Basics of Assessment

Webinar Series Part 2
Logistics

★ Q&A
★ Type your questions into the Chat box on the lower right-hand corner of the screen

★ Recording
★ Each part of the webinar series will be recorded and archived.

★ Supplemental materials
★ All materials needed to participate in each session will be posted on the sign-in page.
# Webinar Series

<table>
<thead>
<tr>
<th>Title</th>
<th>Date</th>
<th>Length</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction: District-Determined Measures and Assessment Literacy</td>
<td>3/14</td>
<td>60 minutes</td>
<td>4-5pm ET</td>
</tr>
<tr>
<td><strong>Basics of Assessment</strong></td>
<td>4/4</td>
<td>90 minutes</td>
<td><strong>4-5:30pm ET</strong></td>
</tr>
<tr>
<td>Assessment Options</td>
<td>4/25</td>
<td>60 minutes</td>
<td>4-5pm ET</td>
</tr>
<tr>
<td>TA and Networking Session I</td>
<td>5/23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determining the Best Approach to District-Determined Measures</td>
<td>7/18</td>
<td>60 minutes</td>
<td>4-5pm ET</td>
</tr>
<tr>
<td>Integrating Assessments into Educator Evaluation: Reporting Student Growth</td>
<td>8/15</td>
<td>60 minutes</td>
<td>4-5pm ET</td>
</tr>
<tr>
<td>Integrating Assessments into Educator Evaluation: Developing Business Rules and Engaging Staff</td>
<td>8/29</td>
<td>60 minutes</td>
<td>4-5pm ET</td>
</tr>
<tr>
<td>TA and Networking Session II</td>
<td>9/19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communicating results</td>
<td>10/24</td>
<td>60 minutes</td>
<td>4-5pm ET</td>
</tr>
<tr>
<td>Sustainability</td>
<td>12/5</td>
<td>60 minutes</td>
<td>4-5pm ET</td>
</tr>
<tr>
<td>TA and Networking Session III</td>
<td>12/12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Audience & Purpose

★ Target audience
★ District teams that will be engaged in the work of identifying and selecting District-Determined Measures

★ Purpose
★ Participants will:
★ develop a working knowledge of the key concepts of assessment relevant for using assessments to measure student growth
★ learn about different types of assessments, how to determine assessment alignment, and elements of assessment quality
Agenda

- District Determined Measures
- Assessment Overview
- Types of Assessments
- Alignment
- Assessment Components
- Assessment Quality
District Determined Measures

- Measures of student learning, growth, and achievement related to the Massachusetts Curriculum Frameworks, Massachusetts Vocational Technical Education Frameworks, or other relevant frameworks, that are comparable across grade or subject level district-wide.

- Pre and post unit and course assessments
- Approved commercial assessments
- Portfolios
- Capstone projects
## Student Impact Rating

- **Student Impact Rating** must be based on at least 2 years of data (trends) across multiple measures (patterns):
  - State-wide growth measures
    - MCAS student growth percentiles
  - District-Determined Measures

### Part VII: Rating Educator Impact on Student Learning

#### Using District-Determined Measures of Student Learning

<table>
<thead>
<tr>
<th>Year</th>
<th>Measure</th>
</tr>
</thead>
</table>
| Year 1 | MCAS SGP, grade 5 mathematics  
Unit assessment on multiplication and division of fractions |
| Year 2 | MCAS SGP, grade 5 mathematics  
Unit assessment on multiplication and division of fractions |
The Opportunity

★ Identifying DDMs can be the impetus for broadening and strengthening the district’s assessment practices.

★ DDMs will provide educators with useful data that will help them improve both student outcomes and their instructional practices.

