



OFFICE OF ELEMENTARY AND SECONDARY EDUCATION

SCHOOL SUPPORT AND ACCOUNTABILITY

2023 State Assessment Conference

1E - Test Administration

September 27, 2023



FOCUS AREA: ASSESSMENT PEER REVIEW

Session 1E- Test Administration

Nathan Dadey, Center for Assessment, ndadey@nciea.org

Phoebe Winter, phoebe.winter@outlook.com

Zach Warner, NYS Ed Dept, zachary.warner@nysed.gov

Chris Rozunick, Texas Education Agency, christine.rozunick@tea.texas.gov



A note About this Conference/Session

- The purpose of this conference/session is to provide an opportunity for State education agency (SEA) staff to interact and engage with relevant experts and other SEA staff about the Department's assessment peer review.
- The observations and opinions of the session presenters are their own.



Outline

1. Framing and overview (15 min)
2. Issues and options (30 min)
3. Responding to peer review requirements (15 min)
4. Q&A (15 min)



1. Overview of Critical Elements



Which Critical Elements?

- 2.3 - Test Administration
- 2.4 - Monitoring Test Administration
- 2.5 - Test Security
- 5.1 - SWD Inclusion
- 5.2 - EL Inclusion
- 5.3 - Accommodations



In Summary

These elements are focused on ensuring that students have a fair opportunity to show what they know and are able to do, and as a result, that inferences made about what students know and can do are defensible.

Doing so involves ensuring:

- Testing conditions are consistently standardized, secure and monitored.
- Test irregularities are minimized, and if they occurred, addressed.
- Ensuring all students can participate and show what they know and can do, including SWDs and ELs, through inclusion policy and accommodations.



Deeper Dive: 2.3 - Test Admin

The State implements policies and procedures for standardized test administration

- 1. Standardization.** Materials and procedures for (a) test distribution and administration, including (b) documentation of irregularities, and (c) requesting and receiving accommodations are clear and clearly communicated.
- 2. Training.** Test coordinators and administrators have been appropriately trained.
- 3. Technology.** For CAT, (a) technology requirements, (b) instructions for use of technology in administration, and (c) solutions to technology issues and contingency plans are clear and clearly communicated.



Deeper Dive: 2.3 - Test Admin

Common Sources of Evidence:

1. Test coordinator manuals, test administration manuals, accommodations manuals, training materials, official communications to educators.
2. Training materials, documentation of training attendance.
3. Technology requirements and specifications, official communications to educators, test coordinator manuals, training materials.

Common Pitfalls:

- Not providing evidence that trainings actually happened or were attended. E.g., evidence that answers questions like “did the trainings happen” and “did anyone attend the trainings?”.



Deeper Dive: 2.4 - Monitoring Admin

The State adequately monitors the administration of its State assessments to ensure that standardized test administration procedures are implemented with fidelity across districts and schools.

- 1. Procedures.** Effective procedures are in place to monitor test administration.
- 2. Implementation.** The monitoring procedures have been implemented with fidelity.



Deeper Dive: 2.4 - Monitoring Admin

Common Sources of Evidence:

- **Monitoring Procedures and records.** Monitoring procedures, monitoring training materials, records of administration monitoring, communications to schools and districts on monitoring, summaries of monitoring reports, summaries of monitoring appeals
- **Self Report and Help Desk Reports.** Process and tools for reporting irregularities, summaries of reports, help desk training materials, help-desk logs

Common Pitfalls:

- Not being comprehensive with policies (e.g., expecting “common sense” instead of explicit detail)
- Not demonstrating that monitoring occurred



Deeper Dive: 2.5 - Security

The State has implemented and documented an appropriate set of policies and procedures to prevent test irregularities and ensure the integrity of test results.

- 1. Irregularity Prevention.** Security procedures are in place (1) for all who access test materials and (2) to ensure security before, during and after admin
- 2. Irregularity Detection.** Documented procedures for reporting, forensic analyses are conducted
- 3. Investigation and Remediation.** Documentation of follow-up investigative procedures, documentation of consequences for violations, procedures for responding to breaches



Deeper Dive: 2.5 - Security

Common Sources of Evidence, in addition to the evidence mentioned in 2.4:

1. Test security handbook, test administration manuals, directions to proctors and test takers, audit trails or other documentation of implementation, CAT exposure controls
2. Forensic analysis procedures, reports of forensic analyses.
3. Investigation procedures, investigation training materials, procedures for issuing and responding to findings of investigations, investigation reports.

Common Pitfalls (similar to 2.5):

- Not being comprehensive with policies (e.g., expecting “common sense” instead of explicit detail)
- Not demonstrating that policies and practices (e.g., forensic analysis) were implemented



Deeper Dive: 5.3 - Accommodations

The State makes available appropriate accommodations and ensures that its assessments are accessible to students with disabilities (SWDs) and English learners (ELs), including ELs with disabilities.

