TO: Clerk of the Senate
              Clerk of the House of Representatives

FROM: Charlene M. Russell-Tucker
              Commissioner of Education

DATE: February 10, 2026

SUBJECT: Report on the Comprehensive Audit of Assessments Administered to Students

In accordance with Section 1 of Public Act No. 24-93, the Connecticut State Department of Education (CSDE) submits this *Report on the Comprehensive Audit of Assessments Administered to Students* to the joint standing committee of the General Assembly having cognizance of matters relating to education.

If you have any questions, please contact Laura Stefon, Chief of Staff at (860) 713-6493.

CRT:ag
cc: Education Committee
      Legislative Library
      Office of Legislative Research
      State Library

Enclosure

Box 2219 • Hartford, Connecticut 06145
An Equal Opportunity Employer

REPORT ON THE

COMPREHENSIVE AUDIT OF ASSESSMENTS

ADMINISTERED TO STUDENTS

Submitted to the Connecticut General Assembly to

fulfill the requirements of Public Act 24-93

JANUARY 2026

\

CONNECTICUT
Education

CONTENTS

EXECUTIVE SUMMARY

Introduction

[Section 1 of Public Act No. 24-93](#) requires the Connecticut State Department of Education (CSDE), in consultation with national assessment experts and public school educators, to conduct a comprehensive audit of the assessments that are administered to students. The statute requires the CSDE to evaluate the assessments inventoried with the goals of eliminating redundant assessments, discouraging classroom activities that focus only on test preparation, reducing testing time, and maximizing assessments that provide actionable information for classroom teachers. The statute also charges the CSDE to develop and implement a program of professional learning for teachers concerning assessment literacy. CSDE consulted with school districts, national experts, and other state education agencies (SEAs) to develop and implement the assessment audit data collection.

Background

An Important Historical Perspective on Educational Assessment

A recent publication by the National Academy of Education titled *Reimaging Balanced Assessment Systems*[i] offers a rich historical perspective on the evolution of educational assessment practices. The 2001 bi-partisan passage of the federal No Child Left Behind (NCLB) act instituted a new mandate of annual large-scale standardized assessment as an educational accountability measure. In the same year that NCLB passed (i.e., 2001), the National Research Council published *Knowing What Students Know*[ii]. This publication emphasized that o**ne assessment cannot serve many purposes. "***Often a single assessment is used for multiple purposes; in general, however, the more purposes a single assessment aims to serve, the more each purpose will be compromised.***"** It articulated a vision for where *"assessments at all levels—from classroom to state—will work together in a system that is comprehensive, coherent, and continuous."*

Despite the 2001 vision for a coherent and aligned system of assessments at the classroom and state levels, what proliferated after the passage of NCLB was not a system of different assessment serving different purposes, but redundant testing in the form of mini-summative benchmark assessments. Many were commercial products while others were district created.

Such tests tried to mimic the state assessment, promoted test-prep/practice, and sought to "predict" performance on the state measure. They were not designed to foster **formative assessment practices**[iii]**,** which is an evidence-based strategy for improving student learning. This cycle re-occurred with the passage of the Every Student Succeeds Act (ESSA) when Connecticut's switch to the SAT as its high school accountability assessment led to new widespread PSAT 8/9 testing in Grade 9 and even some in Grade 8, even though multiple CSDE studies[iv] show that middle school Smarter Balanced results are strong predictors of high school outcomes, rendering the PSAT redundant.

Toward a Balanced Assessment System

When Connecticut adopted the Common Core State Standards in 2010, it joined the Smarter Balanced consortium of states primarily because of its commitment to offer a "balanced" system of assessments. **Different assessments and resources in the system serve different purposes**:

- The broad high quality, **end-of-year state summative** assessment provides a valid measure of a student's overall achievement and can be used for evaluation and accountability purposes. Preliminary results are available starting mid-May, so teachers can review the performance of their students and district/school teams can use the results for planning.
- Over 200 short and focused **interim assessment blocks (IABs)** are of much finer grain than the end-of-year-summative, assess narrow domains of content, and if used in formative ways, can also support instruction. Results are available immediately.
- **Formative assessment tools and resources[v]** are aligned to the assessment targets and interim results. They lend themselves to be used by "teachers during learning and teaching to elicit and use evidence of student learning."

Curriculum, Assessment, and Professional Supports

The CSDE has provided numerous guidance documents and professional development resources such as *The Types and Purposes of Student Assessment in Education: An FAQ* in 2016 and the *Sensible Assessment Practices* guidance in 2020. The CSDE has supplemented these resources with ongoing professional learning and customized technical assistance through the Sensible Assessment Practices webinar series. CSDE's Model ELA and Model Mathematics curricula that are available freely to districts integrates Smarter Balanced IABs/Focused IABs (F-IABs) thoughtfully and consistently across the grades.

Methods

The audit questions are listed in Appendix A. Most questions used dropdowns, where districts could select the best choice from a list of options, including "Other." The assessment audit forms were collected from districts and checked for completion. Out of 199 districts, 175 received personalized follow-up to ensure compliance with the audit instructions.

Districts were instructed to include the designated time for administering the test. Time used for test set up, instructions, or breaks from testing are *not* included in the audit.

To keep the data collection burden reasonable, the CSDE limited this collection to district-required assessments that are not modified by teachers; teacher-written or teacher-modified assessments were not included in this collection as that would have required data collection from thousands of teachers.

Findings

A comprehensive analysis of the data collected in the audit from all districts revealed the following:

- On average, students in Grades 3 through 8 spend around three-and-a-half to four hours taking the Smarter Balanced English Language Arts (ELA) and Mathematics state summative assessments. Students in Grades 5 and 8 on average spend an additional hour-

and-a-half taking the Next Generation Science Standards (NGSS) state summative assessment.

- In the same Grades 3 through 8, students on average spend an additional 20-21 hours taking district-required assessments that are not modified by teachers.
- In the high school Grades 9-11, students spend on average around three-and-a-half hours taking the CT SAT School Day and the NGSS state summative assessments. They spend on average an additional 19 hours taking district-required assessments that are not modified by teachers. It should be noted that teachers in high school are more likely to create and use their own classroom assessments (e.g., mid-terms, finals) which are not included in these estimates.
- With 900 hours of actual school work required in state law, the total time spent testing is approximately two to three percent of that total.
- There is wide variation among districts in the average number of hours spent taking district-required assessments. In Grades 3-11, around one quarter of the districts spend on average more than 26 hours on locally required assessments that are not modified by teachers.
- The most-selected main use for district- and school-required assessments was to *"Track/Report group and subgroup performance (e.g. benchmarking, measuring growth)"*. Many assessments used for this purpose are broad, summative assessments that measure what should be learned in a semester or the entire year; they are ***not*** designed to measure student learning after a shorter unit/lesson that may last 1-3 weeks. As such, their instructional value is limited. The most prevalent among these assessments are the PSAT (90 districts), i-Ready Diagnostic (53 districts), NWEA MAP Growth (33 districts), and STAR Renaissance (31 districts).
- There is some misalignment between the stated use of test and the type of test; for example, many of the broad local summative tests referenced above are also reported as being used to *"Identify specific strengths/weaknesses of students to inform instruction."* Using such broad assessments to evaluate and guide classroom instruction during the year is not a valid use of such assessments. To truly inform instruction, assessments need to be focused, of a much finer grain, measure more narrow content domains, and provide defensible evidence for instructional use.
- There is a need for broad assessment literacy training and intensive coaching to help districts reflect on their assessment choices, eliminate redundant assessments, use the right assessments for the right purpose, and maximize formative assessment practices that truly support teaching and learning.
- Informal conversations with districts over the years, including as part of the audit process, have revealed that state legislation and CSDE policy are sometimes perceived as requiring the use of certain local assessments. The assessments that are implemented are not always aligned to their intended uses. Additionally, the audit revealed the limited use of assessments in K-3 that are not aligned to the "science of reading" as well as assessments solely for mock/practice testing (i.e., test prep).
- Since 2017, CSDE's EdSight Secure platform has provided district/school leaders and coordinators across Connecticut with near real-time historical data about students in a new grade, school, or district. These data include prior test scores, attendance, disciplinary incidents, mobility, course grades, etc. The Early Indication Tool (EIT) in EdSight Secure also

identify students who may need additional support to reach academic milestones and allows for timelier, targeted interventions. These longitudinal data provide valuable and reliable information for immediate use, without needing to administer one more test.

Recommendations

In light of these findings, the strategic actions described below can serve as a foundation for the continued work of the CSDE and for requisite legislative proposals that are designed to accomplish the goals of this audit. To recap, the goals of this audit were to eliminate redundant assessments, discourage test prep, reduce testing time, maximize assessments that provide actionable information for classroom teachers, and provide professional learning on assessment literacy.

The following specific strategic actions are recommended and can lead to more coherent local assessment systems that prioritize assessment **for** learning:

- Establish an **incentive program** for districts who can demonstrate that they have:
    - limited the time that students spend on broad tests during the year;
    - thoughtfully integrated the state-provided IABs/F-IABs and formative assessment tools into the local curriculum in a manner that supports ongoing instructional improvement; and
    - increased teacher competency in the formative assessment process.

    Incentive program awardees should be expected to serve as mentors for others looking to reduce assessment and increase local coherence.

- Issue **updated guidance and policy** that:
    - specifies the appropriate and inappropriate use of assessments;
    - discourages use of assessments solely for test prep;
    - directs discontinuance of assessments that are not aligned to the research;
    - recommends eliminating the fall and spring administration of broad local assessments to reduce redundancy and student testing fatigue;
    - illustrates the use of EdSight Secure longitudinal data (e.g., attendance, behavior, assessment, course grades, mobility) for placement/grouping in the new grade;
    - encourages use of state-provided IABs/F-IABs in place of broad assessments or other end-of-unit classroom summative assessments to evaluate learning during the year; and
    - demonstrates the formative use of IABs/F-IABs along with other state-provided formative tools to support instruction.

- In line with recommendations of the working group established pursuant to Public Act 24-45, explore the feasibility of and appropriate timing for seeking federal waiver to make **changes to the high school accountability model** by reducing the weight assigned to the high school assessment and increasing the weight for college- and career-oriented measures. This can also incentivize districts to reduce broad locally administered assessments.

- Provide **assessment literacy training and intensive coaching**.
  Note that the CSDE is contracting with assessment experts from the *National Center for the Improvement of Educational Assessment* – the preeminent national experts on assessment – to develop and implement a multi-year system of professional learning, coaching, and technical assistance. This project –branded as Student Centered Assessment for Learning (SCALE) – builds on these assessment audits to directly address findings related to inefficiencies, incoherence, and limited usefulness in current assessment practices. SCALE will develop teacher and leader capacity, and build regional coaching and leadership infrastructure through the regional educational service centers (RESCs).

- Offer professional learning opportunities so educators can increasingly leverage the **longitudinal student information in EdSight Secure** and use it for placement, grouping, and support decisions without needing to administer additional assessments.

- Explore a **"Test Authoring" tool** that can allow educators to mix and match existing IAB/F-IAB test items to create their custom assessments that more closely match their curricula/instructional units.

Conclusion

One assessment cannot serve many purposes. For more than two decades, experts have called for a balanced system "where assessments at all levels – from classroom to state – work together in a system that is comprehensive, coherent, and continuous." Unfortunately, state accountability tests have led to more of the same type of broad tests locally. Connecticut has taken many steps toward realizing the vision of a balanced assessment system, but critical work remains. The shift away from broad assessments to focused classroom assessments at the district-level will require a sustained professional learning program, updated policy guidance, successful case studies, incentives, and adjustments and clear communication about the stakes associated with the state's accountability assessment.

Introduction

[Section 1 of Public Act No. 24-93](#) requires the Connecticut State Department of Education (CSDE), in consultation with national assessment experts and public school educators, to conduct a comprehensive audit of the assessments that are administered to students. The statute requires the CSDE to evaluate the assessments inventoried with the goals of eliminating redundant assessments, discouraging classroom activities that focus only on test preparation, reducing testing time, and maximizing assessments that provide actionable information for classroom teachers. The statute also charges the CSDE to develop and implement a program of professional learning for teachers concerning assessment literacy.

In planning for the audit, the CSDE examined assessment inventories developed by national organizations such as Achieve and WestEd and reviewed other state-level efforts in places such as Ohio and Oregon. There was tremendous overlap across the instruments. The CSDE synthesized the information and created an initial audit tool using Excel to be reviewed by Connecticut's Accountability Advisory Committee, a group comprised of district and school leaders who advise the CSDE on assessment and accountability issues. Based on committee feedback, the tool was revised to streamline the questions, ensure clarity to the greatest extent possible, and reduce burden for respondents. Finally, the CSDE consulted with national experts from the National Center for the Improvement of Educational Assessment ([NCIEA](#)) who have done this work in other states to finalize the audit tool.

