A young boy with short dark hair, wearing a plaid shirt, is sitting at a desk and smiling as he looks at a laptop. The background is a bright, out-of-focus indoor setting with a window and a white coffee cup on a saucer on the desk.

Texas Through-year Assessment Pilot (TTAP) Overview

Winter 2022

Executive Summary

- House Bill 3906 (86R) required TEA to develop an integrated formative assessment pilot that supports instruction and can potentially replace the current state summative assessment
- Texas Through-year Assessment Pilot (TTAP) was originally scheduled to launch in school year 2021-22 but was delayed until school year 2022-23 due to COVID-19*
- Approximately 10% of districts across Texas opted to participate in year 1 of the pilot; the first testing window occurred in November 2022
- TTAP is intended to be piloted for at least four years so that it can be rigorously analyzed to determine if it can serve as a valid and reliable summative assessment

*Because summative testing was cancelled in school year 2019-20, it was not possible to field test items for the pilot.

TEA launched an optional, small-scale pilot in SY 2022-23; multiple years of piloting is required to determine if this system can replace our current summative test

A through-year assessment model has many benefits...

- Provides **more timely and frequent feedback** that can be used to support instruction before students move on to the next grade or class
- Offers **multiple opportunities for students** to show what they've learned
- Allows for **in-year growth** information

...but is still relatively new and innovative

- Texas will need to address **technical questions** around design, administration, and scoring
- Pilot will be rolled out over **multiple years** prior to potential adoption (based on STAAR comparability, stakeholder feedback, and legislative input)



| 2022-2023 | 2023-2024 | 2024-2025 | 2025-2026 |
|--|------------------------------|---------------------|--|
| Pilot Year 1 | Pilot Year 2 | Pilot Year 3 | Pilot Year 4 |
| <div data-bbox="186 963 637 1149" style="background-color: #0070C0; color: white; padding: 5px;"> Grade 5 Science Grades 6 and 7 Math Grade 8 Soc Studies </div> | <i>Report to legislature</i> | | <i>Report to legislature – earliest possible decision to potentially replace STAAR with through-year model</i> |

All pilot participation is optional; no new testing requirements, and no requirement for district participation

10% of districts across the state opted into year 1 of TTAP

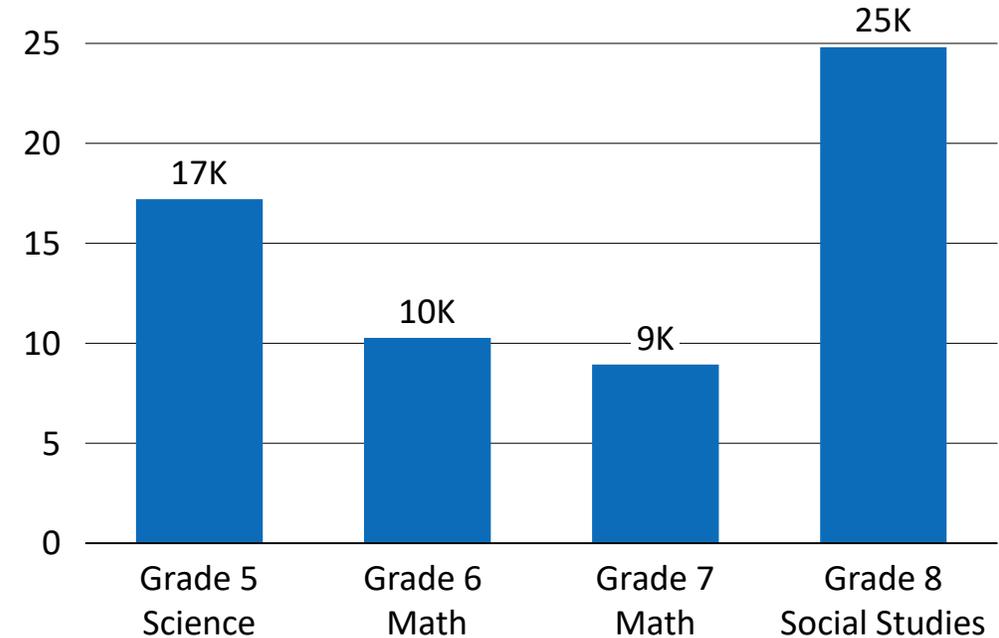
121

Districts participated

19

Regions represented

Number of students per title
(Opportunity 1, November 2022)



Note: Any participation by districts is optional and does not eliminate a district's obligation to administer the STAAR test.

TTAP's design was informed by stakeholders' feedback

Because stakeholders* value...

