



Georgia Assessment Working Committee

Scott Norton
Council of Chief State School Officers

October 3, 2016



Agenda Topics

- ⌘ Locally Selected, Nationally Recognized Assessments
- ⌘ Summative vs. Interim Assessments
- ⌘ Innovative Assessment Pilot



Locally Selected, Nationally Recognized Assessment Provision in ESSA

ESSA's "Nationally Recognized High School Assessment" Provision

- Many of these slides were developed by Brian Gong from the Center for Assessment and were presented to state assessment directors in a August webinar
- The presentation reviews the ESSA provision about district requests to use a "nationally recognized" assessment in lieu of the state high school assessment
- The presentation provides guidance on preparing state criteria and processes for reviewing proposed high school assessment proposals
 - Implications for assessment and accountability
 - Implications for management, credibility, communication

ESSA: nationally recognized high school assessment

- New federal ESSA law (*Every Student Succeeds Act*) has a provision that a district/LEA may request state approval to use a “locally selected, nationally recognized high school academic assessment” in lieu of the state high school assessment
- Once an assessment is approved for one district, any district in state may use that assessment without additional state approval process
- Districts must follow provisions for local notification (e.g., parents) before and during use
- https://www.federalregister.gov/articles/2016/07/11/2016-16124/title-i-improving-the-academic-achievement-of-the-disadvantaged-academic-assessments?utm_content=header&utm_medium=slideshow&utm_source=homepage See Section 200.3

ESSA specific provisions

- “Nationally recognized high school academic assessment” means an assessment of high school students’ knowledge and skills that is administered in **multiple States** and is **recognized** by institutions of higher education in those or other States for the purposes of **entrance or placement** into courses in postsecondary education or training programs.” (Proposed assessment rule, 4/19/16; Posted for comment 7/11/16) – **ACT, SAT, Accuplacer, Smarter Balanced, PARCC, AP, IB, etc.**
- Is aligned with the challenging State academic standards; addresses the depth and breadth of those standards; is equivalent to or more rigorous than the statewide assessments
- Produces **valid and reliable** data on student academic achievement for all high school students and subgroups that are comparable to those produced by the statewide assessment; are expressed in terms consistent with the State’s academic achievement standards; provide for differentiation among schools within the state for accountability
- Has appropriate **accommodations**
- Submit **Peer Review** evidence to Secretary
- Applies to **all high school students in the LEA** except AA-AAS (alternate assessment)
- LEA provisions (notify parents, etc.)

This ESSA provision is unprecedented

- **Adoption of College Entrance Exam as statewide test**
 - Several states have used a college entrance exam (i.e., ACT, SAT, augmented/unaugmented) for the state's high school assessment,
 - but almost always had statewide administration
 - Some states allowed a menu of assessments for high school graduation qualification (e.g., state test *or* AP) with established cutscores; but students almost always took the state test, so there was a **single** assessment system for **school** accountability, with a **multiple** assessment system for **student** accountability
- **Adoption of multiple tests, none of which is administered to all students**
 - This provision could result in **multiple** high school assessments being administered during the year, so **students could take different tests across** districts.
 - The assessments could have **multiple contractors/administrations** and would need to be **combined for reporting and use in accountability.**

State preparation

- It is anticipated that nearly every state will receive a request from at least one district.
- States should:
 - **specify** the technical and operational requirements for an acceptable **multiple**-assessment system to be used for school accountability
 - **anticipate** what it takes to implement,
 - **decide** on appropriate state roles and responsibilities;
 - **establish** application and review processes and criteria to determine whether to approve a district request;
 - plan how to **monitor** implementation
 - put in place a **communication plan**

Questions for Consideration

- Is the assessment acceptable quality?
- How will the assessment work with the state assessment?
- How will the assessments work with other system components (e.g., accountability)?
- What does the process need to work well?
- What are the state's views/values regarding multiple assessments, especially a college entrance exam?

Specify requirements – single assessment

Technical for Accountability

- Alignment, including rigor
- Range
- Security
- Data privacy, data ownership
- Growth
- Testing time
- Grade 3-11 cutscore coherence

Fairness

- Fit with state demographics
- Accommodations
- Retesting
- Test preparation support

Relationship with Instruction

- Informing instruction
- Encouraging high level instruction

Credibility and Student motivation

For more detail, see Martineau, Gong, & Zurkowski. (2016). Preparing for ESSA's "Nationally Recognized High School Assessment" Provision. Presentation at the CCSSO National Conference on Student Assessment. June 21, 2016. Philadelphia, PA.