On January 10, 2025, the CSDE provided all superintendents and district-level test administrators (DAs) with the [audit tool](#) and requested submission by February 21, 2025. Agency staff held a [virtual information session](#) for DAs on January 21, 2025, to clarify the instructions for the audit and answer questions from the districts. The session was [recorded](#) and posted on the agency's YouTube channel. Additionally, the agency developed a [Frequently Asked Questions](#) document based on the questions fielded during the information session and follow-up inquiries. All materials were posted on an [Assessment Audit](#) web page on the CSDE website (see Appendix D).

District representatives were instructed to report all core academic assessments required by the district or school to be administered without modification to some or all students in a standardized manner. Core academic assessments were defined as ELA, math, science, social studies, and related achievement/ability assessments including:

- Vendor-created assessments administered district/school-wide to some or all grades, or used for progress monitoring;
- Locally created assessments that are required to be administered without modification to some or all grades and some/all subjects;
- Smarter Balanced/NGSS Interim assessments that are required to be administered in a standard method; and
- Vendor created K3 benchmark assessments (over and above what is required by the CSDE) that are administered district/school wide.

The audit did not include an inventory of teacher-written or teacher-modified assessments given the anticipated data collection burden this would have created for districts and schools. The data

collection also did not include any of the required state assessments because the CSDE has that information including the amount of administration time necessary for each program.

Background

One Assessment Cannot Serve Many Purposes

The Joint Standards for educational and psychological testing define validity as "the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests." In other words, validity is intricately tied to how the results from the assessment are interpreted and used. There are many reasons why students are assessed and how those results are used. A teacher may want to know if their students can isolate, blend, segment, add, delete, and substitute phonemes; a district/school leader may want to know who should be identified as gifted/talented; the Connecticut State Department of Education (CSDE), the State Board of Education (SBE), legislators, and other community stakeholders may want to know which districts consistently reflect low student achievement statewide and who is exceeding expectations. **One single assessment cannot serve many purposes.**

An Important Historical Perspective on Educational Assessment

A recent publication by the National Academy of Education titled *Reimaging Balanced Assessment Systems*[vi] offers a rich historical perspective on the evolution of educational assessment practices. Knowing this history is important to understand our present state.

The 1983 seminal publication, *A Nation at Risk: The Imperative for Educational Reform* was one of the earliest publications to raise concerns about the academic achievement of America's children. Several national initiatives culminated in the 2001 bi-partisan passage of the federal No Child Left Behind (NCLB) act which instituted a new mandate of annual large-scale standardized assessment as an educational accountability measure.

In the same year that NCLB passed (i.e., 2001), the National Research Council published *Knowing What Students Know*[vii]. This publication emphasized important principles of assessment and articulated a new vision for a "balanced" system of assessments:

- *"Often a single assessment is used for multiple purposes; in general, however, the more purposes a single assessment aims to serve, the more each purpose will be compromised."*
- *"Large-scale assessments not only serve as a means for reporting on student achievement, but also reflect aspects of academic competence societies consider worthy of recognition and reward."*
- *"The power of classroom assessment resides in its close connections to instruction and teachers' knowledge of their students' instructional histories. Large-scale, standardized assessments can communicate across time and place, but ... they often have limited utility in the classroom. Thus, the contrast between classroom and large-scale assessments arises from the different purposes they serve and contexts in which they are used. Certain trade-offs are an inescapable aspect of assessment design."*
- *"...large-scale assessments can provide worthwhile targets for educators and students to pursue. Whereas teaching directly to the items on a test is not desirable, teaching to the*

theory of cognition and learning that underlies an assessment can provide positive direction for instruction."

- *"A vision for the future is that assessments at all levels—from classroom to state—will work together in a system that is comprehensive, coherent, and continuous."*

Pressure to Improve State Test Scores Led to More "Broad" Testing

Despite the 2001 vision for a coherent and aligned system of assessments at the classroom and state levels, what proliferated after the passage of NCLB was not a system of different assessment serving different purposes, but redundant testing in the form of mini-summative benchmark assessments. Many were commercial products while others were district created.

Such tests tried to mimic the state assessment, promoted test-prep/practice, and sought to "predict" performance on the state measure. They were not designed to foster **formative assessment practices**[viii]**,** which is an evidence-based strategy for improving student learning because it changes how teachers and students engage with assessment during the teaching and learning process.

This cycle re-occurred under ESSA when Connecticut's switch to the SAT as its high school accountability assessment led to new widespread PSAT 8/9 testing in Grade 9 and even some limited use in Grade 8, even though multiple CSDE studies[ix] show that middle school Smarter Balanced results are strong predictors of high school outcomes, rendering the PSAT redundant.
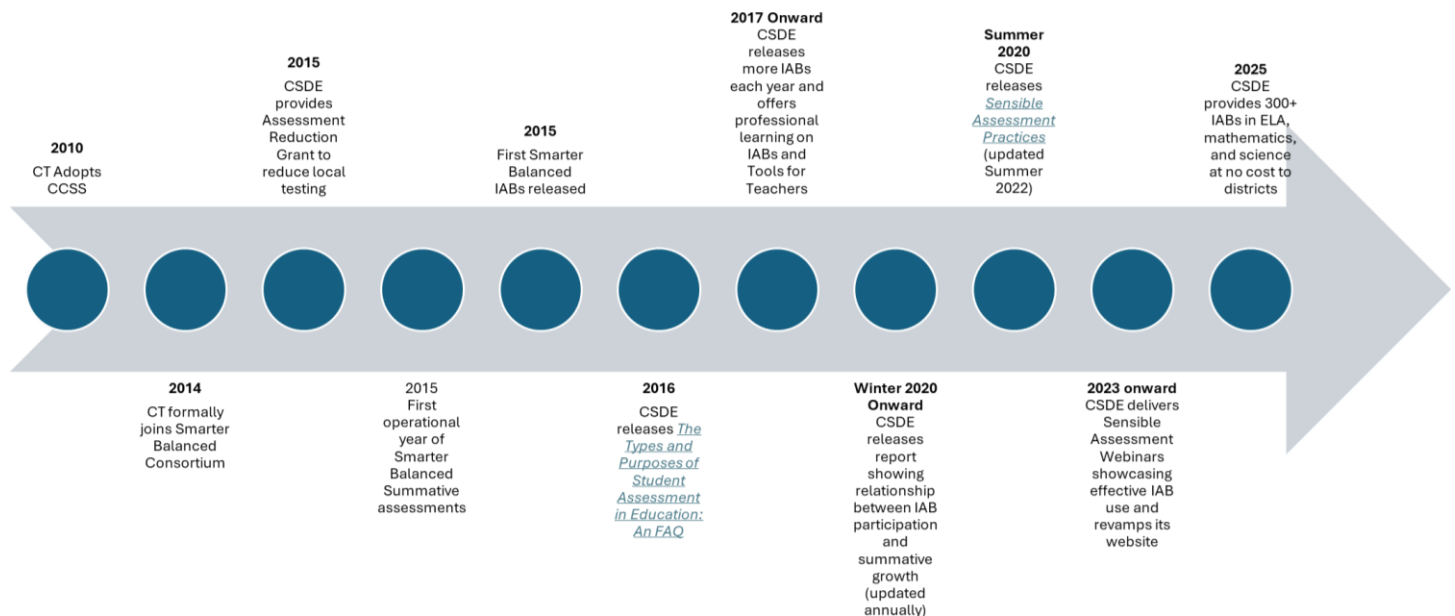
Toward a Balanced Assessment System

Connecticut adopted the Connecticut Core State Standards in 2010. The adoption of new content standards in English language arts (ELA) and mathematics required significant changes to the state assessment to ensure alignment with the new content standards.

When Connecticut adopted the Common Core State Standards, it also joined the Smarter Balanced consortium of states primarily because of its commitment to offer a "balanced" system of assessments. In addition to the end-of-year summative assessment in English language arts (ELA) and mathematics that would be aligned to the new standards, the consortium offered interim assessment blocks (IABs) to evaluate student learning, and resources and guidance to support teachers' formative assessment practices in the classroom. In such a balanced assessment system, **different assessments and resources would serve different purposes**:

- The broad **end-of-year state summative** assessment would provide a high-quality, valid measure of a student's overall achievement on the standards, as well as growth on those standards from the prior year. It would be valid to use it for evaluation and accountability purposes but would *not* be suitable for instructional uses.
- Short **interim block assessments** would be of much finer grain than the summative and assess narrow domains of content that align well with 1-3 weeks of classroom instruction. They would be more sensitive to student learning on those instructional units, while remaining tightly aligned with the summative assessment targets.
- **Formative assessment tools and resources** would be aligned to the assessment targets and interim results. They would lend themselves to be used by "teachers during learning and teaching to elicit and use evidence of student learning to improve student

understanding of intended disciplinary learning outcomes and support students to become self-directed learners."

Figure 1: CSDE's Timeline for Implementing a Comprehensive, Balanced Assessment System

Over the years, Smarter Balanced has released many short interim assessments.

- In ELA, there are 19 IABs/F-IABs in each grade from 3 through 8, and high school for a total of 133 ELA assessments. Together, these assessments include a combined 1,477 ELA test items.
- In math, educators can access anywhere from 10 to 16 short interim assessments per grade for a total of 94 math assessments that include a combined 1,202 math test items.
- Every interim assessment item is developed using the same comprehensive, and rigorous vetting process applied to summative test items.
- These online assessments are scored immediately. The assessment delivery platform allows educators to view student responses to individual test items. Teachers can see the test item, the underlying assessment target and standard, and the item's difficulty level. They can also be used in non-standard and formative ways.
- In addition to the ELA and Math interim assessments, the CSDE also offers 116 short interim science assessments aligned to the Next Generation Science Standards (NGSS). All these assessments are available to Connecticut educators at no cost.

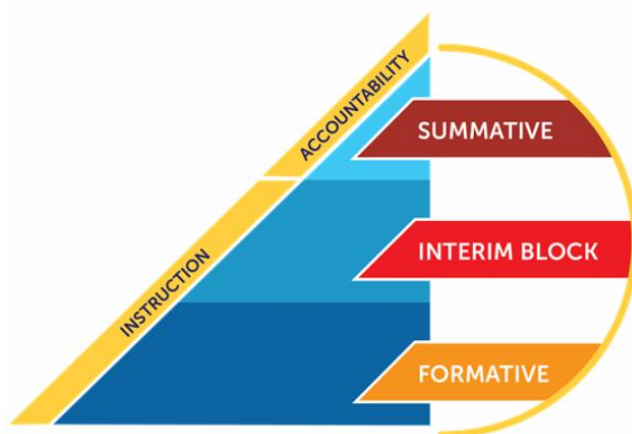## Curriculum, Assessment, and Professional Supports

As early as the 2014-15 school year, the CSDE recognized that local assessment practices require scrutiny to avoid over testing and ensure usefulness. To encourage local districts to conduct a careful review of assessments used throughout the year, the CSDE announced an *Assessment*

Reduction Grant. Districts could apply for the grant designed to eliminate tests that were outdated or did not support improving student learning. District applications explained how they would conduct a comprehensive inventory; engage educators, parents and the community in the evaluation and reduction process; and provide professional learning to assist educators in understanding the characteristics of high-quality assessments that can inform instruction. The stated goal was to limit the inclusion of unnecessary assessments.

Since that time, the CSDE has urged district and school leaders to carefully examine the types and purposes of assessment. In 2016, the CSDE issued *The Types and Purposes of Student Assessment in Education: An FAQ* to clarify that the state assessment is a big picture/high-level measure that is best designed for evaluation and accountability, and not for supporting the instructional needs during the year, and that a balanced assessment approach is necessary to meet all needs.

