

What does it mean to be “High Quality”?



Criteria for High-Quality Assessment

25

Five Criteria for High-Quality Assessment

1. Assessment of Higher-Order Cognitive Skills
2. High-Fidelity Assessment of Critical Abilities
3. Standards that Are Internationally Benchmarked
4. Items that Are Instructionally Sensitive and Educationally Valuable
5. Evidence of Validity, Reliability, and Fairness



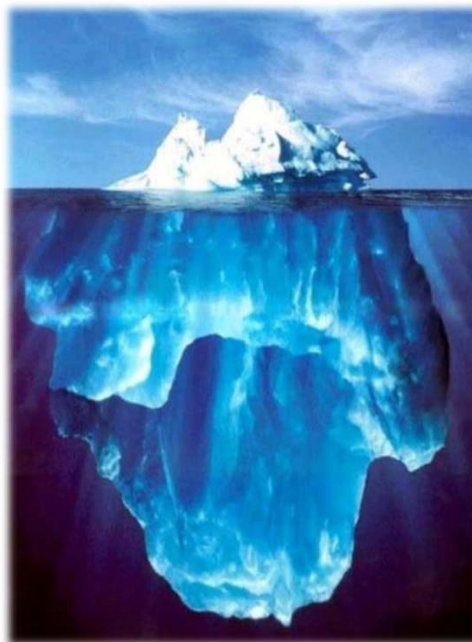
CCSSO Criteria for High Quality Assessments

- Support states as they “develop procurements and evaluate options for high-quality state summative assessments aligned to their college- and career readiness standards.”
- Grouped into five broad categories:
 - A. Meet Overall Assessment Goals and Ensure Technical Quality
 - B. Align to Standards – English Language Arts/Literacy
 - C. Align to Standards – Mathematics
 - D. Yield Valuable Reports on Student Progress and Performance
 - E. Adhere to Best Practices in Test Administration

A. Meet Overall Assessment goals and Ensure Technical Quality

- Indicating progress toward college and career readiness
- Ensuring that assessments are valid and required for intended purposes
- Ensuring that assessments are reliable
- Ensuring that assessments are designed and implemented to yield valid and consistent test score interpretations within and across years
- Providing accessibility to all students, including English learners and students with disabilities
- Ensuring transparency of test design and expectations
- Meeting all requirements for data privacy and ownership

QUESTIONS RELATED TO THE TRANSITION TO ASSESSMENT ALIGNED TO COLLEGE AND CAREER READY STANDARDS



35

Why did proficiency rates drop this year in many states?

- Given differences between the most recent tests used by many states and their previous state tests, this should actually be expected.
- The visible “drop” in proficiency is not actually a drop.
- What we see arises from increasing expectations for student achievement and relatively little change in student performance, proficiency, or school effectiveness.
 - States have adopted more challenging academic standards and raised expectations for what students should know and be able to do when they graduate from high school.
 - If states had maintained their former achievement expectations, students would have performed at least as well as students in previous years.
- The new standards and expectations for student achievement better reflect the demands of college and careers.

36

What do proficiency rates tell us? What do the new scores mean?

- Proficiency rates provide insight into the rigor of the test, student performance, and the status of implementation of new content standards.
 - Assessment results provide information about how students perform on the new content standards and expectations for achievement
- The new scores indicate whether and to what extent, students are on track to be successful in college and careers

37

Can results from a test aligned to new standards be compared to results from previous years?

- It is not possible to make a direct or simple comparison between state results on a new assessment and results on your past assessment.
 - The change in assessments, scales, and achievement standards represents a clean break from the past assessment
- Even when statistical linking occurs, interpreting student performance on the new test in terms of old achievement levels and scales is not appropriate because the assessment aligns to new expectations.

w

Discussion Topics

- The Nature of Educational Assessment
- Federal Law, College Readiness Standards, and High Quality Assessments
- **Considering a Balanced/Comprehensive Assessment System as Texas and Its Districts Move Forward**

39

Need for a “Theory of Action”

A common problem at state and/or district levels is that the assessment components are not conceptually coherent.

They often conflict and as a consequence their use doesn't lead to the desired outcomes of educational improvement.

It is essential to make **EXPLICIT** one's assumptions and “theory of action” for the system of assessments.

A Theory of Action

What is it?

- An empirically and logically stated argument
- A set of underlying assumptions
- A testable hypothesis

...that outlines how and why a given assessment, system, or program, as designed, will support the achievement of specified goals.

- It requires specification of each component of the assessment/evaluation system, the connection between components, and the manner in which they jointly fulfill the requirements of the system.