<https://ccsso.confex.com/ccsso/2016/webprogram/Session4740.html>

Specify requirements – multiple assessments

- Issue is not only “Is assessment high quality?” but “Is it comparable and compatible enough?”

Technical Requirements - comparable

Content: Assessment is aligned to state assessment blueprints and specifications

Scores & ALDs: Scale scores, ALDs, and cutscores are comparable; allow aggregation for reporting (e.g., avg. scale score, percent proficient), calculation of growth scores, use in making accountability determinations

Accommodations: Must be equivalent to those allowed on state assessment

Scaling/Equating: Sound maintenance of score meaning, comparable to state assessment; also, changes in tested populations must not compromise state’s ability to scale and equate state test over time

Precision: Must be possible to provide metrics of assessment precision for data combined across tests

Reports: Reports provide at least same information as state assessments (e.g., scoring; student)

Updates/changes: stay in synch with state tests

Operational Requirements - compatible

Test Window: Test windows must be acceptably similar to state assessment test windows to support claims of fairness and comparability

Test Mode/Platform/Training: Adequate training must be provided for each assessment

Administration Support: Adequate support must be provided during test administration

Data Processing: Data must be provided to state in detail, in format, and on time for state processing.

Data Security/Ownership: Data must be (in)accessible in compliance with state policy for state assessments

Responsive: Vendor should handle any issues with its testing as responsively as would be done with state testing (e.g., rescoring request).

Transparency/Communication Support: Vendor should provide at least as much support in materials, expert labor for critical issues as is expected of vendor for state assessment (e.g., released items; anomaly)

Other implications

- Consider implications of multiple high school assessments for rest of educational system:
 - Grades 3-8 ALDs and cutscores
 - Accountability system
 - “Career readiness” efforts
 - School/district support
 - Public support
 - Assessment contract planning (e.g., volume pricing)
 - State department of education capacity (e.g., management support and communication)

Decisions for Assessment Review Process

- What are technical and operational criteria? What is acceptable evidence? How to evaluate practical impacts of dealing with multiple assessments?
- When will reviews be available? (Starting what year?) When will reviews be conducted (annually)?
- How many likely applications to review?
- What will review process be? How independent? How elaborate? Advisory? Who will generate the specific review protocols/forms? What will the appeal process be? How is evaluation review related to state's Peer Review?
- Who will reviewers be? What qualifications? How recruited? Paid?
- Who generates information for adoption/evaluation review?
- Who pays for what? State or district? Who negotiates with assessment provider?
- How will state formulate policies and communicate? How will state update policies?
- How will state ensure needed capacity for dealing with additional complexity?

Example: state roles in review process

More
district



More
state

- State establishes criteria, sets up application process. State says districts/publishers are largely responsible for assembling evidence for application, generating acceptable assessment data for state use, and paying for the additional assessment option
- State establishes criteria, sets up application process. State takes large role in gathering evidence and implementing tests.
- State reviews and adopts multiple assessments without district request. State takes responsibility for gathering evidence, negotiating contracts, paying for the multiple assessments, and implementing tests.

Summary

- Be **clear** about state's values, uses, resources
- Specify the criteria **both** for an acceptable high school assessment *and* for acceptable **multiple** assessments, considering systemic coherence
- The criteria must address not only technical quality but probably **all operational aspects** in state's assessment contracts.
- The process is not only technical review but **procurement**.
- **Communicate** before, during, and after adoption/review process
- The state will need **specific** criteria, evidence evaluation guidelines (rubrics), policies, and programs.
- The state should be **proactive**.



Using Interim Assessments Instead of Summative?: Considering an ESSA Option

Overview

- ESSA provision about interim assessments
- Guidance for considering interim assessments instead of summative
- Discussion

ESSA interim assessment provision

- New federal law (ESSA—*Every Student Succeeds Act*) allows state to consider using “a single summative or multiple interim assessments” to comply with assessment and accountability requirements of ESSA
- Interim assessments would need to pass Peer Review (just as Summative Assessments do now)

Why states may be interested

- Overall reduction in testing time (if state and districts already administer summative and interims)
- Have state-sponsored test provide more instructionally useful information (evaluate districts/teachers' use of interim assessments)
- May provide ways to allow more complex, extensive assessments for summative use (e.g., performance and/or curriculum-embedded assessments prior to end-of-year summative)

Guidance for considering interim assessments instead of summative

- What is the state's summative claim ?
- What information is provided by the interim assessment?
- What administrative conditions are required for each?

What is the state's summative claim?