The CSDE subsequently rebranded the Student Assessment section of the website to clearly communicate the importance of a **balanced assessment system** that ensures different types of assessments are used more specifically for the purposes for which they are designed.
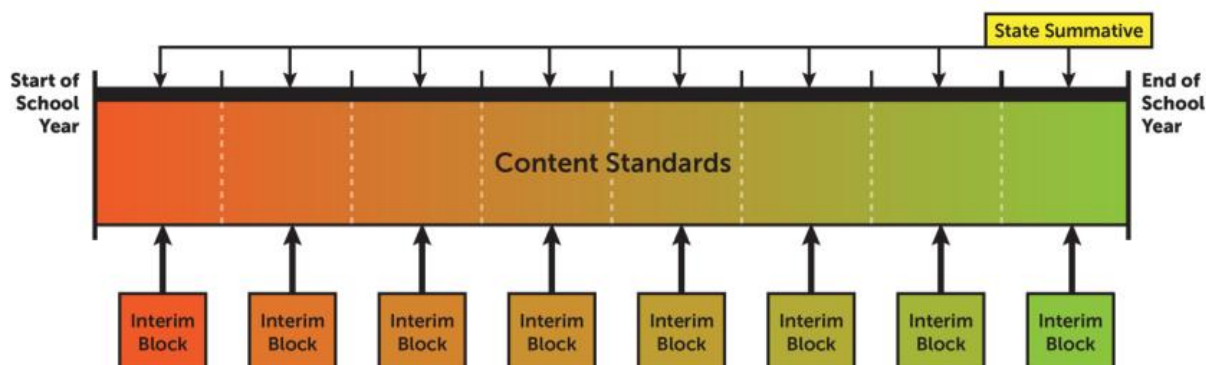
Figure 2: A Balanced Assessment System Framework

This image is an adaptation of a graphic originally developed by Perie, Marion, Gong, and Wurtzel (2007)

For example, the summative Smarter Balanced assessments are designed to evaluate overall academic achievement at the end of the year and inform accountability, evaluation, and support. They are suitable for use in district and state accountability systems, program/curriculum evaluations, and district/school identification for support and recognition. Preliminary results from the summative Smarter Balanced and NGSS assessments are also released starting mid-May. These preliminary results are available either immediately or within days of testing, so teachers can review the performance of their students before the end of the school year in order to support them for transition to the next grade. District and school team are also able to use these results to evaluate their year and plan for the upcoming year.

Unlike the state summative assessments, IABs/F-IABs are short assessments that help teachers check student progress during the year, gather information about learning, and alter upcoming curriculum plans for future units.

In Summer 2020, the CSDE published the first edition of *Sensible Assessment Practices* (updated in Summer 2022), which provides a set of recommended actions to consider at key points in time throughout the school year and suggests leveraging existing data sources to be used in lieu of additional testing. The publication also clarified the difference in content coverage between the end-of-year summative assessment and the IABs (Figure 3).

Figure 3: Content Coverage of State Summative Assessment vs IABs

The CSDE has supplemented these publications with ongoing professional learning and customized technical assistance. For example, here are some sessions offered over the past two years as part of the Sensible Assessment Practices webinar series.

- **Non-Standard Uses of Smarter Balanced/NGSS Interim Assessments and District Sharing**
- **Smarter Balanced Interims and District Sharing**
- **NGSS Interims and District Sharing**
- **Using Smarter Balanced/NGSS Interim Results and District Sharing**
- **Sensible Assessment Smarter Balanced Tools for Teachers - Back to Basics**
- **Tools for Teachers and District Sharing**

These CSDE publications and professional learning sessions review essential terminology and emphasize that a single test cannot serve multiple purposes. Educators make inferences about what students know and can do based on test scores. Every inference should have strong evidence to support it. This is referred to as "validity evidence" and is an important consideration when selecting assessments for use in school districts.

CSDE's Model ELA and Model Mathematics curricula that are available freely to districts integrates Smarter Balanced IABs/F-IABs thoughtfully and consistently across the grades.

Methods

The audit questions are listed in Appendix A. Most questions used dropdowns, where districts could select the best choice from a list of options, including "Other." The assessment audit forms were collected from districts and checked for completion.

Out of 199 districts, 175 received personalized follow-up to ensure compliance with the audit instructions. CSDE staff checked each audit submission to ensure

- All rows were complete
- Districts did not report state-required assessments
- Student counts were reasonable for the district
- The amount of testing time was reasonable for the assessment (for assessments with standardized assessment time)
- Every district is required to do progress monitoring as part of the SRBI/MTSS process. In cases where a district did not report any assessments for used for progress monitoring for students in intervention, the district was asked to supply a narrative describing the progress monitoring process for students in the district.

Districts were instructed to include the designated time for administering the test. To the extent that time is used for test set up, instructions, or breaks from testing, that time is *not* included in the audit. Thus, the actual time dedicated to testing may be greater than the amount reported on the audit.

Assessment Terminology for Data Collection

The assessment audit data collection tool required districts to report a primary purpose for each assessment. Purpose is a key variable to ensure that each assessment administered to students aligns with its stated primary purpose.

Districts categorized assessments into the following purposes and were provided with the definitions listed below.

- Summative

  o An assessment administered at the end of instruction to measure whether students have learned what was expected to be learned.

- Interim/Benchmark

  o An assessment typically administered during the year to all students to measure general progress

- Universal Screening

  o An assessment typically administered to all students to identify those who may need additional intervention

- Progress Monitoring

     o Assessments administered for students in intervention to see if they are making progress

- Diagnostic
  - Assessments administered to diagnose specific strengths/weaknesses or make a particular diagnosis (e.g., identify risk for dyslexia)
- Other

Summative assessments should be administered less frequently than interim/benchmark assessments because the former measures overall learning at the end of instruction while the latter is meant to measure students' knowledge and skills ideally relative to a specific domain or sub-skill (e.g., "Read Informational Text" or "Numbers and Operations in Base Ten"). Summative assessments, interim/benchmark assessments, and universal screening assessments are typically administered to all students. Progress monitoring and diagnostic assessments should be reserved for groups of students. They are not intended to be used with all students.

Assessments not Included in the Audit

Districts were instructed not to include any state-required assessments in the data collection tool. Additional information about the time spent on these assessments is included in the results section. In addition, many students elect to take Advanced Placement or International Baccalaureate exams. These exams were not included in the audit. Finally, students who strive to earn the Seal of Biliteracy prior to graduation must demonstrate their language proficiency through an exam. These data were not included as part of the audit because the CSDE already collects this information through other sources.

Assessments outside of the core academic areas were not included in the audit. These include but are not limited to school climate assessments, physical fitness or arts assessments, and assessments of social and emotional competencies such as the Devereux Student Strengths Assessment (DESSA). Individualized testing after referral for special education services is not included in the audit, but districts were advised to report screening assessments administered to a group of students for potential referral to special education.

The CSDE provided guidance that project-based assessments spanning multiple class sessions where students work under the guidance or oversight of a teacher, either individually or in small groups, to investigate an issue and develop presentations, reports or other products should not be reported. This would include capstone projects that may be evaluated using a rubric.

Districts were also instructed not to include any teacher-written or teacher-modified assessments. These assessments are part of the regular teaching and learning cycle. CSDE acknowledges that a significant amount of instructional time may be spent on the formative assessment process, short quizzes, and unit tests; that time is not reflected in this audit.

Average Time Spent Testing

To calculate the average amount of time spent tested, the total number of students tested each time was multiplied by the number of times the test was administered and the time in minutes of each administration, then divided by the number of students enrolled in the district (see formula 1).

The official enrollment counts come from the October Public School Information System (PSIS) collection, which reflects the school enrollment on October 1, 2024. For example, if a district reported that 10 students enrolled in intervention took a 15-minute test 8 times a year, then the total number of minutes spent testing by those students is 10*15*8 = 1,200. If there are 100 students in the district, then the average time spent testing is 12 minutes per year. In cases where districts reported multiple grades on one line, the total number of students assessed was divided equally among the grades, and each grade was capped at the number of enrolled students.

Formula 1: Average Testing Time

$$Average\ Testing\ Time = \frac{Tested\ Students * Times\ administered * Minutes\ spent\ testing}{Number\ of\ students\ Enrolled}$$

Average time spent testing was calculated for each district individually, and for the whole state. Averages were calculated by several analysis categories: Grade, Alliance District, and Content Area. Averages were calculated at the state level and for each district; then the 25th and 75th percentile of (tested) district averages were calculated. The 25th and 75th percentiles indicate the percentiles for average time spent testing for students by districts who did some testing – districts who did not test in a category are not included in the percentile calculation. The 25th percentile indicates the amount of testing where 25% of districts tested *less* than that amount. The 75th percentile indicates the amount of testing where 25% of districts tested *more* than that amount. As such, the 25th to 75th percentile range gives a picture of what the middle 50% of districts are doing. See the charts in figures 1 through 3, where the blue bars indicate the state-wide average, and the orange lines indicate the 25th to 75th percentile of district averages.

Results

State Academic Assessments Average Time

State-required academic assessments are administered for accountability in grades 3-8 and 11 in ELA and Math, and grades 5, 8, and 11 in science. In addition, English/multilingual learners will take the LAS Links English Proficiency Assessment.

Table 1: State Academic Assessment Average Testing Time

Assessment	Approximate Average Testing Time
ELA Grades 3-8 (Smarter Balanced)	1.7 hours
Math Grades 3-8 (Smarter Balanced)	2.1 hours
Science Grade 5 (NGSS)	1.6 hours
Science Grade 8 (NGSS)	1.5 hours
ELA Grade 11 (SAT)	1.1 hours
Math Grade 11 (SAT)	1.2 hours
Science Grade 11 (NGSS)	1.1 hours

In addition, all students in grades Kindergarten through grade 3 must take a literacy universal screening assessment three times per year; districts can select from a menu of assessments.

District Required Testing: Assessments and Average Time

Districts reported using a variety of assessments. The assessment audit form had a list of commonly used assessments, though districts could also select "other" or "Locally developed". A selection of other assessments used are listed in Appendix B.

Most Used Assessments

Table 2 shows the most used assessments, the number of districts that use that assessment, and a range of time reported spent on that assessment. The reported range is the interquartile range (middle 50%) for districts that required the assessment for all students in at least one grade. This represents a reasonable range for the amount of time a district might spend on the assessment if they are choosing to use the assessment. The same assessment may take more or less time depending on the number of times administered and how long students take to complete the assessment. Some districts appear to spend more time on locally developed assessments which can represent midterms and finals.

Each vendor was categorized by whether it is a broad assessment or a focused assessment. Note that some vendors offer multiple formats for assessment. Broad assessments have a large scope, drawing from multiple content strands – sometimes spanning across grade levels. Conversely, focused assessments have a smaller scope. Focused assessments pinpoint whether a student has learned a specific skill; they are better for assessing the content being taught at the moment and consequently better suited to help teachers in day-to-day instruction.

Note that, in general, the focused assessments take less time to administer. This is a combination of the focused assessments taking less time per administration and being used fewer times per

year. When an assessment has a wider range of assessment time, it reflects that some districts administer the assessment more times per year than others.

Table 2: Most Used Assessments for District- and School-Required Assessments
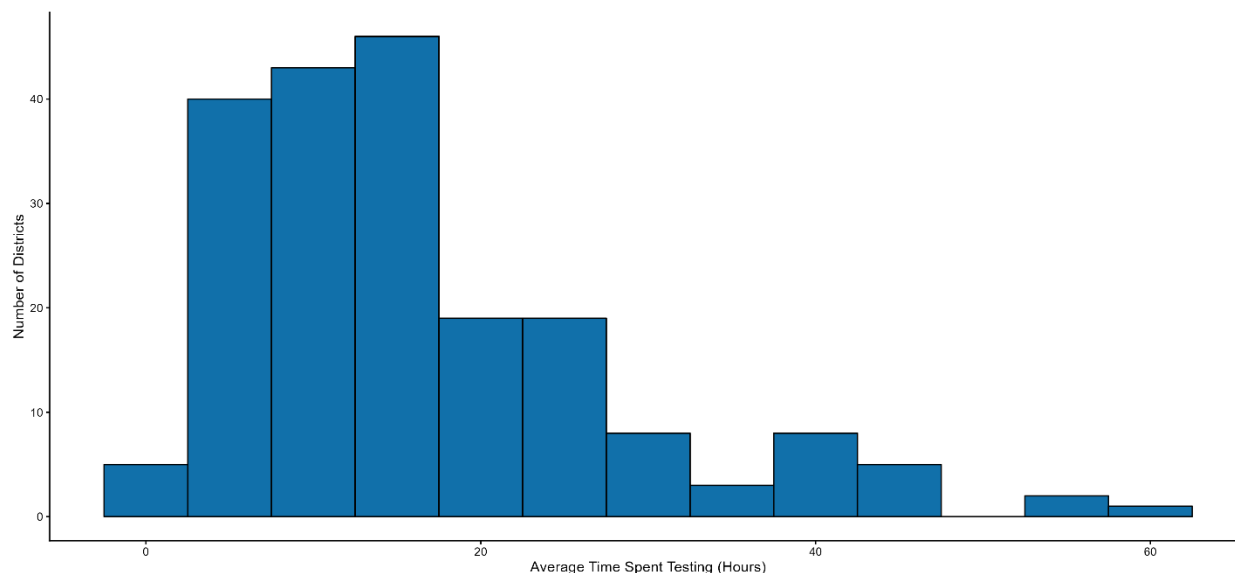
Assessment / Vendor	Subject	Number of Districts	Inter quartile range of assessment time per year for tested students	Assessment Scope
Acadience	ELA and Math	31	45 minutes to 1.5 hours	Focused
Aimsweb	ELA and Math	40	1.5 to 4.5 hours	Focused
DIBELS	ELA	127	15 minutes to 1 hour	Focused
IXL	ELA and Math	41	2 to 4.5 hours	Broad
NGSS Interim Assessments	Science	88	1 to 2 hours	Focused
NWEA	ELA and Math	48	2.5 to 7.5 hours	Broad
PSAT	ELA and Math	129	3 to 4 hours	Broad
STAR/Renaissance Learning	ELA and Math	47	3 to 4.5 hours	Broad
Smarter Balanced Interim Assessments	ELA and Math	137	3 to 6 hours	Focused
I-Ready	ELA and Math	85	4 to 9 hours	Broad
Other Vendor	-	175	3 to 14 hours	-
Locally Developed	-	108	3 to 17 hours	-

Total Average Time Spent

Districts had wide variation in how much time was spent on district-required testing overall. Figure 4 shows a histogram of districts based on the district-level average time spent testing. Most districts spent under 20 hours testing (on average) during the year. With 900 hours of actual school work required annually per state law, this represents about 2 percent of that total time.