Pieces of the TOA Puzzle for a CAS

- Purpose
- Theory of Learning
- Prioritized Goals of the System
- Intended Uses of Results

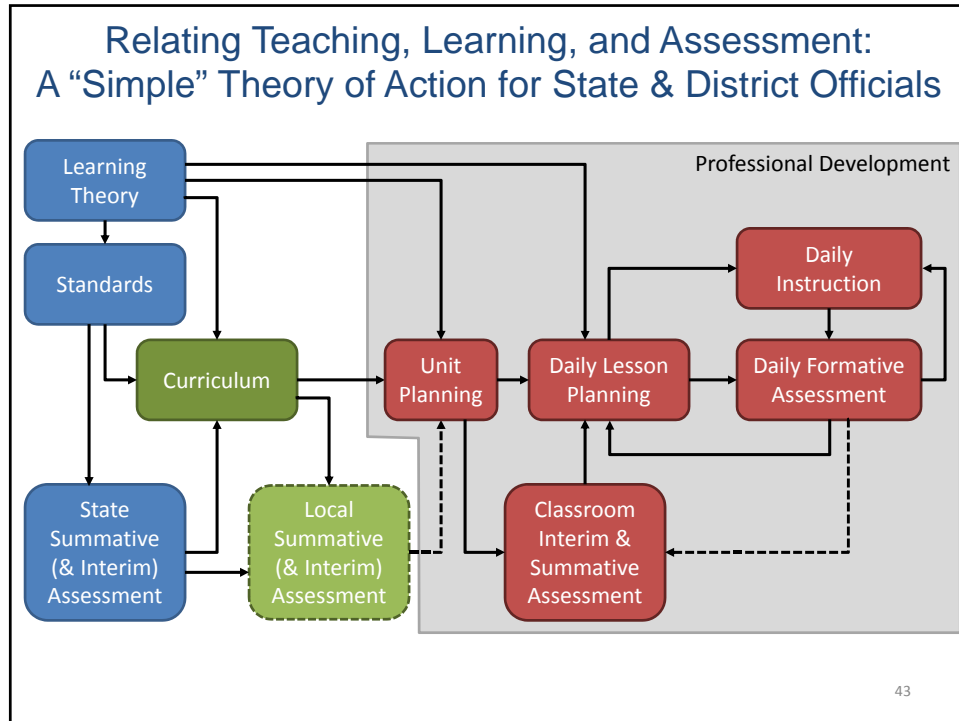


Must be well articulated prior to assessment system design.

- Overarching Theory as to manner in which the assessment system will bring about desired change (Key Design Principles)
- Design of the system and its component parts
 - Assessments, Tasks
 - Alignment of each component to goals/intended uses/Key Design Principles
- Mechanism by which component are intended to provide for specified goals.
- Expected relationship among components
- Inferences/assumptions underlying the system working as intended.

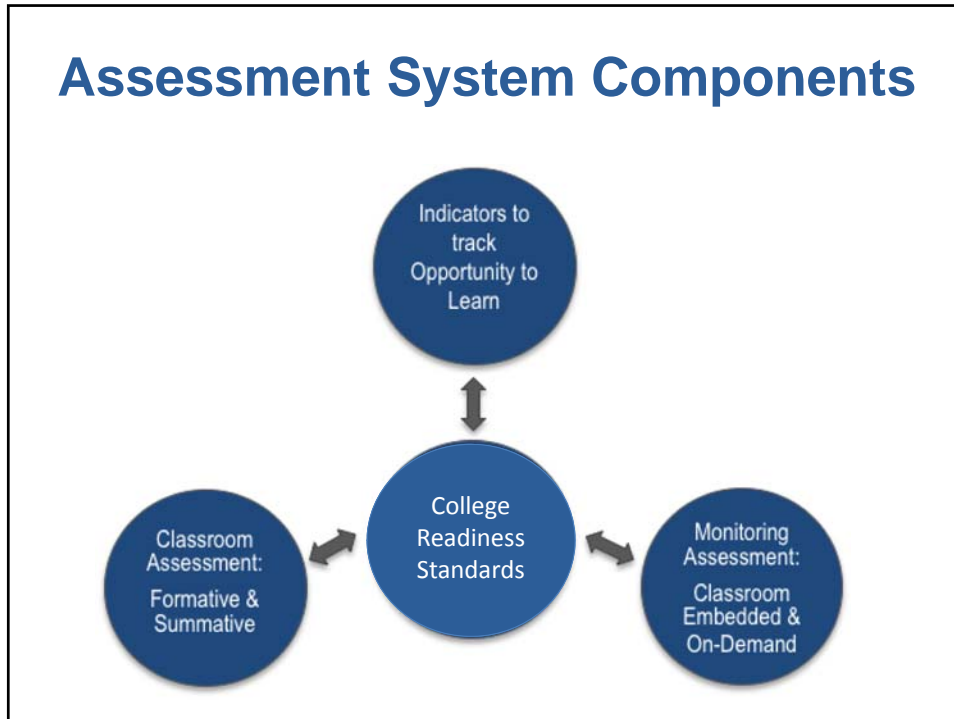


Articulated as part of assessment system design.



