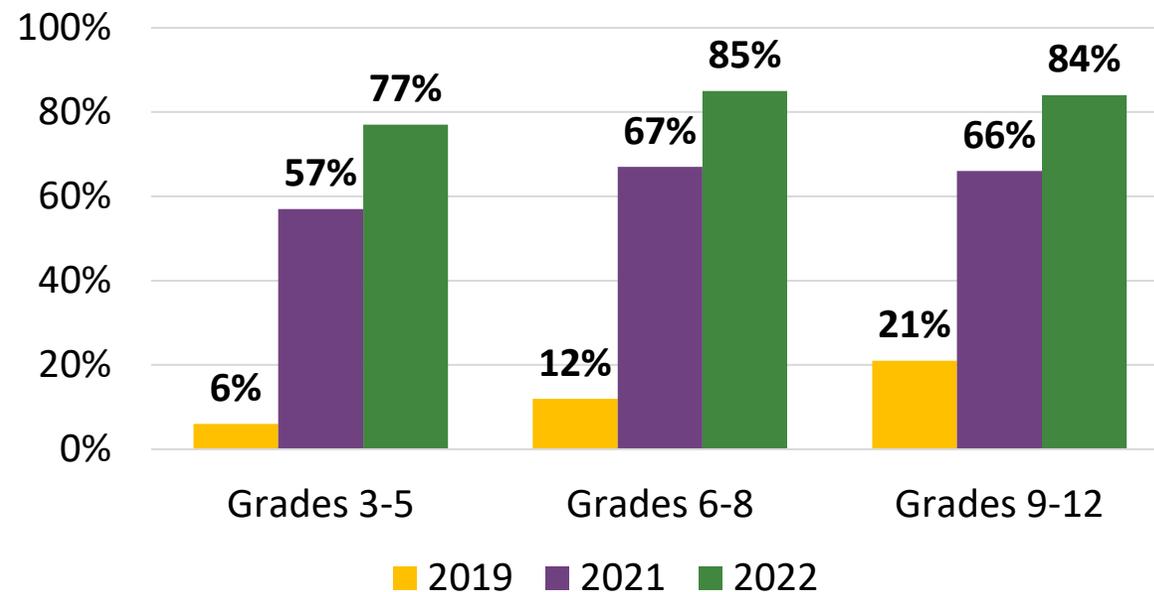
Online testing has steadily increased since 2019 with 82% of students testing online in 2022

STAAR online participation increased by **70 percentage points** since 2019.



Click to see STAAR FAQ video: [“How will the transition to fully online testing affect students’ performance on STAAR?”](#)

We saw increases in students testing online **across all grade bands**.



Click to see STAAR FAQ video: [“How do we know that young students will be able to type constructed responses on the redesigned STAAR tests?”](#)



Actions to Prepare for Success in the Spring

We'd like to provide ways for students to meaningfully interact with the online testing platform before online testing in Spring 2023

- Ideally, we want these interactions to be existing assessments, not additional assessments that take up valuable instructional time
- We don't want educators to create assessments just for the sake of giving students practice with the system
- The data generated by the assessment should be actionable

As a result, we are providing multiple ways for LEAs to administer meaningful assessments in the online testing platform:

Beginning-of-Year Diagnostic Assessment

LEAs can administer **released STAAR tests** as beginning-of-year diagnostics (this ended on 10/21/22).

This should only be used if the LEA plans to use the resulting data.

Interim Assessments

LEAs can administer **STAAR interim assessments** 1-2 times per year to monitor student progress.

These shouldn't be used if the LEA uses other interims or benchmarks.

Formative Curricular-embedded Assessments

LEAs that have adopted **TEA's core OER instructional materials** can administer curricular-embedded assessments in **TFAR**. Other LEAs can recreate their existing unit tests in **TFAR**.

These should be aligned to instructional materials.



Available STAAR Redesign Resources Currently Include:

- A. Full length practice tests that resemble the redesigned STAAR, and new question type samplers by content area and grade level
- B. Overview of new question types by content area and grade level
- C. Scoring and reporting guides by content area for new question types (constructed response scoring guides **released in October**)
- D. Updated blueprints by content area and grade level
- E. Policy on which students qualify for a special paper administration
- F. Updated STAAR Redesign FAQs
- G. Training and materials to support communications with teachers **to be released this month.**



Some STAAR Redesign Key Takeaways From Today

- Redesign STAAR goes live this school year, Spring 2023.
- The STAAR redesign was informed by an unprecedented stakeholder input process, including over 700 Texas educators and 200 Texas students. In addition, all STAAR items continue to be reviewed and approved by groups of current Texas teachers.
- Based on the input of educators, the new STAAR is more reflective of quality instructional practices, while continuing to serve as an accurate summative assessment of the TEKS.
- As part of typical development processes, the test is equated so that, as a whole, it is neither easier nor harder than before (in reading). But we are newly assessing writing and will engage in standard setting to ensure we have accurate definitions of what it means to be on grade level.
- Because we must do standard setting, test results will be reported to both district staff and families later than usual. Also because of this, *A–F* scores will be issued about a month later than usual.
- The Agency has provided some tools to help ensure a smooth experience for your students in the spring, including Interims and TFAR on the same testing platform.