The districts that spent significantly more time usually had very detailed assessment audit forms, with multiple different assessments used at each grade level. Spending a large amount of time on district-required assessments could reflect a systematic approach to assessment, where the district has agreed-upon assessments used at each grade level to standardize the teaching and learning experience. Conversely, they may represent opportunities for reducing assessments with overlapping content (e.g., curriculum embedded assessments and state-provided IABs/F-IABs) or reducing the frequency of assessments (e.g., only administering a winter benchmark with an off-the-shelf assessment instead of fall, winter, and spring).

By Grade level

Figure 5 shows the average amount of time spent on district required testing by grade level group. Pre-Kindergarten students spend very little time engaged in district- or school-required tests. Students in kindergarten through 11th grade spend the most time testing. Twelfth-grade students spend less time testing, though there is considerable variation between districts. In particular,

- **Pre-kindergarten students** spent .8 hours testing, with the 75th percentile at 3.6 hours. Only 52 districts did any standardized testing for pre-kindergarten students.
- **Kindergarten through Grade 2 students** spent 14.8 hours testing, with the middle 50% ranging from 7.7 hours to 20.3 hours.
- **Grade 3 to 5 students** spent 20.2 hours testing, with the middle 50% of districts ranging from 11.6 to 25.3.
- **Grade 6 to 8 students** spent 21.5 hours testing, with the middle 50% of districts ranging from 9.5 hours to 26 hours.
- **Grade 9 to 11 students** spent 19.2 hours testing, with the middle 50% of districts ranging from 4.8 hours through 26.3 hours.
- **Grade 12 students** spent 10.7 hours testing with the middle 50% of districts ranging from 4.4 hours to 24.9 hours.

Further examination of the high school grades showed significant variation due to whether districts reported midterm and final exams on the audit form. Districts were instructed to only report exams that were given in a standardized manner where teachers cannot change the content of the test. Thus, those districts that have standard course exams, even if written by the teachers, appear to have more testing in grades 9 to 12, compared to districts that have midterm and final exam blocks but do not provide their teachers with common assessments. For students in grades 9-12, 54 districts reported administering standardized midterm or final assessments. For those districts, students spent between 5 and 15 hours on midterms and finals, with some students spending as much as 35 hours on midterms and finals.

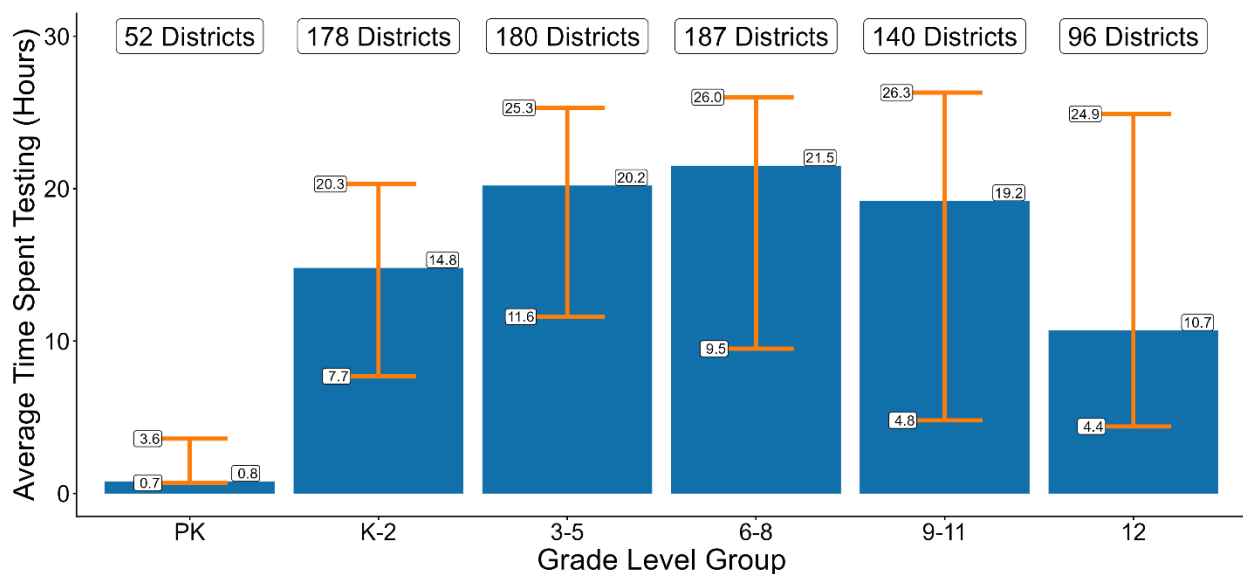*Figure 5: Time Spent testing by grade level group*

Figure 5 Note: The blue bars indicate the *state-level* average time spent testing. The orange lines indicate the middle 50% of *district-level* weighted average time spent testing for districts that test in that grade range.

By Content Area

Every district reported at least some district-required testing in ELA and almost every district reported some district-required testing in Mathematics. Those two subjects also reported the highest average testing time. Much less testing time is spent on science, social studies, and other content areas, though many districts do spend some time testing in those subjects. See Figure 6. In particular:

- **In ELA,** students spent 6.9 hours on testing, with the middle 50% of districts ranging from 3.4 to 8.5 hours.
- **In Mathematics,** students spent 6.7 hours on testing, with the middle 50% of districts ranging from 3.2 hours to 8.5 hours.
- **In Science,** students spent 2.1 hours testing, with the middle 50% ranging from .7 hours to 3.5 hours.
- **In Social Studies,** students spent 1.2 hours testing, with the middle 50% ranging from 0.9 hours to 2.8 hours.
- **In other subjects,** students spent 0.7 hours testing, with the middle 50% of districts ranging from 0.1 to 1.3.

It is not surprising that ELA and mathematics consume more time testing than science and social studies, as those are the focus of annual state-level testing and accountability. However, this highlights an opportunity for a reduction in time spent testing.
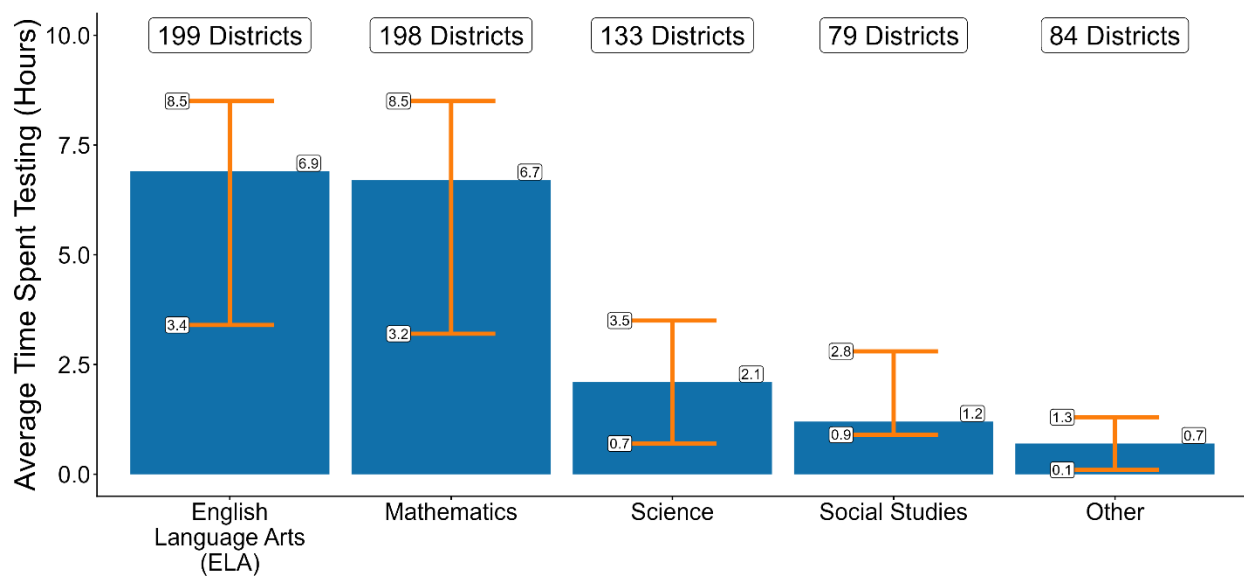
Figure 6: Time Spent Testing by Content Area

Figure 6 Note: The blue bars indicate the *state-level* average time spent testing. The orange lines indicate the middle 50% of *district-level* weighted average time spent testing for districts that test in that subject.
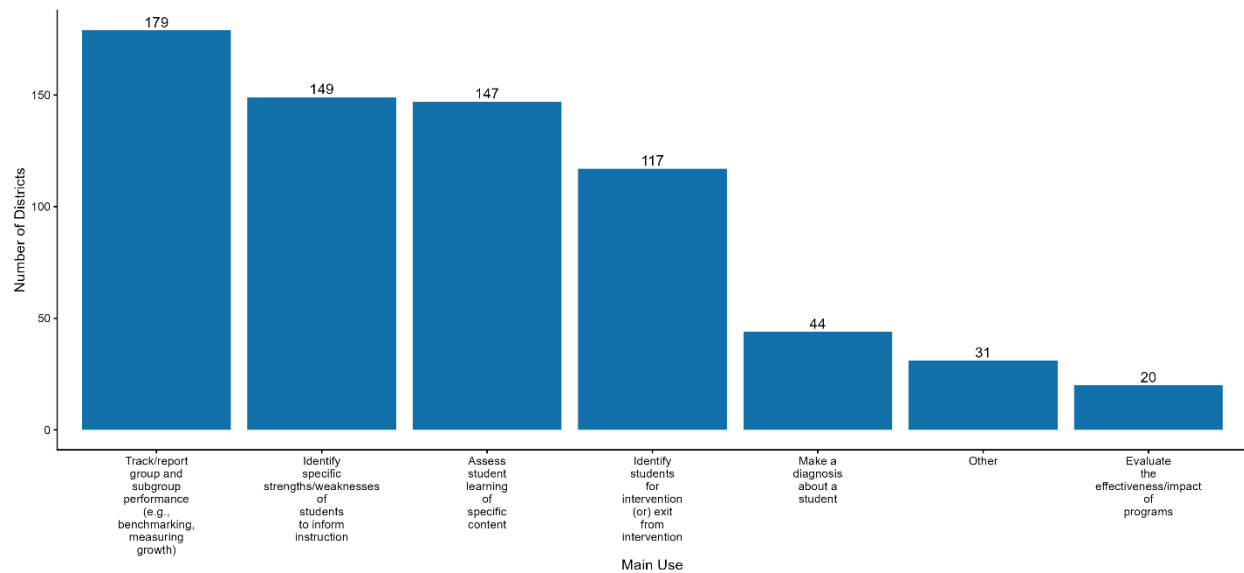
District-Required Assessments and their Reported Main Use

The assessment audit form asked for the main ways each assessment was used by the district and/or school. Districts selected one main use and could select one secondary use from the list below.

- Assess student learning of specific content
- Evaluate the effectiveness/impact of programs
- Identify specific strengths/weaknesses of students to inform instruction
- Identify students for intervention (or) exit from intervention
- Make a diagnosis about a student
- Track/Report group and subgroup performance (e.g. benchmarking, measuring growth)
- Other

Figure 7 shows the number of districts who reported using at least one assessment for each main use. The most common use for assessments was *Track/report group and subgroup performance*, with 179 districts reporting at least one assessment used for this purpose. The next three most common uses were all about supporting teaching and learning; Identify Specific Strengths/Weaknesses of Students to Inform Instruction (149 districts), Assess Student Learning of Specific Content (147 districts), and Identify Students for Intervention (or) Exit from Intervention (117). Relatively few districts reported assessing to Make a Diagnosis About a Student (44), Evaluate the Effectiveness/Impact of Programs (20), or Other (31).