Need a Coherent System of Assessments – 3 Major Components

- A system of assessments should include classroom assessment, monitoring (large-scale) assessments, and indicators of opportunity to learn.
 - **Classroom assessment should be an integral part of instruction and should reinforce the type of learning envisioned in standards.**
 - **Monitoring (large-scale) assessments will need to include an on-demand component and a component based in the classroom (classroom-embedded) in order to fully cover the breadth and depth of the standards.**
 - **Indicators of opportunity to learn should document that students have the opportunity to learn in the way called for in standards and that schools have appropriate resources.**



**1st Major Challenge in
Design of the Monitoring
Component:
Intended uses of the Information**

The Complex Space of Monitoring Functions

TABLE 5-1 Questions Answered by Monitoring Assessments

Types of inferences	Levels of the Education System			
	Individual Students	Schools or District	Policy Monitoring	Program Evaluation
Criterion-referenced	Have individual students demonstrated adequate performance in science?	Have schools demonstrated adequate performance in science this year?	How many students in state X have demonstrated proficiency in science?	Has program X increased the proportion of students who are proficient?
Longitudinal and comparative across time	Have individual students demonstrated growth across years in science?	Has the mean performance for the district grown across years? How does this year's performance compare to last year's?	How does this year's performance compare to last year's?	Have students in program X increased in proficiency across several years?
Comparative across groups	How does this student compare to others in the school/state?	How does school/district X compare to school/district Y?	How many students in different states have demonstrated proficiency in science?	Is program X more effective in certain subgroups?

2nd Major Challenge in Design of the Monitoring Component: Possible Sources of Evidence

State Assessments for Monitoring

Combine two types of external assessment strategies, in conjunction with OTL indicators:

On-Demand Assessments

- Developed by the state
- Administered at a time mandated by the state

Classroom-Embedded Assessments

- Developed by the state or district,
- Administered at a time determined by the district/school that fits the instructional sequence in the classroom

Possible Options for the On-Demand Assessment Components

- Mixed item formats, including extended constructed response
 - Such as AP exams
- Mixed item formats with performance tasks
 - might involve both group and independent activities
 - might involve some hands-on tasks
- Use matrix sampling, depending on the intended use and the need to report scores for individuals versus for groups.

Possible Options for the Classroom-Embedded Assessment Components

- **Replacement units** (curriculum materials + assessments) developed outside of the classroom (by state or district)
- **Item banks of tasks**, developed outside of the classroom
- **Portfolio collections of work samples**, with tasks specified by state or district

How might states and districts organize the different assessments that they seem to need?

- ❑ **Desired end product is a multilevel system**
 - Each level fulfills a clear set of functions and has a clear set of intended users of the assessment information
 - The assessment tools are designed to serve the intended purpose
 - Formative, interim, summative
 - Design is optimized for function served
- ❑ **The levels are articulated and conceptually coherent**
 - They share the same underlying concept of what the targets of learning are at a given grade level and what the evidence of attainment should be.
 - They provide information at a “grain size” and on the “time scale” appropriate for translation into action.

What are the key design elements of such a comprehensive system?

- ❑ The system is designed to track progress over time
 - At the individual student level
 - At the aggregate group level

- ❑ The system uses tasks, tools, and technologies appropriate to the desired inferences about student achievement
 - Doesn't force everything into a fixed testing/task model
 - Uses a range of tasks: performances, portfolios, projects, fixed- and open-response tasks as needed

What else is needed for change to occur?

Much of the change in the productive use of assessment requires training in the use of new tools and systems

- A substantial professional development effort is needed across levels of the system
 - Teachers, principals, and district leadership
- Processes for the effective collection and use and interpretation of assessment information need to be implemented
 - Focus of many assessment literacy efforts
- New technologies and data systems may need to be created and accommodated in the system's business practices

One Vision for how integration of assessment and instruction happens in districts and schools

District Level

- District has a vision for high quality teaching and learning.
- High quality tasks are embedded into the K-12 curriculum.
- Standards Based Grading is aligned.
- District supports ongoing professional learning for staff.

School Level

- Common collaboration for grade level teams is in place.
- Administrators support this work.

Classroom Level

- Students engage in ongoing problem solving and challenging assessments.
- Teachers engage in formative assessment processes.

What else is needed to make assessment useful in promoting student achievement?

Assessment Should not be the
“Tail that Wags the Educational Dog”