- Student is ready for next stage (close in time to when student will enter next stage) because student has learned the requisite knowledge/skills
- Student has learned the requisite knowledge/skills at some time (which may be subject to forgetting or additional learning)
- Student is likely ready for the next stage based on predictive performance, but no claim that the student has learned full set of requisite knowledge/skills

What information is provided by interim assessment?

Learning sequence of 10 topics/content standards during year									
A	B	C	D ₁₂₃₄	E	F ₁₂₃	G	H	I	J
Sept	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June

Four interim assessment instruments & content topics assessed	State test & content assessed
---	-------------------------------

C, D4, F ₂ , etc.	C, D4, F ₂ , etc.	C, D4, F ₂ , etc.	C, D4, F ₂ , etc.	C, D4, F ₂ , etc.
---------------------------------	---------------------------------	---------------------------------	---------------------------------	---------------------------------

In this model, the interim assessment instruments mirror the end-of-year state test in terms of content, balance of emphasis, format, administration conditions, etc. Each test administered during the year covers the same content and has the same design. This design provides high "practice" and high "prediction" from the interim to the end-of-year state test. It is also an excellent design for program evaluation of the impact on learning of an instructional program between pre- and post-tests.

A, B	C, D	E, F	G, H	C, D4, F ₂ , etc.
------	------	------	------	---------------------------------

In this model, the interim assessment instruments focus on the content that was instructed. Each interim measure covers only the content in the most recent instructional period, and thus each test's content differs from the others. This may be the best design for assessing recent instruction and informing remedial work on what was recently instructed. It may not be an effective predictor of student performance on the state test if students forget after instruction.

A, B	A, B, C, D	A, B, C, D, E, F	A, B, C, D, E, F	C, D4, F ₂ , etc.
------	---------------	---------------------	---------------------	---------------------------------

In this model, the interim assessment instruments are designed to assesses what was instructed, but is cumulative, i.e., the assessment includes all topics instructed up to that point in time. This model values student retention of knowledge previously taught. It may not be an effective or efficient way to predict student performance on the state test.

What administrative conditions are required for summative/interim?

Condition	Summative	Interim
Administration & scoring	Highly standardized	Flexible in timing
Administration Security	High	High (district high stakes) to low (classroom use)
Curriculum-specific	Low	Low for commercial; high for district-custom
Accommodations	State mandated	Variable; rarely validated
Transparency (e.g., public, reviews)	High (typically)	Low (typically)
Control	High (if state custom)	Low state control if interim controlled by district
Participation	High for accountability	Low for individual student; Med-high for district eval.

While many states are deciding against the interim option, **interest remains**, especially from teachers. e.g.,

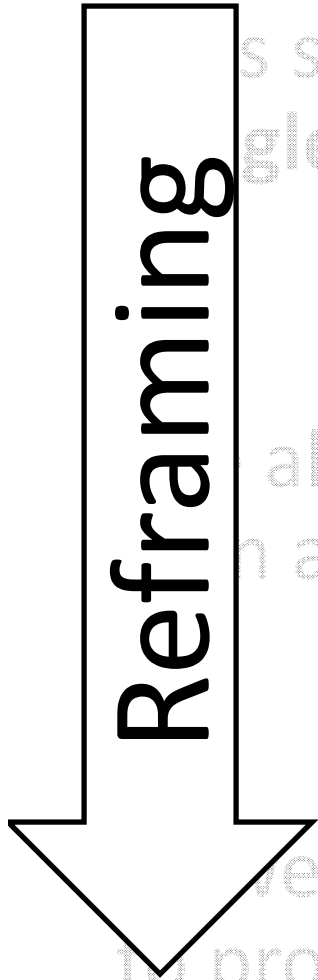
Could new plans to adapt a test teachers love work for Indiana?

By Shaina Cavazos @ShainaRC scavazos@chalkbeat.org
PUBLISHED: June 14, 2016 - 4:47 p.m. EDT

Indiana educators who want the state to consider replacing ISTEP with an exam that would give teachers immediate feedback about their students might have new reason to be optimistic.

Chalkbeat reported last week that teachers have called on the state to use something like the popular MAP test, which is beloved by teachers for providing real-time information about what students know. At the time, the company that makes the test said MAP isn't designed to meet testing requirements demanded for federal accountability and Indiana law.

Now, however, the Oregon-based testing company that makes MAP says it's looking for a way to respond to demand from states like Indiana that have school accountability systems that require a single yearly score to indicate which kids are performing at grade level.



Replace

Predict

Complement

Part I: Replace

Some pressing concerns

- Security
- Administration
- Claims
- Combining/Aggregation

The claims to be made about students are key.