Figure 7: Number of Districts Reporting at Least One Assessment by Main Use
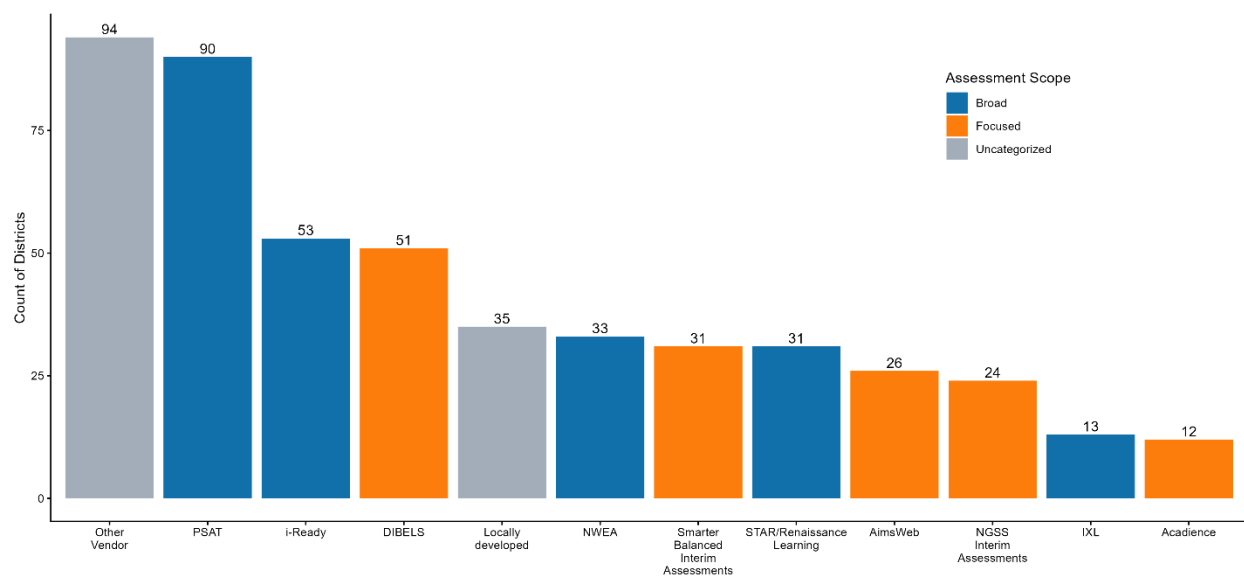

Figure 7: Number of Districts Reporting at Least One Assessment by Main Use

Main Use: Track/report Group and Subgroup Performance

Figure 8 shows the vendors that districts selected for the main use of *Track/report group and subgroup performance,* colored by assessment scope (i.e. Broad, Focused, or Uncategorized).

Figure 8: Count of Districts Using Each Vendor for Track/report Group and Subgroup Performance

PSAT makes up the greatest single assessment used for *Track/report Group and Subgroup Performance* with almost half of districts using the PSAT for this purpose. A variety of other

assessments are also used to *Track/report Group and Subgroup Performance*. PSAT is most commonly used in grades 9 and 10.
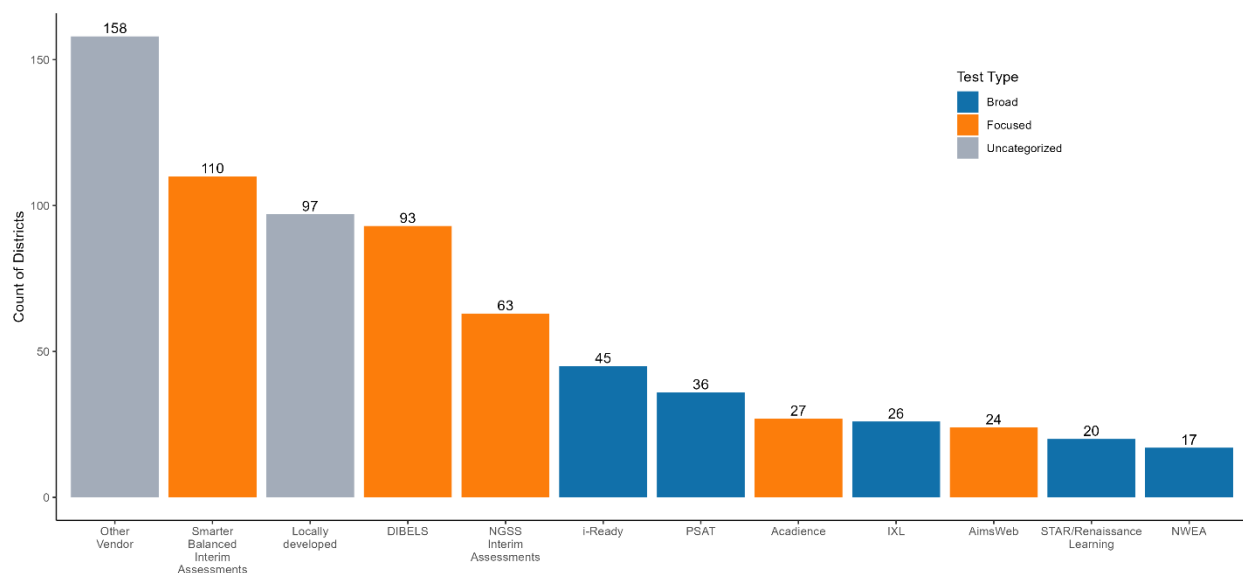
Figure 8 shows that broad assessments are more likely to be used for tracking and reporting subgroup performance than focused assessments.

Main use: Teaching and Learning

The three main uses focused on instruction were also commonly reported: Assess Student Learning of Specific Content, Identify Specific Strengths/Weaknesses to Inform Instruction, and Identify Students for Intervention (or) Exit from Intervention. Figure 9 shows the vendors selected for these three main uses colored by assessment scope.

Focused assessments are more commonly used to support teaching and learning compared to broad assessments. In addition, locally developed assessments are much more prevalent in supporting teaching and learning as compared to tracking and reporting subgroup performance. However, it is still common to use broad assessments to support teaching and learning. These assessments measure what should be learned in a semester or the entire year; they are ***not*** designed to measure student learning after a shorter unit/lesson that may last 1-3 weeks. As such, their instructional value is limited.

Figure 9: Count of Districts using Each Vendor for Main Uses Related to Teaching and Learning

Other Assessments Reported by Districts

Mis-aligned Literacy Screeners

Connecticut General Statutes Section 10-14ii requires the CSDE to review and approve comprehensive reading curriculum models or programs that are required to be implemented by all public schools. These curriculum models or programs are evidence-based and aligned to the "science of reading." C.G.S Section 10-14t also requires districts to select a K-3 literacy universal

screening assessment from a list compiled by the CSDE. Districts must then administer that screener to all students in Grade K-3 three times a year to identify students who are below proficiency in reading, at risk for reading difficulties and require intervention, and assist in identifying, in whole or in part, students at risk for dyslexia, or other reading-related learning disabilities.

Based on information submitted through the audit process, a small number of districts report using additional assessments in K-3 that are not approved by the CSDE as a universal screener and do not align with Connecticut's approved list of K-3 reading curriculum models and programs. Districts using these systems should discontinue implementation of such mis-aligned assessments and instead devote efforts to offer professional learning so staff can be better prepared to use and support aligned assessments.

Gifted and Talented Assessments

Under current regulations, public school districts in Connecticut are required to identify gifted and talented students, K-12. Although school districts are not mandated to provide services, many schools offer programs at some grade levels. The process for identifying students who are gifted and talented is flexible and determined by district personnel. Thirty-eight districts reported a gifted identification test or screener administered to some or all students within a grade. These assessments take between 5 minutes and three hours and average about an hour in length.

Mock or Practice Assessments

Eleven districts reported implementing "Mock" or "Practice" assessments for AP or SAT exams. For the districts that administered those assessments, students spent between 1 and 5 hours on mock exams. CSDE recommends limiting the activities that are solely devoted to test preparation at the expense of instructional time.

Discussion

The axiom "One Assessment Cannot Serve Many Purposes" was discussed extensively more than two decades ago in the National Research Council's publication, *Knowing What Students Know*. It spoke of the power of classroom assessment to inform teacher knowledge of student learning given its proximity to the classroom. It contrasted that with the important purpose of large-scale state assessments that report on student achievement and growth and serve as "north star" for what matters. The report imagined a vision where *"assessments at all levels—from classroom to state— will work together in a system that is comprehensive, coherent, and continuous."*

In the decade immediately after NCLB's passage, what emerged however was not a balanced system of assessments as was envisioned by *Knowing What Students Know*. Instead, as pressure rose on educators to improve student performance on test scores, there emerged an increased use of large-scale, broad commercial tests **during the year.** The theory-of-action was that to improve performance on the state assessments, one needs to periodically "benchmark" student performance on measures *like* the state assessment, track "growth" from fall to spring on that measure, and use the benchmark results to "predict" performance on the state summative. As more resources were directed toward broad, large-scale assessments, minimal attention and resources were devoted to improving classroom assessment and formative assessment practices. Unfortunately, broad assessments draw from a wide variety of content strands, tend to be

computer adaptive, and cannot be used to reliably determine student mastery of specific content. The past 20 years have unequivocally demonstrated that more broad "practice" testing in and of itself does not improve student achievement.

This cycle re-occurred under ESSA when Connecticut, in Spring 2016, switched from administering the Smarter Balanced assessment in Grade 11 to using the SAT as its high school accountability assessment. The primary rationale for this switch was that 70 percent of students were already taking the SAT and that administering the SAT would reduce double-testing. Ironically, however, the opposite occurred as this decision led to more large-scale standardized testing in the high school grades. As the stakes on SAT performance increased, a new assessment – the PSAT 8/9 – began to be administered to students in the earlier grades. The assessment audit submissions show that 101 districts now administer the PSAT 8/9 in the fall of Grade 9 – a mere 3-4 months after the students had already taken the Smarter Balanced Grade 8 assessment even though multiple CSDE studies[x] show that middle school Smarter Balanced results are strong predictors of high school outcomes. Another 19 districts even administer the PSAT 8/9 in Grade 8, thus unnecessarily double-testing the student. Often a whole day of high school instruction is lost to the administration of the PSAT. Additionally, 36 districts claim to use PSAT results to *Assess Student Learning of Specific Content* or *Identify Specific Strengths/Weaknesses*. The broad content of PSAT is not designed for such uses.

Another purported use of broad, large-scale benchmark testing is to measure progress during the year from fall to spring. However, researchers have questioned the veracity of this practice. One researcher[xi] examining "summer loss" showed that *"the strongest predictor of whether a student would experience summer gains or losses was the size of gain the student had made during the previous academic year. That is to say, the more students learned during the school year, the more likely they were to lose ground during summer break. Knowing how much a student gained in the prior year alone explained between 22 and 39% of the variation in summer learning patterns."* Another study demonstrated that this observed phenomenon of depressed fall scores *"is not surprising, but another example of a famous statistical artifact: regression to the mean."[xii]* Administering, scoring, tabulating, and reporting results from such off-the-shelf broad benchmark assessments takes valuable time away from classroom assessments and instruction. It should be noted that the CSDE's Smarter Balanced growth model measures growth from Spring to Spring i.e., from the end of one grade to the end of the next grade, providing a more valid view of longitudinal student growth.

Connecticut's adoption of the Common Core State Standards in 2010 and the implementation of the Smarter Balanced assessment system has considerably changed the breadth of available tools and resources. The "smarter" and "balanced" system offers **different assessments and resources for different purposes**. The broad **end-of-year state summative** assessment is short, computer-adaptive and provides a high-quality measure of achievement and growth on the state standards; it is a valid measure to use for evaluation and accountability. The system also offers over 200 **IABs/F-IABs** that are focused, fixed-form tests and are designed explicitly to measure student learning on short instructional units. They are tightly aligned with the summative assessment targets and built to the same rigor and quality. CSDE's Model ELA and Model Mathematics curricula that are available freely to districts integrates Smarter Balanced IABs/F-IABs thoughtfully and consistently

across the grades. **Formative assessment tools and resources** provide teacher-created, standards-aligned, and vetted lessons that incorporate the formative assessment process.

The CSDE's Sensible Assessment Practices guidance urges school leaders to leverage information they already have about students rather than administering additional tests. CSDE's EdSight Secure platform provides over 4,000 district/school leaders and coordinators across Connecticut with near real-time historical data including prior test scores, attendance, disciplinary incidents, mobility, course grades, etc. as soon as the student is enrolled in their grade, school, or district from any public school in Connecticut. Since 2017, districts and schools have also had access to the Early Indication Tool (EIT) in EdSight Secure. EIT is a K-12 system that uses available data to identify students who may need additional support to reach academic milestones (e.g. reaching proficiency on state assessments, on-time graduation, demonstrating college/career readiness) and facilitates timelier, targeted interventions. The EIT uses statistical methods to assign a level of support to every student (i.e. low, medium, high). The longitudinal data, coupled with the EIT support level provide valuable and reliable information for immediate use, without needing to administer one more test.