 A more cohesive assessment system that can replace existing benchmarking assessments

 Assessments that minimize the disruption of instructional time

 Providing students with multiple opportunities to demonstrate proficiency

 Preserving local scope and sequence of curriculum

 Providing measures of in-year growth to track student performance within the year

 More timely and frequent feedback

The through-year assessment pilot will...

1 Be **administered three times a year** (fall, winter, spring), serving as viable replacement to locally adopted district benchmarks

2 Limit the amount of test time across the year by leveraging a **computer adaptive model**

3 Explore a cumulative scoring model in which **earlier performance can help but not hurt** students' final scores

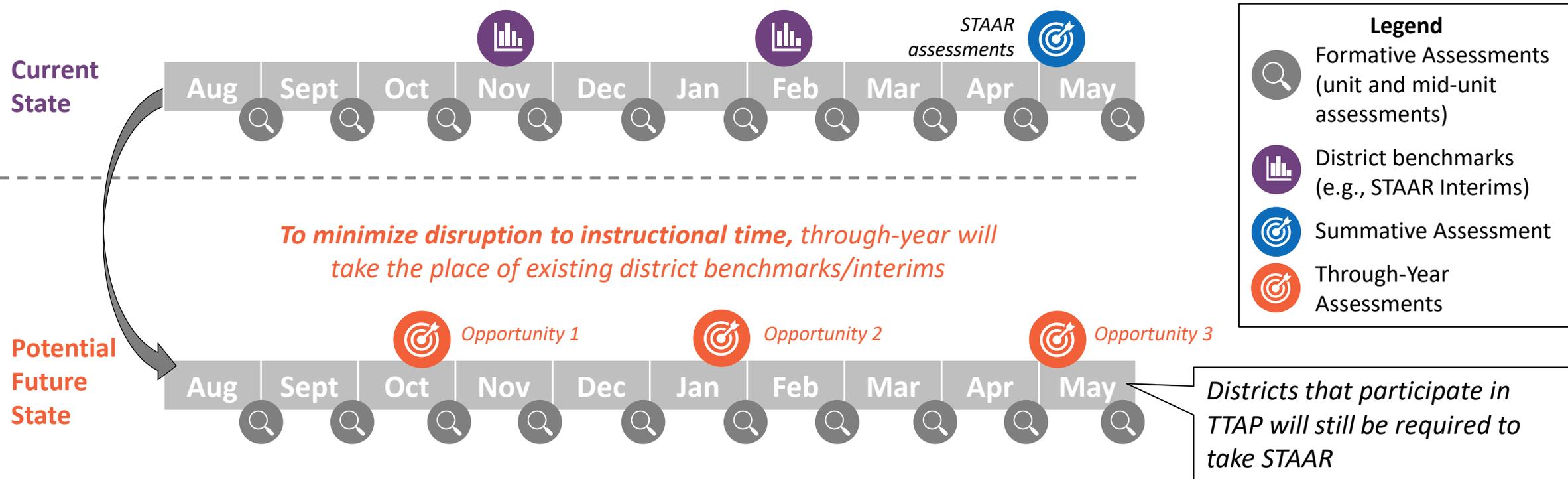
4 Be **full scope** for every testing opportunity (covering entire curriculum proportionately to the STAAR blueprint)

5 Be **fully online**, yielding **immediate reports containing different types of data** after each test opportunity

*Stakeholders engagements include – Educator Advisory committee and subcommittee meetings, CAO council presentation, superintendents survey, teacher and parent focus groups, student focus groups