Determining whether the claim(s) will be about what a student knows and can do at the **end of the year** is also key.

Example End of Year Claim

At the end of the year, students can solve a variety of problems based on content from each standard.

Example “Unit-Based” Claim

Shortly after instruction on the standard, students can solve a variety of problems based on content from that instructed standard.

Example End of Year Claim

The current summative assessment can be used as a criterion for these claims.

Example “Unit-Based” Claim

Little guidance for these types of claims.

Part II: Predict

How well can we predict end of grade
performance

and

What is the shortest test needed to reach
that level of prediction?

Part III: Complement

Can we get more information by combining the results of the interim and summative assessments?

The background image shows a close-up of a math worksheet. A pencil is pointing to a long division problem: $23 \overline{) 109,345}$. The quotient $4,754$ is written above the line. The worksheet also contains other text like "Divide." and "Date".

Students often take hundreds of items across multiple assessments.

However, assessment results are generally examined in isolation.

The Goal: find a way to combine information
across multiple assessments
for *all* assessment given *in* an academic year.

**ESSA Innovative Assessment
Demonstration Authority:
Regulations &
Technical Requirements**

What is the opportunity at hand?

To innovate new approaches to assessment and accountability that provide:

- More **valid, varied, and richer measures** of student learning and progress;
- Greater understanding of a **broader set of skills** that more clearly signal the nature of college and career readiness for students, educators, parents, and policymakers;
- More **timely and useful data** that allows educators to make real-time adjustments to student instruction; and
- Better **alignment with student-centered models of learning and instruction**, such as personalized or competency-based education, that are designed to achieve high learning expectations for all students.

In other words...

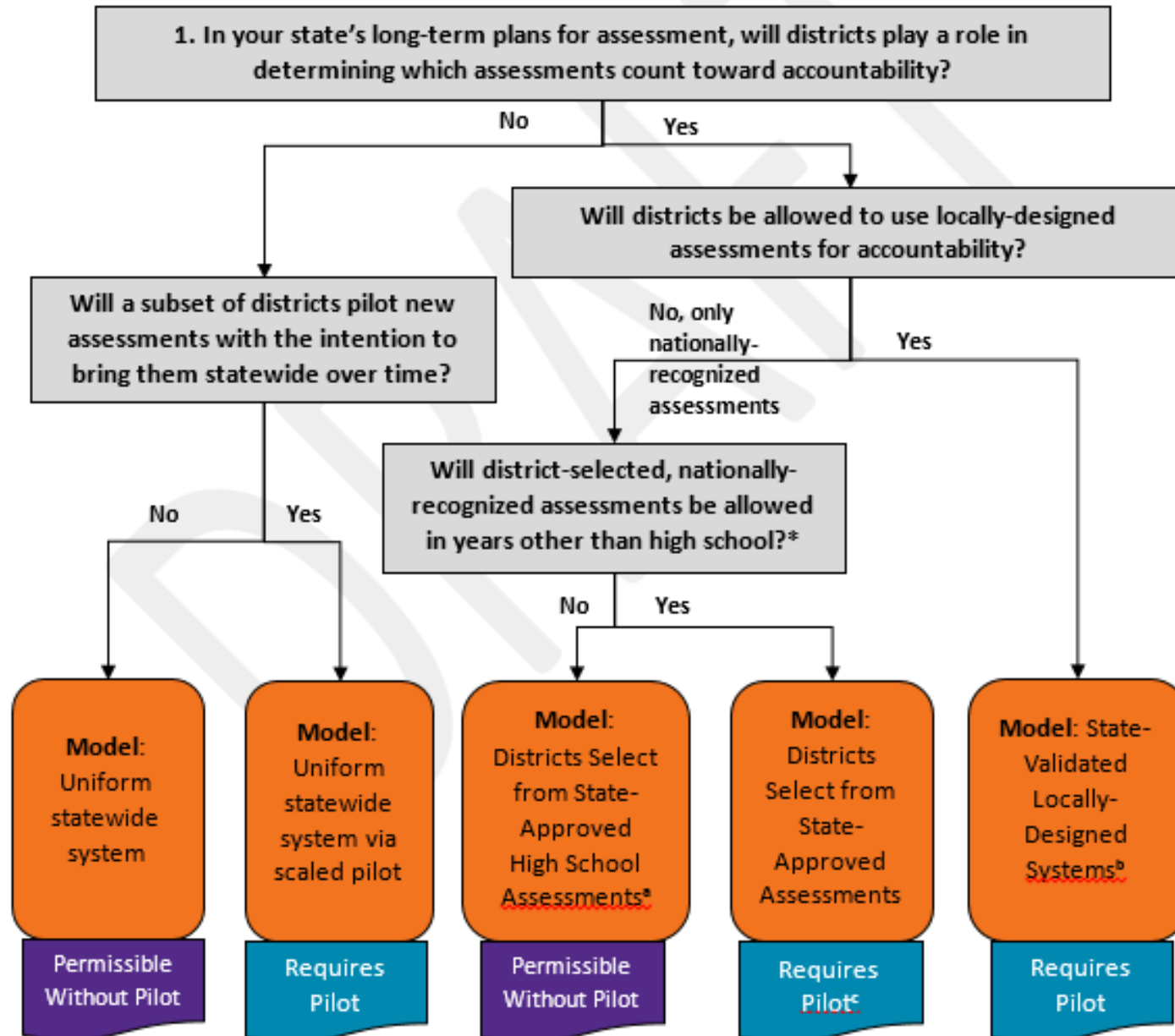
To innovative and continuously improve a comprehensive system of assessments worth taking.