★ DDMs will yield data educators can use throughout the 5-step evaluation cycle.
Agenda

- District-Determined Measures
- **Assessment Overview**
- Types of Assessments
- Alignment and Rigor
- Assessment Components
- Assessment Quality
Assessment Overview

★ **Assessment** is a general term that refers to the process of gaining information about student learning
  - Process includes administration procedures, scoring, reporting, etc.
  - A DDM is an assessment

★ **Instrument** refers to a specific type of data collection tool or mechanism used in an assessment process
  - There are many types of instruments
  - A test is one example
Value of Good Assessment

Better assessment

Better teaching

Better learning and greater confidence

Better student outcomes

Better opportunities in life
Assessment Approaches

- **Indirect**
  - Gather information from sources other than actual samples of student work

- **Direct**
  - Gather information from actual samples of student work
Agenda

★ District-Determined Measures
★ Assessment Overview
★ Types of Assessments
★ Alignment and Rigor
★ Assessment Components
★ Assessment Quality
Types of Assessments

On-Demand

Performance, Project

Portfolio

Hybrid

Summative Assessments:
EOC, EOY, Interim, Capstone
On-Demand Assessments

- An assessment that takes place at a predetermined time and place, usually under standard conditions for all students being assessed

  - E.g., SAT, district and state tests, and most in-class unit tests and final exams
Assessments based on observations of behaviors or based on work performed on a complex activity
- Natural vs. Artificial
- Unstructured vs. Structured
- Participant vs. External Observer
- Self-rated vs. Other-rated (teacher, peer, observer)
Portfolio

★ A purposeful and systematic collection of student work
★ Should include:
  ★ student participation in the selection of portfolio content,
  ★ the criteria for selection are aligned to standards and grade-level expectations through a rubric or other scoring device,
  ★ the criteria for judging merit, and
  ★ evidence of student self-reflection
★ May include both finished work (Product) and work in progress (Process)
  ★ When using portfolios for DDMs, include the student’s finished products whenever possible
★ May focus on one or more curricular areas
Hybrid Assessment

★ An on-demand assessment that combines of two or more types of assessments
★ Usually a paper-and-pencil or online test with a performance, portfolio, or project assessment
Reflection #1

★ Take a minute to jot down sources of existing assessments that might be used for DDMs.
Agenda

- District-Determined Measures
- Assessment Overview
- Types of Assessments
- Alignment and Rigor
- Assessment Components
- Assessment Quality
Alignment and Rigor

- Alignment refers to the extent to which the assessment aligns with curriculum as expressed in the curriculum map.
- Rigor is the level of cognitive complexity of the item or of a set of items.
  - Bloom’s revised taxonomy or other taxonomy.
- Understanding alignment and rigor is critical for selecting or developing an assessment.
- Documentation is the key!
  - E.g., Table of Test Specifications.
Alignment

- DDMs reflect key learning objectives by grade and content area
  - Information on key objectives is found in the district’s curricular maps and other curricular planning tools
Alignment

★ Identify the key content you want to assess

★ Standards
  ★ E.g., Use trigonometric ratios and the Pythagorean Theorem to solve right triangles in applied problems (Mathematics.G.SRT.3.08)

★ Learning objectives
  ★ E.g., Students will correctly apply Pythagorean Theorem when prompted.
  ★ E.g., Students determine when to correctly apply trigonometric ratio models.
## Table of Test Specifications

<table>
<thead>
<tr>
<th>Cognitive Complexity</th>
<th>Lower-Order</th>
<th>Higher-Order</th>
<th>Subtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mathematics.G.SRT.3.08</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LO #1: Students will correctly apply Pythagorean Theorem when prompted.</td>
<td>25%</td>
<td>30%</td>
<td>55%</td>
</tr>
<tr>
<td>LO: #2 Students determine when to correctly apply trigonometric ratio models.</td>
<td>25%</td>
<td>20%</td>
<td>45%</td>
</tr>
<tr>
<td>SubTotal</td>
<td>50%</td>
<td>50%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Rigor – Revised Bloom’s Taxonomy

- Remembering
- Understanding
- Applying
- Analyzing
- Evaluating
- Creating
Rigor – Revised Bloom’s Taxonomy

The diagram below represents a ramp and some of its dimensions.

What is the height, $h$, of the ramp?
Rigor – Revised Bloom’s Taxonomy

True/False: Will a set of skis that are 6’ (six feet) high fit into an empty closet with the following dimensions?