The average testing time spent on the state academic accountability assessments is only around 5 hours annually. Students spend an additional 20 hours taking district-required assessments that are not modified by teachers. Some of these district-required assessments are focused assessments that can be useful to classroom teachers for instructional purposes; others are broad, large-scale tests that are benchmarking achievement and have limited instructional value. The primary use of assessment during the school year should be to support teaching and learning. When that focus is prioritized, it will be possible to eliminate redundant assessments and increase instructional time. While the pressure felt by educators from state tests is real, the state test is not requiring districts to repeatedly administer broad large-scale tests to benchmark growth during the year, and predict performance on the state test. Though the overall time spent on district-required testing is only around two to three percent of the total statutorily required minimum hours for actual school work, the audit findings reveal there is an opportunity to increase the use of classroom assessments and formative practices in lieu of broad large-scale tests. Such a change can redirect teacher energies toward greater application of those formative assessment practices in the classroom to truly change teaching and enhance student learning.

This shift toward focused classroom assessments and formative practices at the district-level will *not* happen by only tinkering with the state's requirements around the accountability assessments. Instead, a comprehensive, systemic approach is needed that is comprised of the following elements:

- a sustained program of professional learning and technical assistance for district teams to re-imagine their local assessment practices;
- updated policy guidance from CSDE that prioritizes focused classroom assessments and formatives practices, while giving permission to district leaders to discontinue broad, large-scale tests;
- successful case studies of school districts that have already made significant strides in this direction; and
- adjustments and clear communication about the stakes associated with the state's accountability assessment.

Recommendations

In light of the findings from this audit, the strategic actions described below can serve as a foundation for the continued work of the CSDE and for requisite legislative proposals that are designed to accomplish the goals of this audit. To recap, the goals of this audit were to eliminate redundant assessments, discourage test prep, reduce testing time, maximize assessments that provide actionable information for classroom teachers, and provide professional learning on assessment literacy.

The following specific strategic actions are recommended and can lead to more coherent local assessment systems that prioritize assessment **for** learning:

- Establish an **incentive program** for districts who can demonstrate that they have:
  - limited the time that students spend on broad tests during the year;
  - thoughtfully integrated the state-provided IABs/F-IABs and formative assessment tools into the local curriculum in a manner that supports ongoing instructional improvement; and
  - increased teacher competency in the formative assessment process.

  Incentive program awardees should be expected to serve as mentors for others looking to reduce assessment and increase local coherence.

- Issue **updated guidance and policy** that:
  - specifies the appropriate and inappropriate use of assessments;
  - discourages use of assessments solely for test prep;
  - directs discontinuance of assessments that are not aligned to the research;
  - recommends eliminating the fall and spring administration of broad local assessments to reduce redundancy and student testing fatigue;
  - illustrates the use of EdSight Secure longitudinal data (e.g., attendance, behavior, assessment, course grades, mobility) for placement/grouping in the new grade;
  - encourages use of state-provided IABs/F-IABs in place of broad assessments or other end-of-unit classroom summative assessments to evaluate learning during the year; and
  - demonstrates the formative use of IABs/F-IABs along with other state-provided formative tools to support instruction.

- In line with recommendations of the working group established pursuant to Public Act 24-45, explore the feasibility of and appropriate timing for seeking federal waiver to make **changes to the high school accountability model** by reducing the weight assigned to the high school assessment and increasing the weight for college- and career-oriented measures. This can also incentivize districts to reduce broad locally administered assessments. Specifically consider the following:
  - Reduce the weight for the high school assessment in the accountability index; the weight is currently at about 52% due to federal law and U.S. Department of Education requirements;
  - Incorporate new measures for career readiness within Indicator 6;

- o Increase the weight in other areas such as postsecondary readiness (Indicator 6) or on-track to high school graduation (Indicator 7);
  - o Reallocate funds that are currently supporting PSAT administration in Alliance Districts for expanding meaningful career-oriented programming/credential, and postsecondary partnerships.

- Provide **assessment literacy training and intensive coaching**.
  Note that the CSDE is contracting with assessment experts from the *National Center for the Improvement of Educational Assessment* – the preeminent national experts on assessment – to develop and implement a multi-year system of professional learning, coaching, and technical assistance. This project –branded as Student Centered Assessment for Learning (SCALE) – builds on these assessment audits to directly address findings related to inefficiencies, incoherence, and limited usefulness in current assessment practices. SCALE will develop teacher and leader capacity, and build regional coaching and leadership infrastructure through the regional educational service centers (RESCs).

  SCALE will directly address the nine threats to balanced assessment systems outlined in the Center for Assessment and National Academy of Education Practical Guidebook and listed in the table below. This Guidebook is a companion to the National Academy of Education volume, *Reimagining Balanced Assessment Systems*.

  Common Threats to Efficiency:

  Threat 1: Too much testing overall, particularly early or later in the year

  Threat 2: Redundant assessments

  Threat 3: Unused assessment results

  Common Threats to Usefulness:

  Threat 4: No clear match between the assessment purpose, design, and use

  Threat 5: Assuming all tests can inform instruction

  Common Threats to Coherence:

  Threat 6: Inconsistency between assessments and instructional vision

  Threat 7: Policies and politics that distort practice

  Threat 8: Over-emphasizing the role of summative assessment

  Threat 9: Under-emphasizing the role of formative assessment

- Offer professional learning opportunities so educators can increasingly leverage the **longitudinal student information in EdSight Secure** and use it for placement, grouping, and support decisions without needing to administer additional assessments.

- Explore a **"Test Authoring" tool** that can allow educators to mix and match existing IAB/F-IAB test items to create their custom assessments that more closely match their curricula/instructional units.

Conclusion

One assessment cannot serve many purposes. For more than two decades, experts have called for a balanced system "where assessments at all levels – from classroom to state – work together in a system that is comprehensive, coherent, and continuous." Unfortunately, state accountability tests have led to more of the same type of broad tests locally. Connecticut has taken many steps toward realizing the vision of a balanced assessment system, but critical work remains. The shift away from broad assessments to focused classroom assessments at the district-level will require a sustained professional learning program, updated policy guidance, successful case studies, incentives, and adjustments and clear communication about the stakes associated with the state's accountability assessment.

Appendix A: Assessment Audit Form

1a. Name Of Assessment/Vendor: Select from the dropdown list

        Options: Acadience, AimsWeb, Amira, DIBELS, DRA, easyCBM, FastBridge, Horizon, i-Ready, Imagine Language and Literacy (Imagine Learning), IXL, Lexia RAPID (Lexia Learning), NGSS interim Assessments, NWEA, PSAT, Smarter Balanced Interim Assessments, STAR/Renaissance Learning, Locally developed, Other

1b. If reporting an assessment as "other" in column A, provide more information here.

2. Grade(s)
Select Y for each grade where the test is administered. Leave all other grades blank.

3. Content Area: Select the content area assessed.
If an assessment measures more than one content area (e.g. ELA and mathematics), complete separate entries for each content area.

        Options: English Language Arts, Mathematics, Science, Social Studies, Other

4. Who Requires the Assessment?

        Options: School, District

5.Who is assessed?

        Options: All Students, Subgroup of Students

6. Primary Purpose of Assessment:

        Options: Summative, Interim/Benchmark, Universal Screening, Progress Monitoring, Diagnostic, Other

7. Total number of students tested each time.
Enter the total number of students tested each time the test is administered. If #6 is Progress Monitoring, then the number of students may vary for each administration, so for this item, enter the approximate average number of students receiving intervention at any time throughout the year.

8. Number of Times the Test is Administered Annually to Each student.
For example, if the same/similar test is administered 3 times annually, enter 3.

9.Time (in minutes) Per Administration:
Select the value that is closest to the actual, designated time for administering this test. If the test is untimed or varies for different students, then select the time taken by the average student.

10. Are the results reported publicly, including in a local Board of Education presentation, the district or school website, and/or newsletter, etc.?
        Options: Yes, No

11. Primary User
Who is the predominant user of this assessment?

Options: District/School Leader, Teacher, Other District/School Staff

12a. Main Use

What is the most important way in which the results are used by the Primary User?

Options: Track/report group and subgroup performance (e.g., benchmarking, measuring growth), Identify specific strengths/weaknesses of students to inform instruction, Assess student learning of specific content, Identify students for intervention (or) exit from intervention, Make a diagnosis about a student, Evaluate the effectiveness/impact of programs, other

12b.Secondary Use

What is next most important way in which the results are used, if applicable, by the district/school

Options: Track/report group and subgroup performance (e.g., benchmarking, measuring growth), Identify specific strengths/weaknesses of students to inform instruction, Assess student learning of specific content, Identify students for intervention (or) exit from intervention, Make a diagnosis about a student, Evaluate the effectiveness/impact of programs, other

Appendix B: List of "Other" assessments

The following assessments were reported by districts under the "Other" category.

English Language Arts (ELA)

- Achieve 3000
- AimsWeb
- American Reading Company
- Amplify
- AP
- ARC
- Assessment for Reading Decodable Texts
- BAS
- Baseline Writing Unit Assessment
- Basic Reading Inventory
- Basic Skills Assessment
- Benchmark
- BOEHM Test of Basic Concepts
- Book Level
- Bookworms
- CBM for Writing (TWW; CWS; WSC)
- CKLA
- CLMS
- Common Lit
- Comprehension and Writing about Reading
- Comprehensive Assessment of Reading Strategies
- Concepts About Print
- Connecticut Documentation & Observation for Teaching System
- CORE
- Corrective Reading
- C-TOPP
- Curriculum Based Measure
- Decodable Running Records
- Degrees of Reading Power (DRP)
- Developmental Assessment of Spelling
- Diagnostic Decoding Surveys
- Dictation
- Differentiated Reading Inventory
- Discovery Conference
- Dolch High Frequency Word List
- DRA3
- DreamBox
- DRP
- DSA- Developmental Spelling Assessment
- Early Literacy Concepts
- Ekwall Shanker - listening comprehension

- EL Education
- ESRI administered to students in reading classes
- Exact Path
- Fountas and Pinnell
- Fundations
- Gallistel - Ellis
- GE Decoding and Encoding Progress Monitoring
- Heggerty
- High Frequency Words
- Highscope Child Observation Record
- HMH
- Imagine Learning
- Independent Reading Level Assessment
- Informal Decoding Assessment
- Informational Writing Post
- Into Lit Unit Assessments
- Into Reading
- IRLA
- Jerry Johns Basic Reading Inventory
- Jump Rope Readers phonic decoding
- KDG Decodable Running Records
- Kilpatrick
- KTEA-Achievement
- Language Live Benchmark Assessments
- LETRS
- Letter and Sound Identification
- Leveled Literacy Intervention Running Records
- Lexia
- Lexia Core 5
- LinkIt
- Marie Clay
- Masi-R Oral Reading Fluency
- McGraw Hill Wonders 2020 Unit Assessments
- Narrative Writing Post
- NMSQT
- No Red Ink Grammar Benchmark
- Nonsense Word Survey
- OLSAT / Gifted and Talented Screener
- On Demand Writing Prompt
- Opinion Writing Post
- Oral Narrative Discourse Assessment (Word Scientists)
- PAST
- Pearson: TELL
- Phonemic Awareness
- Phonic Decoding Assessment
- Phonics for Reading
- Phonics Screener
- Phonological Assessment Screening Test (PAST)

- Pre-ACT
- Preschool Early Literacy Indicators (PELI)
- Promoting Awareness of Speech Sounds (PASS)
- Qualitative Reading Inventory (QRI)
- Qualitative Spelling Inventory
- Quick Phonics Screener
- Quill Baseline Diagnostics
- Quill Growth Assessment
- Read 180
- Read Live fluency and comprehensions
- Reading A-Z
- Reading Fluency assessment
- Reading Interview
- Reading Plus Diagnostic
- Reading Plus G7-12 Benchmark
- Reading Units of Study Post Assessments
- Response to Literature Written Response
- Rewards
- Road to Reading - phonics
- Running Records
- Sadlier-Vocabulary Workshop end of unit assessment
- San Diego Quick Assessment
- SAT
- SAT Bluebook Practice
- SAVVAS
- SBAC Benchmark
- Seeing Stars
- Shaywitz Dyslexia Screener
- SIPPS - Intervention/ SRBI
- Slosson Oral Reading
- Spelling Connections Baseline and Year-End Assessment
- Spelling Connections Unit Assessments
- SPIRE
- Springboard
- Support Coach
- Target Spelling
- Teachers College Letter Sound Identification (Kindergarten)
- Teacher's College Reading Assessment - ELA Growth
- Teaching Gold
- TELL (English Proficiency for EL/ML Learners)
- Test for Adult Basic Education (TABE)
- Test of Narrative Retell (ESOL)
- Test of Written Spelling
- TOSCRF
- TOWL 4
- TOWRE2
- TOWRE-2
- TS Gold Preschool