Design Decisions

1. District role in assessment decisions: Will the assessment system be uniform statewide, or will there be some flexibility for districts to select or pilot other forms of assessment?
2. Timing and adaptability of assessments: Will there be flexibility in the timing of when assessments are administered? Will assessment items be tailored to each student's current level of mastery?
3. Types of assessments: What types or forms of assessment will comprise the assessment system?

Do you need to apply for the Demonstration Authority?



Why would states want to apply?

States may want to apply for the Demonstration Authority if:

- ➔ They intend to develop a new approach for assessing meaningful student learning outcomes, **AND**
- ↳ They will pilot in a subset of schools or districts before going statewide, **AND**
- ↳ They will use the innovative assessments for accountability (meaning: ***pilot districts will not need to double-test students***)

Statutory Highlights

May Pilot in a Subset of Districts

- Approved states may pilot with a subset of districts before scaling the system statewide by the end of the Demonstration Authority.

Can Be Entirely Performance-Based

- Approved states may design an assessment or system of assessments that consists of all performance tasks, portfolios, or extended learning tasks.

Can Administer when Students are Ready*

- Approved states may assess students when they are ready to demonstrate mastery of standards and competencies as applicable.*

**However, systems must be able to report grade-level proficiency for accountability*

Timeline for Regulatory Process

Draft Regulations
July 11, 2016

**60 Day Comment
Period Closes**
**September 9,
2016**

Final Regulations
**Before End of
Administration**

Application
**Next
Administration?**
**When a cohort of
states is ready?**

Proposed Regulatory Clarifications

- **Definitions:** “Innovative Assessment System” and “Demonstration Authority”
- **Readiness:** A peer review panel will determine readiness based on application requirements AND selection criteria
- **Flexibility:** Waives these provisions from Sec. 1111:
 1. assessment must be statewide,
 2. assessments must occur every year in grades 3-8 and once in grades 9-12 (states would have to administer statewide assessment in off years)
- **Accountability:** Must produce comparable determinations for State’s long-term goals & academic achievement indicator. Apply consistent identification methodology for schools.
- **Transition:** Will require peer review sign-off that the system and each assessment meets the requirements in Sec. 1111.

Substantive Proposals

- Timeline
- Comparability
- Selection Criteria

Substantive Proposal 1 - Timeline

A state must be ready to implement its new system in a subset of districts immediately upon approval and use those results for accountability purposes.*

7 states may apply initially and each will propose its timeline up to 5 years.

After year three, IES will publish a progress report on the initial states.

USED may extend the demonstration authority to additional states.

A state may request a 2-year extension.

A state may request an additional 1 year* waiver.

USED (with peer review) will determine whether a state can transition permanently.

* Proposed in USED draft regulations

Substantive Proposal 2 - Comparability

A state will need to demonstrate that **its innovative assessment system and the results** of that system **are comparable to the Statewide assessment.**

States can demonstrate this one of four ways:

- 1) Assess all students using an existing State assessment at least once in each grade span
- 2) Assess a sample of students in the same school year on both the innovative and State assessment
- 3) Incorporate common items on both innovative and statewide assessments
- 4) An alternative method that will provide for an equally rigorous and statistically valid comparison

Substantive Proposal 3 – Selection Criteria

In reviewing applications, the Secretary will consider whether a state has met all of the statutory application requirements **and** sufficiently addressed these criteria:

- 1) Project Narrative
- 2) Prior Experience, Capacity, Stakeholder Support
- 3) Timeline and Budget
- 4) Supports for Educators and Students
- 5) Evaluation and Continuous Improvement