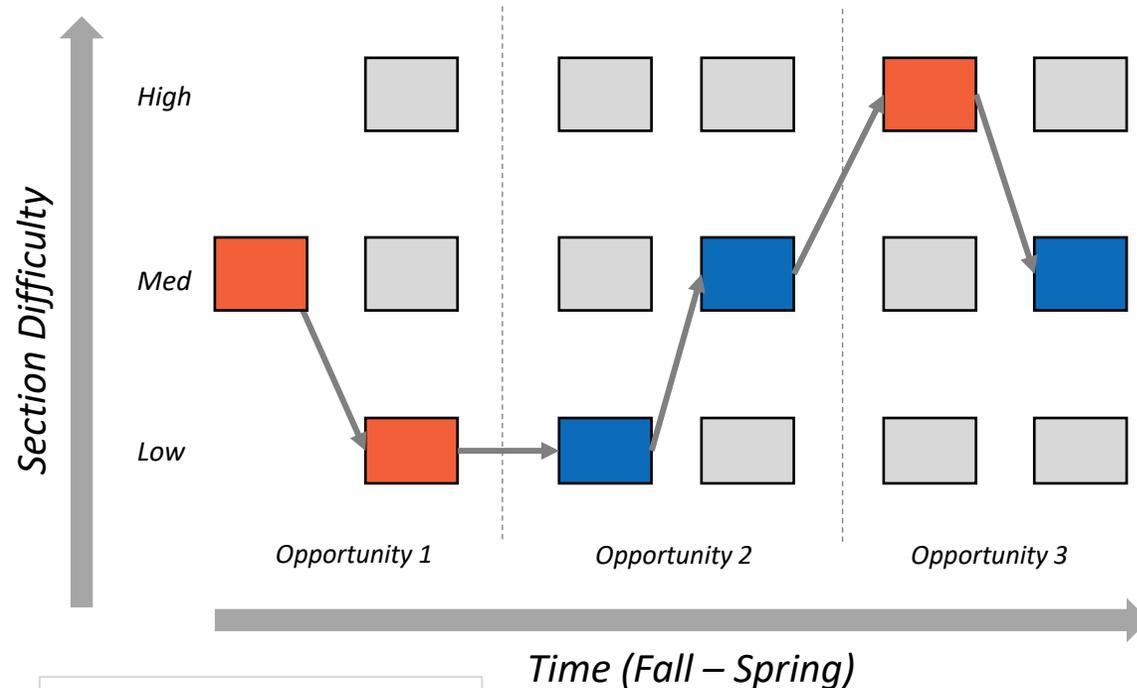
TTAP is designed to replace both benchmarking/interims and summative tests, combining them into one cohesive system across the year

1 Administered **3x year (fall, winter, spring)**, serving as viable replacement to locally adopted district benchmarks



TTAP's design minimizes the time required for testing because students respond to items tailored to their ability levels

2 Limit the amount of test time across the year by leveraging a **computer adaptive model**



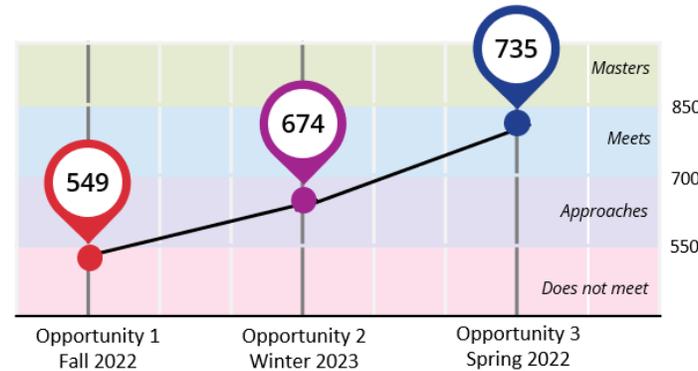
A computer adaptive model...

- ✓ Matches students with more appropriate items/sections based on their demonstrated ability
- ✓ Allows for shorter tests that fit within a class period (in Opportunity 1 and 2), minimizing disruptions to learning
- ✓ Will be administered online to ensure quick turnaround of results

Giving students multiple chances to show what they know would allow for additional opportunities to increase their end-of-year cumulative score

3 Explore a cumulative scoring model in which **earlier performance can help but not hurt** students' final scores

Scenario 1: Student scores the strongest in the third testing opportunity. It benefits them the most if their **final score is used** as the cumulative score for the year. Their prior testing performance did not hurt their cumulative score.



735
Meets Grade Level Expectations

Scenario 2: Student scores stronger in prior testing opportunities, compared the last test. It benefits them to use a **weighted average formula** to calculate the cumulative score. Their prior testing performance helped their cumulative score.

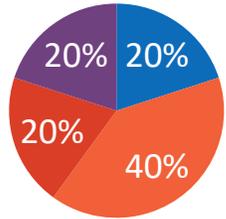


610
Approaching Level Expectations

*Pilot participants will not be given a cumulative score in Year 1;
Scoring will undergo further data study and are subject for further iteration after gathering initial data.*

A full scope test enables in-year growth reporting and allows districts to preserve their local scope and sequence

4 Be **full scope** for every testing opportunity (covering entire curriculum proportionately to the STAAR blueprint)



STAAR (~36 items)