- UFLI
- Vanderbilt First Sound Frequency
- Visualize and Verbalize
- WADE
- Westport Decoding Assessment
- WILSON End of Step Assessment
- Wilson Fundations
- WIST
- Wit & Wisdom
- Wonders
- Woodcock Johnson Reading Mastery
- Wordly Wise (Spelling)
- Words Their Way Spelling Inventories
- Writing on Demand

Mathematics

- Add+Vantage Math Recovery
- Addition Running Record
- AIMSweb
- ALEKS
- Algebra Readiness
- Amplify mClass
- AP/ECE
- Applications of Mathematics End of Unit Assessments
- ASSISTments
- AVMR
- Big Ideas Mathematics
- Bridges
- CLMS
- CML Tests
- Comprehensive Assessment
- Connecticut Documentation & Observation for Teaching System
- Counting Proficiency Assessment
- CPM End of Unit Assessments
- Delta Math Intervention
- Desmo Unit Tests
- District Number Corner Math Assessment
- Division Running Record
- DMA
- Do the Math
- DreamBox
- Ed Gems
- Elementary Math
- Elementary Milestone Tasks (Unit Assessments)
- ENI-R
- Envision
- EnVision Math Curriculum

- Eureka
- Exact Path
- Exemplars
- Fact and Skill Fluency
- Final Exam
- First Steps
- Focus Math
- Forefront
- Go Math mid-unit assessments
- Go Math Unit Tests
- Hands on Standards
- Hanna Orleans
- Happy numbers
- Heinemann Listening to Learn: Math Inventory
- Heinemann: Do the Math
- High Leverage Assessment (HLA) - All Learners Network
- Highscope Child Observation Record
- HMH Unit Assessments Math
- Illustrative Math
- Into Math
- Investigation
- IOWA Math
- i-Ready Standards Mastery
- IXL
- K-5 Eureka Math Squared Topic Assessments
- Kendall-Hunt Illustrative Math End of Unit Assessment
- Key Math 3
- Maneuvering the Middle: Math Intervention PM Assessments
- Maneuvering the Middle - Unit Assessments
- Mastertrack
- Math +/- Fluency
- Math 180 Block Assessment
- Math Benchmarks
- Math Fact Lab
- Math Placement Continuum
- Math Recovery
- Math Running Records
- Math Series Benchmark
- Math X/ division Fluency
- Mathematics Diagnostic Testing Project
- McGraw Hill
- MClass Math
- MFACTS: Mathematical Fluency and Calculation Tests
- Multiplication Running Record
- Number Corner
- OLSAT
- PK Bridges Mathematics Assessment
- Pre-ACT

- Ready Unit Assessments
- Reflex Math
- REVEAL MATH PLACEMENT TEST GRADE 9
- SAT
- SAT Bluebook Practice
- Savvas
- SBAC Benchmark
- School Based Math Survey
- ST Math
- Subtraction Running Record
- Success Maker
- Symphony Math
- Teaching Gold
- Test for Adult Basic Education (TABE)
- Test of Mathematical Abilities for Gifted Students
- Think! Math
- TOMAGS
- TransMath assessments
- TS Gold Preschool
- Universal Number Sense Screener
- Vanderbilt Computational Fluency
- Vanderbilt Oral Counting and Number ID, K CBM Computation Fluency
- Vmath (Progress Assessment Measure) Fall, Winter, Spring (Math)

Other

- AAPPL
- AIMS Performance Task
- ALIRA (ACTFL Latin Interpretive Reading Assessment)
- Amplify
- AP/ECE
- ASQ & ASQE
- ASVAB
- BASC-Social/emotional
- BATTELE
- Bridge the Gap
- Brigance
- Carolina Science
- CASL-2
- CELF 5
- Classroom Timeline
- Clinical Evaluation of Language Fundamentals-5 Screening Test
- Clinical Evaluation of Language Fundamentals-Preschool 3 Screen
- CLMS
- CogAT
- Connecticut Documentation & Observation for Teaching System (CT DOTS) Connecticut Office of Early Childhood (OEC)
- COR Advantage-Assess Academic Development, Social and Emotional Development

- CREC Science end of unit assessment
- CREC Unit Assessments
- DBQ Project/Online Writing Assessment
- Defined Learning Performance Tasks
- Devereux Early Childhood Assessment
- DIAL
- Early Childhood Outcome (ECO - Brigance)
- Early Screening Inventory
- ESI
- Exact Path
- Final exams
- HMH- End of unit assessments
- IAB - Science
- InView Test of Cognitive Skills
- IQWST Science
- KABC-II
- Kaufman Test of Educational Achievement
- Kindergarten Language Screen- 2
- Language Testing International
- MS Office Cert. Test
- Naglieri
- National Latin Exam
- NNAT
- Observational Behavioral CFAs
- OLSAT
- Open Sci Ed
- OWL-II
- Personal Finance Certification Test
- PPVT 4
- Renzulli-Hartmann Rating Form
- REWARDS program fluency passages and end-of-unit checkups
- Savvas
- Smithsonian Assessments
- Solid Works Cert. Test
- Speech Ease Screening Inventory
- Speed DIAL
- STAMP World Language Assessment
- TACL 4
- TAPS-3
- TCI
- TELD-3
- Test for Adult Basic Education (TABE)
- Timecapsule
- TOMAL 2
- TOPL 2
- Torrance Test of Creative Thinking
- TUVA
- VMI

- WAIS-IV
- WIAT IV
- WIDA Model
- WISC V
- WISC-V- Cognitive
- WJ IV
- WPPSI IV

Appendix C:

Assessment Audit Webpage

Slides shared during district webinar on January 21, 2025

Slide 1

The Assessment Audit Data Collection

January 21, 2025

Slide 2

Assessment Audit
Section 1 of Public Act 24-93

"The Department of Education shall, in consultation with national assessment experts and local and regional boards of education in the state, conduct a comprehensive audit of the assessments that are administered to students."

Such audit shall include, but not be limited to,

(1) issuance of guidance to local and regional boards of education for conducting an inventory of the assessments administered to students at the classroom, school and school district levels,

(2) development of a program of professional learning for teachers concerning assessment literacy, and

(3) an evaluation of the assessments inventoried by local and regional boards of education with the goals of eliminating redundant assessments, discouraging classroom activities that focus only on test preparation, reducing testing time and maximizing assessments that provide actionable information for classroom teachers.

"Not later than January 31, 2026, the Department of Education shall submit" .... a report to the Education Committee.

Slide 3

Tool Development and Implementation

- Reviewed inventories from Achieve, Ohio, California Collaborative for Educational Excellence, Delaware, Illinois, Kentucky, Michigan, Oregon, and WestEd.
- Consulted with Connecticut's Accountability Advisory Committee
- Incorporated feedback from national experts
- Data collection tool designed to balance the need for detailed information with burden on districts and schools
  - Includes assessments required only by districts and schools, not individual teachers.

Slide 4

What assessments must be submitted?

- All core academic assessments that are required by the district or school to be administered without modification to some or all students.
- Core academic assessments means ELA, Math, Science, Social Studies, or related achievement/ability assessments including:

- Vendor-created assessments administered district/school-wide to some or all grades, or used for progress monitoring;
- Locally created assessments that are required to be administered without modification to some or all grades and some/all subjects;
- Smarter Balanced/NGSS Interim assessments that are required to be administered in a standard method; and
- Vendor created K3 benchmark assessments (over and above what is required by the CSDE) that are administered district/school wide.

Slide 5

What assessment must NOT be submitted?

- Required state assessments i.e., KEI, K3 Literacy Benchmark Assessments, LAS Links, CAAELP, Smarter Balanced, NGSS, SAT, CTAA, or CTAS
- National Assessment of Educational Progress (NAEP)
- Advanced Placement / International Baccalaureate Exams
- Seal of Biliteracy exams
- Assessments that are created or modified by individual teachers for use within their classrooms
- Assessments used by teachers that are not required by the district or school
- Non-core academic assessments (e.g. DESSA, climate assessments, physical fitness, or arts assessments)

Slide 6

Getting Started

- Every district submits a single Excel file along with the district assessment calendar.
- The district assessment calendar is a key resource, but District Administrators (DAs) in TIDE should consult others in the district and at the school-level to ensure this is a "comprehensive audit."
- Review the instructions tab within the Excel file carefully and communicate the expectations to others who may be providing information to you for submission.
- Responses should be recorded and saved in the AuditQuestions worksheet.
- Three sample responses are provided, please do not remove those rows. District response should begin on Row 16.

Slide 7

Items 1a & 1b: Name of Assessment/Vendor

- 1a includes a dropdown list with 19 options including "Locally Developed" and "Other."
    - Acadience, DRA, DIBELS, NWEA, PSAT, etc.
- 1b is an opportunity to provide additional detail.
    - If selecting "Other" for 1a, provide more information here.
    - Example of specificity: If NWEA is selected from the 1a dropdown, including MAP Growth for 1b provides additional detail.

Items 2 & 3: Grade(s) and Content Area

- Report all applicable grades for the assessment by selecting Y for each grade
- Select a single content area using the dropdown. Options are Mathematics, ELA, Science, Social Studies, Other.
    - If the assessment reported includes more than one subject area, there should be one entry for each content area.
    - Example: World Language assessment that is not an AP exam or used for the Seal of Biliteracy would be reported as Other.

Slide 9

Items 4 & 5: Who requires the assessment and Who is assessed?

- Who requires the assessment? Using the dropdown, select one of two options.
    - District
    - School
- Who is assessed? Using the dropdown, select one of two options.
    - All Students
    - Select Group of Students

Slide 10

Item 6: Primary Purpose of Assessment

- Summative:
    - An assessment administered at the end of instruction to measure whether students have learned what was expected to be learned
- Interim/Benchmark
    - An assessment typically administered during the year to all students to measure general progress
- Universal Screening
    - An assessment typically administered to all students to identify those who may need additional intervention
- Progress Monitoring
    - Assessments administered for students in intervention to see if they are making progress
- Diagnostic
    - Assessments administered to diagnose specific strengths/weakness or make a particular diagnosis (e.g., identify risk for dyslexia)
- Other

Slide 11

Items 7, 8, & 9: Number of Students, Number of Times, and Time (in minutes)

- Total number of students tested each time the test is administered.

- If the test is used for progress monitoring, the number of students may vary for each administration. In these cases, enter the approximate average number of students receiving intervention at any time throughout the year.
    - NOTE: If a test is administered to all students in a grade, be sure the number reported for item 7 aligns with enrollment.
- Number of times the test is administered annually to each student.
    - If the test is administered in the Fall, Winter, and Spring, enter 3.
- Time in minutes per administration.
    - Select from the dropdown the value that is closest to the actual, designated time for administering the test. If the test is untimed or varies by students, select the time taken by the average student.
    - Options range from 5 minutes to 3 hours.
    - NOTE: If one assessment is measuring content in two areas, requiring two separate entries, adjust the time accordingly.

Slide 12

Item 10: Public Reporting

- Are the results of the assessment reported publicly?
    - Public reporting includes local Board of Education presentations, district or school websites, newsletters, etc.
    - Two response options: Yes or No

Slide 13

Item 11: Primary User

- Primary User: Who is the Primary User of the assessment? Using the dropdown, select one of three options.
    - District/School Leader
    - Teacher
    - Other District/School Staff

Slide 14

Items 12a & 12b: Main and Secondary Uses of the Assessment

- Main Use: What is the most important way in which the results are used by the Primary User?
- Secondary Use: What is the next more important way in which the results are used, if applicable, by the district/school?
    - Select the appropriate dropdown to report uses
    - Track/report group and subgroup performance
    - Identify students for intervention (or) exit from intervention
    - Assess student learning of specific content
    - Identify specific strengths/weaknesses of students to inform instruction
    - Make a diagnosis about a student
    - Determine professional learning needs
    - Evaluate the effectiveness/impact of programs

- - o   Determine resource allocation
    - o   Other

Slide 15

Submission Instructions

- Save the completed Excel file with .xlsx as the extension.
- The file name should be the district name as displayed in cell B3 of the AuditQuestions worksheet.
    - o   For example, Bloomfield's submission should be named 0110011 – Bloomfield School District.xlsx
- Send the Excel file <u>and</u> a copy of the district's assessment calendar to ctstudentassessment@ct.gov
- Due date: Friday, February 21, 2025

Slide 16

Submission Checklist

- ❑   Excel file name matches cell B3 with .xlsx as the extension.
- ❑   Cells B3, B5, B6, and B7 are complete.
- ❑   All sample rows are preserved. District responses begin on Row 16.
- ❑   All assessments have a validation of YES in column AB.
- ❑   None of the assessments listed as "not" to be reported in the instructions are included in the Excel.
- ❑   Assessments that measure two content areas are entered twice and the timing for each administration in a single content area (item 9) is reported accurately.
- ❑   Assessments administered to all students (item 5), include a count of students (item 7) that roughly aligns with enrollment data.
- ❑   Assessment calendar provided

Slide 17

<u>Assessment Audit Web Page</u>

Questions?