- 1. Accommodations are Provided.** SWDs under IDEA and ELs under section 504 are provided with accommodations.
- 2. The Accommodations are Appropriate.** The accommodations provided are appropriate and effective at reducing barriers, do not alter the construct, and are comparable to those not receive the accommodations
- 3. Exceptional Accommodations Requests.** Process for addressing requests for accommodations beyond those typically allowed, process communicated to the field



Deeper Dive: 5.3 - Accommodations

Common Sources of Evidence:

1. Accommodations manual; technical manual; accommodations guidance; IEP/504/EL requirements, templates, training materials; official communications, FAQs.
2. Technical manual, measurement invariance studies, literature reviews, cognitive laboratories, expert/advocate judgment, studies of differential boost
3. Official communications, exceptional accommodations request process and approval criteria, audit trails

Common Pitfalls:

- Not demonstrating that students who needed accommodations received accommodations



2. Considering Multiple Approaches



On Standardization and Multiple Approaches

Standardization does **not** mean that every student receives the:

- Exact same test.
- Under the exact same conditions of measurement.

Instead, variation in tested content and conditions of measurement are **purposefully connected to the inferences to be made about students, through a consistent administration procedure.**



Consistent Administration

For multiple approaches:

- Define what an “consistent administration” is, in detail, which involves:
 - Explain the allowable **variations in the administration process.**
- These variations may depart from previous practice.
- Challenges arise when the timing or content of administrations vary from one student to the next
 - Especially when both content and timing vary from student to student



Example Application

Students are administered multiple “testlets” after students have been instructed on a topic based on educator judgment.

- The process for identifying when students have had sufficient instruction to be assessed has been defined, documented and communicated to educators
- Educators have received training on this identification process
- Educators implement this process with fidelity
- The state monitors the administration for patterns that suggest educators are not implementing the process with fidelity (e.g., no testlets given during the semester) or that patterns inappropriately vary across student groups.



Considering Multiple Approaches

Defining & Implementing a Consistent Administration

Scaling Up Current Practice.
Increasing the frequency or amount of various aspects of current administration practice.

Developing New Practices.
Addressing new challenges or issues that arise from the approach used.

An Aside on Burden



Some of these approaches result in increased burden on administrators and coordinators



- **Managing administrations**
 - Potentially involving new responsibilities for teachers, leaders, administrators and coordinators
 - Often, district level training on assessment begins in early summer, requiring the appropriate lead time from the program.
- **Understanding and supporting these systems**, particularly if the approach is **different across subject areas** (both for those who support administration and those who have to understand it, e.g., principals, district test coordinators)



3. Issues and options



Policies and Practices for Consistent Administration

- Although best practices for administering the types of assessments currently seen under ESSA are fairly established, **work is still underway** to research and put forth fair administration policies for new types of assessment.
- Traditionally, these policies and practices have focused on standardization in terms of trying to come as close as possible to **all students taking the same test in the same way** (with accommodations often supporting students in that aim)
- These approaches to assessment likely require to **shifts to administration policies and practices** that ensure that the intended interpretations supported by the way the assessment is administered to individual students.



3A. Through-Year Assessment



Through-Year Considerations

Appropriate administration policies and practices could include:

1. consistent provision of accommodations (with decision rules to make changes if appropriate);
2. collection of classroom identifiers to facilitate reporting;
3. tracking and addressing missing data throughout the year (within and across schools);
4. thoughtful consideration about makeups (both within-year and end of year);
5. decision rules about what constitutes a valid case/score; and
6. significant help desk-style support throughout the year.
7. considerations of within-year exposure



Through-Year Considerations

Appropriate administration policies and practices could include:

1. consistent provision of accommodations (with decision rules to make changes if appropriate);
2. collection of classroom identifiers to facilitate reporting;
3. tracking and addressing missing data throughout the year (within and across schools);
4. thoughtful consideration about makeups (both within-year and end of year);
5. decision rules about what constitutes a valid case/score; and
6. significant help desk-style support throughout the year.
7. considerations of within-year exposure



Missing Data

Addressing missing data involves a combination of

1. **Administration Policy** designed to obtain as complete and administration as possible, including whether make ups are permitted, and if so:
 - When make ups are administered (during or at the end of year)
 - Who makes decisions around make ups
2. **Rules or adjustments that addressing missing data**, including rules around what constitutes a valid score for accountability (i.e., what defines participation and non-participation)
 - Whether there is an allowable number of missed administrations
 - And associated rules with missing administrations (e.g., reduced reporting)
 - Special rules for mobile students



Missing Data – Identifying Types of Mobile Students

There a limited number of missing cases can the state could identify, monitor and develop rules for, e.g., a student

- Is absent for the entire testing window
- Moves up or down a grade within the testing window
- Moves from one classroom, school, or district to another within the testing window
- Moves into the SEA's jurisdiction within the testing window
- Moves out of the SEA's jurisdiction within the testing window
- Tests, but receives an invalidated result due to incompleteness or void

Each possible type of mobile student should be accounted for via rules.