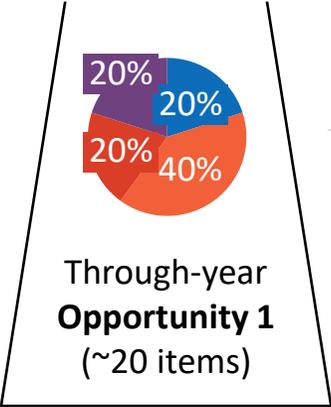
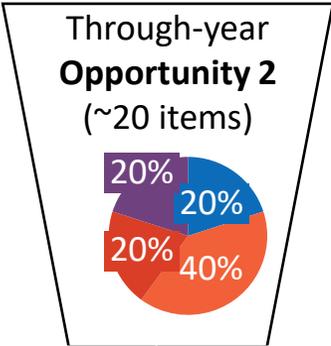
STAAR Grade 6 Math Blueprint Example

Reporting Category 1: Numerical Representations & Relationships (20%)

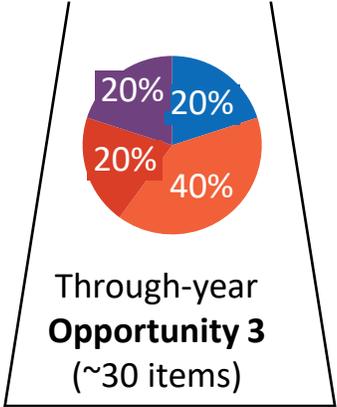
Reporting Category 2: Computations & Algebraic Relationships (40%)

Reporting Category 3: Geometry & Measurement (20%)

Reporting Category 4: Data Analysis & Financial Literacy (20%)



Full scope means the same proportion of content across reporting categories is covered at each point of the year

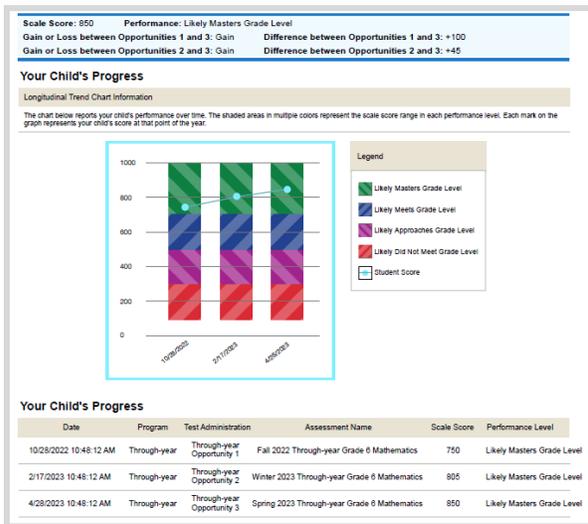


Although students will be tested on content they have not yet covered, the full scope approach is the only way we can provide **in-year growth measures** and allow districts to **preserve local scope and sequence**.

The data provided in TTAP will be packaged in different ways for different audiences

5 Be fully online, yielding immediate reports containing different types of data after each test opportunity

Individual score report printouts allow for *students to get an overview of their performance* while focusing on the most pertinent pieces of data at different points of the year. Teachers can also provide this to *parents to facilitate conversations about their child's progress during the year.*



CRS (Centralized Reporting System) allows *teachers and campus/district administrators to gain a bird's-eye view of student performance, as well as the ability to drill-down* into certain demographics or at the student-level. The level of data provided is dependent on the role of the individual (i.e., teachers can only see their own classrooms and aggregate data).

Average Score and Performance Distribution for STAAR Through-year Assessment Pilot Grade 6 Math (Through-year Opportunity 1), by Campus and Reporting Category: Demo District TX 1, 2022-2023
 Filtered By Campus: All Campuses | Test Administrations: Through-year Opportunity 1 | Standards Keys

| Campus | Total | Student Count | Average Scale Score | Performance Distribution | 5 Items on which Students Performed the Best | 5 Items on which Students Performed the Worst | 1. Numerical Representations and Relationships | | | | |
|------------------|-------|---------------|---------------------|---|--|---|--|---|---|------|------|
| | | | | | | | (1) Low Difficulty Average Percent Correct | (2) Medium Difficulty Average Percent Correct | (3) High Difficulty Average Percent Correct | 2 | 4 |
| | | | | | | | 6.1.7.B | 6.1.4.G | 1 pt | 1 pt | |
| ESC | | 25 | 484 | Percent Count: 64% 16, 20% 5, 12% 3, 4% 1 | | | 45% | 38% | 39% | 0.6 | 0.48 |
| District | | 25 | 484 | Percent Count: 64% 16, 20% 5, 12% 3, 4% 1 | | | 45% | 38% | 39% | 0.6 | 0.48 |
| Demo Campus T... | | 25 | 484 | Percent Count: 64% 16, 20% 5, 12% 3, 4% 1 | | | 45% | 38% | 39% | 0.6 | 0.48 |