Frequently Asked Questions

This FAQ document was compiled based on questions posed during the January 21, 2025, information session and through follow-up correspondence and conversations with districts. The document was updated twice throughout the data collection process. The final version was posted on February 3, 2025.

The questions and answers below are designed to address the questions posed during and following the Assessment Audit Data Collection Information Session held on January 21, 2025. They are meant to supplement the instructions included within the Assessment Audit Excel workbook.

1. My district uses many different tools for progress monitoring for students in intervention. The tool selected for use with each student depends on the needs and goals for that student. How can a district with a wide array of options used in varying intervals for individual students synthesize and report progress monitoring assessments in the audit Excel file?

All districts are required to provide information regarding progress monitoring assessments. We appreciate that this is complicated, so we are offering two response options. For districts that are able to synthesize the information for progress monitoring, report the information in the Excel as explained during the information session. For districts with a complex system for progress monitoring that cannot be adequately reported in the Excel, please provide a **brief** narrative description of your district's progress monitoring assessment practices outside of the Excel tool. When submitting your district's completed Excel via email along with your district assessment calendar, include the narrative explanation of progress monitoring in the body of the email message.

2. Should we include assessments that we use for psychological testing for identification for special education services?

No, do not report assessments used for individualized testing after referral for special education services. Screening assessments administered to a group of students for potential referral to special education *should* be reported. *Previous guidance during the information session indicated that all individualized testing for students with disabilities should be included; this document reflects updated guidance.*

3. Should dual language (Spanish reading) be included?

Yes, if the assessment is required by the district or school. Select "Other" for content area.

4. Should we report preschool assessments such as the Pre-K PELI assessment?

Yes, if this is a district- or school required assessment, it should be reported. Pre-K is a selectable grade.

5. Should we include world language assessments?

Yes, if it is an assessment required by the district or school, it should be reported. Select "Other" for content area.

6. Should we report core academic common summative assessments (unit level) given to all students, but created by a collaborative learning team (teachers teaching the course)?

If the teachers are allowed to individually and independently change the assessments, then those should not be reported. If the district/school is directing the teaching team to agree on a common set of unit assessments, those *should be* reported. They can all be reported on one line as a "locally developed" assessment, with the number of times administered set as the number of unit tests.

7. Should we report the built-in curriculum assessments that are part of the K-3 mandated program we chose?

If the district or school requires teachers to administer the built-in assessments without modification, *and* the assessments go above and beyond the required K-3 mandatory assessments, they should be reported.

8. Should we report ESOL intake assessments?

Any placement assessments administered that are over and above the assessments required in the CSDE English Learner/Multilingual Learner Identification Process should be reported.

9. Should we report Kindergarten screening tools that are completed pre-enrollment at point of registration?

No. Do not report assessments that are completed prior to or as part of registration.

10. Should we report open-ended (i.e. project-based) assessments that may yield a variety of products but that are assigned in a standardized way and/or that are scored using a standardized rubric? What about capstone?

Project-based assessments typically span multiple class sessions where students work under the guidance or oversight of a teacher, either individually or in small groups, to investigate an issue and develop presentations, reports or other products. Assessments of such products should not be reported.

This should not be confused with constructed response items like a common writing assessment, performance task, structured laboratory, or essay prompt that could span more than one class session, but where the assessment is well defined, and all students are responding to the same assessment prompt. Such performance tasks should be reported if they are required by the district/school to be administered without modification by the teacher.

Capstone projects should not be reported.

11. Do midterm exams need to be reported?

Some midterms need to be reported, and some do not. If the test is required by the district or school to be administered, and the teacher cannot change the questions or format of the test, it generally should be reported. If the only requirement is that a midterm be administered, but the format or items are not specified, that should not be reported. Here are some examples of assessments that should and should not be reported:

Midterms that should be reported:

✓ A district Math leader creates a geometry midterm that must be used without modification by all geometry teachers in every school in the district.
✓ A principal directs a team of teachers to create a common midterm that they will all use without modification.
✓ A teacher is the only geometry teacher in the district. They create a standard geometry midterm that they use every year. The principal decides to formalize all school midterms, so this exam is copied and kept in the office for the future. A new teacher would be required to use that midterm.
✓ A principal requires that a team of teachers create a common midterm. The team of teachers decides to create a common multiple-choice section but allow the individual teachers to modify the open-ended section – *in this case, only the common multiple-choice section should be reported*.

Midterms that should *not* be reported:

⊗ A district requires a midterm to be administered but does not specify the form of that midterm.
⊗ A teacher is the only geometry teacher in the district. They create a standard geometry midterm that they use every year. If they left the district, the new teacher would be able to change the midterm.

12. Can I combine multiple assessments in a single row of the Excel or do I need to list them separately?

You may combine multiple assessments for the same subject area, as long as they are used for the same purpose, required by the same people, and are reasonably similar in time tested. If the assessments have very different testing times or number of times tested for different grade levels, they should be split into two rows.

Examples of assessments that may be combined:

✓ All high school mathematics midterms and finals used for summative assessments. Report the average number of students taking the assessments and the number of assessments (in this case, 2). Include only the standardized exams (for example, if Algebra 1, Geometry, and Algebra 2 have common assessments, but precalculus and statistics do not, only report the first three courses). See FAQ question 18 for details on reporting the number of students and grade levels.

- ✓ All end-of-unit Science tests. Report the average number of students taking the tests, and the number of assessments (in this case, the number of units).
- ✓ If a district administers four math interim assessment blocks (IABs) for 3rd grade, and six math IABs for 4th grade, the 3rd and 4th grade math IABs can be reported on one row, with the average number taken by 3rd and 4th grade students (in this case, five).

Examples of assessments that may *not* be combined

- ⊗ The PSAT must be reported on two lines, one for Math and one for ELA.
- ⊗ If a district administers one math IAB for 3rd grade and 10 math IABs for 4th grade, then they should use two rows, one for 3rd grade and one for 4th grade.

13. My school requires that 3rd and 4th grade teachers administer *at least* three ELA IABs throughout the year, however, we allow teachers to select which ELA IABs they administer. Some teachers choose to administer more ELA IABs. The assessments must be administered in a standardized fashion. Does this need to be reported, and how should we report?

Yes, this needs to be reported. Select "Smarter Balanced Interim Assessments" for the name of vendor. Select "English Language Arts (ELA)" for content area, enter "3" for item 8 in the Excel (Number of times test is administered annually).

14. We are a small district. We have only one middle school. Do we pick "district" or "school" for question 4?

Think about who is requiring the assessment. Is the superintendent, district-level curriculum coordinator, or someone else at the district level requiring the assessment? If so, select "District". Is the principal requiring the assessment? If so, select "School". Another way to think about this: if the principal or school administrator left, would the assessment still be required? If so, select "District".

15. Can I report a range for number of times administered and time in minutes as it may vary by grade?

No. If the number of times administered or time in minutes is close, take the average. If the number of times administered or time in minutes is very different by grade, use two different rows to report.

16. When a standard local assessment is administered across a district, but per IEP's may be modified for a few students, should those still be reported as most students will receive it without modification?

Yes. Include all students, including the students who take a modified test in the student count for the standard local assessment. For time spent testing, put the average amount of time spent on the assessment for all students.

17. How should we report ELA assessments that are administered in another language?

If the majority of students are taking an ELA assessment, and a subset of students are taking a translated version of the assessment for the same purpose, report all students together. Select "English Language Arts (ELA)" for the content area.

18. We require all geometry students to take a common midterm. Most geometry students are 10th graders, and most 10th graders are geometry students, but there are some 9th grade students taking geometry, and some 10th graders take algebra II. How should I fill out the sheet?

Mark "Y" for the grade level that is most common for the assessment, in this case, 10th grade. If there are many 9th grade students taking the assessment, you may check off 9th grade as well. Select "select group of students" for question 5. Report the total number of geometry students for question 7 in the Excel. If you have multiple common midterms in math, they may all be reported together (see FAQ question 12).

19. If we use only some subtests of a standardized battery, what should we report? For example, DIBELS and Acadience have multiple subtests. In some cases, we use only one or two subtests, not the whole battery.

All subtests can be reported on one row, the total testing time should reflect the testing time for the subtests used. You may choose to specify the subtests used under question 1b.

20. Do we have to specify which specific IABs are used?

No. You may choose to specify which IABs are used under question 1b if it helps you keep track of the assessments you are reporting, but it is not required.

21. If our district intends to change our assessment protocol in 2025-26, should we report our intended plan or do we report our practices for 2024-25?

The Excel entries should reflect practices implemented in 2024-25 only.

[i] Marion, S. F., Pellegrino, J. W., & Berman, A. I. (Eds.). (2024). Reimagining Balanced Assessment Systems. National Academy of Education.

[ii] National Research Council. 2001. Knowing what students know: The science and design of educational assessment. Committee on the Foundations of Assessment. Pelligrino, J., Chudowsky, N., and Glaser, R., editors. Board on Testing and Assessment, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press. nationalacademies.org/read/10019/chapter

[iii] The Council of Chief State School Officers (CCSSO)'s collaborative called the Formative Assessment for Students and Teachers (FAST) describes formative assessment as:

> "a planned, ongoing process used by all students and teachers during learning and teaching to elicit and use evidence of student learning to improve student understanding of intended disciplinary learning outcomes and support students to become self-directed learners.
>
> Effective use of the formative assessment process requires students and teachers to integrate and embed the following practices in a collaborative and respectful classroom environment:
> * Clarifying learning goals and success criteria within a broader progression of learning;
> * Eliciting and analyzing evidence of student thinking;
> * Engaging in self-assessment and peer feedback;
> * Providing actionable feedback; and
> * Using evidence and feedback to move learning forward by adjusting learning strategies, goals, or next instructional steps."

[iv] See The Relationship between the Smarter Balanced Grade 8 and the PSAT 8/9 see Encouraging Participation in Rigorous Courses: Rationale, Methods And Results

[v] Tools for Teachers provides teacher-created, standards-aligned instructional resources (lesson plans) that incorporate the formative assessment process with embedded formative assessment, accessibility strategies, and printable worksheets.

[vi] Marion, S. F., Pellegrino, J. W., & Berman, A. I. (Eds.). (2024). Reimagining Balanced Assessment Systems. National Academy of Education.

[vii] National Research Council. 2001. Knowing what students know: The science and design of educational assessment. Committee on the Foundations of Assessment. Pelligrino, J., Chudowsky, N., and Glaser, R., editors. Board on Testing and Assessment, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press. nationalacademies.org/read/10019/chapter

[viii] The Council of Chief State School Officers (CCSSO)'s collaborative called the Formative Assessment for Students and Teachers (FAST) describes formative assessment as:

> "a planned, ongoing process used by all students and teachers during learning and teaching to elicit and use evidence of student learning to improve student understanding of intended disciplinary learning outcomes and support students to become self-directed learners.
>
> Effective use of the formative assessment process requires students and teachers to integrate and embed the following practices in a collaborative and respectful classroom environment:
> * Clarifying learning goals and success criteria within a broader progression of learning;

- Eliciting and analyzing evidence of student thinking;
- Engaging in self-assessment and peer feedback;
- Providing actionable feedback; and
- Using evidence and feedback to move learning forward by adjusting learning strategies, goals, or next instructional steps."

[ix] See The Relationship between the Smarter Balanced Grade 8 and the PSAT 8/9 see Encouraging Participation in Rigorous Courses: Rationale, Methods And Results

[x] See The Relationship between the Smarter Balanced Grade 8 and the PSAT 8/9 see Encouraging Participation in Rigorous Courses: Rationale, Methods And Results

[xi] Kuhfeld, M. (2019). Surprising new evidence on summer learning loss. Phi Delta Kappan, 101(1), 25-29. Rethinking summer slide: The more you gain, the more you lose - Kappan Online

[xii] Briggs, D. C. (2022). Gain Scores and the Regression Fallacy. CADRE Research Brief. University of Colorado Boulder https://www.colorado.edu/cadre/2022/03/10/gain-scores-and-regression-fallacy.