Addressing Missing Data: TTAP

- With three distinct opportunities, we assume we will have to deal with missing data and are considering this as we investigate what a comprehensive or summed score might be.
- Operationally, we currently only offer retest opportunities for high-stakes end of course assessments and do not do so for grades 3-8.



Through-Year Considerations

Appropriate administration policies and practices could include:

1. consistent provision of accommodations (with decision rules to make changes if appropriate);
2. collection of classroom identifiers to facilitate reporting;
3. tracking and addressing missing data throughout the year (within and across schools);
4. thoughtful consideration about makeups (both within-year and end of year);
5. decision rules about what constitutes a valid case/score; and
6. significant help desk-style support throughout the year.
- 7. considerations of within-year exposure**



Exposure and Memorability

- Having multiple administrations within-year means considering whether and how content exposure in an earlier administration may, or may not, compromise later administrations.
 - E.g., should some content be saved for later windows?
- Increased exposure due to within-year windows or multiple-forms, potentially leading to increased testing.
- For items and tasks that are more involved, memorability may also be an issue.
 - Consider whether and how to investigate the memorability of unique tasks.



TTAP

- Current Texas pilot uses Multi-stage Adaptive model with all testlets for a given year created at one time to ensure item coverage is similar across all three opportunities for any given year.
- Once an item is used for one opportunity, it may not be reused for another opportunity that year.
- Operationally we are required to release 100% of our items but this does not yet apply to our pilot program.



3B. Performance Assessment & Content Based Matrix Sampling



Performance Assessment

Appropriate administration policies and practices could include:

1. guidelines for how to integrate the performance assessment into ongoing instruction;
2. suggestion/rules for consistent distribution and collection of materials;
3. support for determining appropriate student accommodations based on the specific attributes of the assessment and the student's needs; and
4. clear directions on which administration procedures must be followed exactly and where local flexibility is allowed without impacting the interpretations.



Matrix Sampling

Matrix-sampling approaches likely pose the fewest challenges given the similarities to current assessments administered under ESSA. Many of the traditional policies and practices are in place. Additional considerations could include:

1. Sampling plan to ensure that content coverage is appropriate and comprehensive across students (and allows for inferences at the subgroup level);
2. Limitations to testing windows to ensure any instructional timing effects are minimized when results are aggregated; and
3. audit procedures to examine, among other things, the influence of differences in instruction across districts and how that may impact the aggregated results.

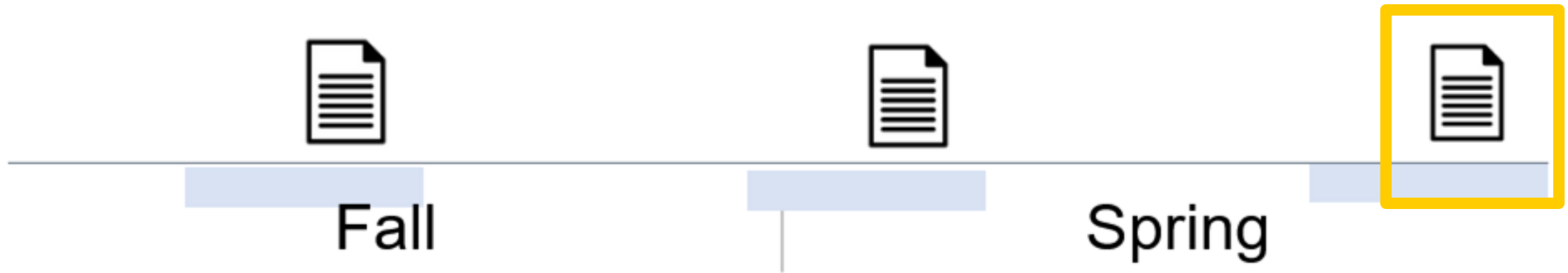


3. Responding to peer review requirements



Reminder on the Limits of Peer Review on Administration

- Again, peer review is only concerned with assessments that produce **summative annual determinations**
 - This extends to administrations as well.





Consistent Administration

- Define what an “consistent administration” is, in detail, especially within the technical report, relevant manuals, communications and peer review notes.
 - Explain the allowable variations in the administration process.
- For peer review, explain how these variations support the interpretation and uses of the assessment.
 - Doing so may involve new criteria or redefining old criteria.



Considering Criteria

- E.g., consider a program where multiple within year assessments were given based on educator judgment.
- Evidence that doing so supports a consistent administration could involve:
 - A rationale of how the process for a consistent administration supports the validity argument.
 - An explanation of the appropriate administration process.
 - Clear boundaries and guardrails for that process.
 - Training on that process.
 - Documentation that the process was implemented with fidelity.



Monitoring

- With designs that increase the number of within year administrations, monitoring may need to increase
- Alternative approaches may be used to alleviate burden, which may involve:
 - sampling in person monitoring,
 - Leveraging technology to increase monitoring capacity, e..g, remote monitoring of in person administrations



3. Q&A



QUESTIONS?





STILL MORE QUESTIONS?

- Submit your questions using the QR code
- Attend session 1G (*Preparing for Assessment Peer Review*) Wednesday afternoon for answers