All TTAP teachers are trained on how to interpret and use the data properly in tandem with formative data



The data provided after each progress monitoring opportunity will provide valuable insights to support instruction

5 Be **fully online**, yielding **immediate reports containing different types of data** after each test opportunity

Available Year 1

- Opportunity scale score
- Opportunity performance level
- Reporting category information
- In-year growth (*Opp 2 & 3 only*)*
- TEKS-alignment and difficulty level* for each question
- Item-level performance

Available Year 2+

- EOY prediction (*Opp 1 & 2 only*)
- Cumulative score (*after Opp 3*)*
- Cumulative performance level (*after Opp 3*)*
- Cumulative score comparison to district and state (*after Opp 3*)*

*Data element that is unique to TTAP

A photograph of a school staircase with several students. The students are wearing backpacks and casual clothing. The staircase has metal railings and is set against a large window that looks out onto a green landscape. The image is slightly faded and serves as a background for the text.

Appendix

Pilot Design Question: Cumulative scoring methods

| | Pros | Cons |
|---------------------------------------|--|--|
| Final Score Only | <ul style="list-style-type: none"> • Most psychometrically valid approach • Accounts for fall and winter performance based on where students start in the spring | <ul style="list-style-type: none"> • Does not fully meet pilot’s intention of giving students multiple opportunities to demonstrate proficiency |
| Weighted Average | <ul style="list-style-type: none"> • Helpful for students who perform well earlier in the year if all three tests are weighted into a cumulative score | <ul style="list-style-type: none"> • Unfairly punishes students who do not demonstrate proficiency early in the year |
| Maximum Score | <ul style="list-style-type: none"> • Takes the best scores out of the three individual testing opportunities | <ul style="list-style-type: none"> • Could provide perverse incentives to test takers (e.g. strong early performance may disincentivize students to give best effort later) |
| Final Only OR Weighted Average | <ul style="list-style-type: none"> • Most preferred method by stakeholders (earlier performance could help, but not hurt student cumulative score) | <ul style="list-style-type: none"> • Would require further study |

Preferred

Pilot Design Question: Curricular scope

| | Pros | Cons |
|---------------|---|--|
| Year 1 Design | <ul style="list-style-type: none"> Enables districts to keep local curricula and doesn't penalize students who switch districts during the school year Allows for within-year growth measures | <ul style="list-style-type: none"> Students will be tested on content they have not yet been taught during fall and winter |
| | <ul style="list-style-type: none"> Students aren't tested on content they haven't yet been taught | <ul style="list-style-type: none"> Requires all districts to adopt statewide curricula Does not allow for within-year growth measures |

 Full scope state assessments aligned to state standards

 Local formative assessments aligned to local curriculum



Although through-year assessments are full scope, districts will continue to use curricular-aligned formative assessments throughout the year

Two other states plan on using a through-year assessment model in the 2022-2023 school year

Florida's model is similar to our existing STAAR Interim Assessments if they were required rather than optional

| | Texas | Nebraska | Florida | Texas - Next Phase |
|------------------------------|---|---|--|--|
| Testing System | STAAR and optional STAAR Interims | NSCAS Growth | FAST | Through-year Assessment Pilot |
| Status 2022-23 | Operational | Operational | Operational | Pilot |
| Content areas & grade levels | Grades 3-8 math and RLA, 5 & 8 science, 8 S.S. and EOC tests | Grades 3-8 math and RLA | Grades 3-8 math, 3-10 RLA, 5 & 8 science | Select grades and content areas |
| Windows | 1 required spring summative 2 optional interims during the fall and winter | 3 tests administered during the fall, winter & spring | 3 tests administered during the fall, winter & spring | 3 tests administered during the fall, winter & spring |
| Design | Full scope, static tests for all spring summative tests (items released) Full scope, multi-stage computer-adaptive tests for interims (items released) | Full scope, item-level computer-adaptive tests (items not released) | Full scope, item-level computer-adaptive tests (items not released) | Full scope, multi-stage computer-adaptive tests (partial item release) |
| Cumulative Scoring | Cumulative score is a student's spring score | Cumulative score is a student's spring score , but a student's 'starting place' on the spring test is informed by the results from fall and winter | Cumulative score is a student's spring score ; will provide recommendation to legislature by Jan 31, 2025, of how to incorporate fall and winter scores in cumulative score | Cumulative score is a student's spring score or a weighted average of all opportunities, whichever is highest |

Texas pilot is the only one attempting to incorporate results from the first two tests into a student's final score