<table>
<thead>
<tr>
<th>Dimensions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Feet</td>
<td>Wide</td>
</tr>
<tr>
<td>3 Feet</td>
<td>Deep</td>
</tr>
<tr>
<td>5 Feet</td>
<td>High</td>
</tr>
</tbody>
</table>

Applying
Reflection #2

★ Think about the potential DDMs you wrote down in Reflection #1.

★ Of these, consider which one shows the best alignment to the curriculum and the degree of rigor.

★ Identify one that is promising but that does not show perfect alignment. Consider how could it be improved.
Agenda

- District-Determined Measures
- Assessment Overview
- Types of Assessments
- Alignment and Rigor
- Assessment Components
- Assessment Quality
Assessment Components

- Table of Test Specifications
- Administration Protocol
- Instrument
- Scoring Method
- Documentation
Administration Protocols

- Often found in Administration Manuals
- Needed to ensure all students have a fair opportunity to demonstrate what they know and can do in the assessment

Proctoring directions
- All examinees receive same set of directions and conditions for taking the instrument (including resources – e.g., calculator – breaks, pacing, etc.)

Security provisions
- Ensure specific items are not overly familiar to examinees and proctors (unless portfolio assessment is used)

Student accommodations
- Ensure the majority of students can participate in the program, with appropriate supports
Items

★ Selected Response
  ★ True–False
  ★ Multiple Choice
  ★ Matching

★ Constructed Response
  ★ Short answer
  ★ Restricted constructed response
  ★ Extended constructed response (includes essay)
  ★ Portfolio item
  ★ Performance item
Selected Response Item

Which of the following could not be the exact number of right angles in a quadrilateral?

A. 1  
B. 2  
C. 3  
D. 4

The Stem and Stimulus:
Succinctly describes the problem for examinees.

The Options:
Provide a correct answer (C) and incorrect answers that will "distract" examinees who do not know the material.
One lap around a track is $\frac{1}{4}$ mile. On Monday, Stacy ran 11 laps.

a. What was the total number of miles that Stacy ran on Monday? Show or explain how you got your answer.

b. It took Stacy $\frac{1}{2}$ hour to run 11 laps on Monday. What was Stacy’s average speed, in miles per hour, on Monday? Show or explain how you got your answer.

c. Stacy’s goal is to run at an average speed of 1 mile per 10 minutes. What is the number of laps that she must run in $\frac{1}{2}$ hour to reach her goal? Show or explain how you got your answer.
Scoring Items

★ Scoring objective items
  ★ Scoring key or short guide
  ★ Based on clearly defined scoring key and set of scoring rules
  ★ Limits error variance

★ Scoring subjective items
  ★ Longer scoring guide with rubrics or calibrated scoring papers
  ★ Based on personal judgment
  ★ Increases potential for error
Sample Holistic Rubric

In 200 words or less, describe how you would explain to a homeowner the concept of eminent domain and how it is related to the Fifth Amendment.

<table>
<thead>
<tr>
<th>Points</th>
<th>Student Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>This is a superior piece of writing. The prompt is directly addressed, and the response is adapted effectively to audience and purpose. It is exceptionally developed, containing compelling ideas, examples and details. The response, using a clearly evident organizational plan, actively engages the reader with a unified and coherent sequence and structure of ideas. The response consistently uses a variety of sentence structures, effective word choices and an engaging style.</td>
</tr>
<tr>
<td>5</td>
<td>This is an excellent piece of writing. The prompt is directly addressed and the response is clearly adapted to audience and purpose. It is very well developed, containing compelling ideas, examples and details. The response, using a clearly evident organizational plan, engages the reader with a unified and coherent sequence and structure of ideas. The response typically uses a variety of sentence structures, effective word choices and an engaging style.</td>
</tr>
<tr>
<td>4</td>
<td>This is an effective piece of writing. While the prompt is addressed and the response adapts to audience and purpose, there are occasional inconsistencies in the response’s overall plan. The response is well developed, containing effective ideas, examples and details. The response, using a good organizational plan, presents the reader with a generally unified and coherent sequence and structure of ideas. The response often uses a variety of sentence structures, appropriate word choices and an effective style.</td>
</tr>
</tbody>
</table>
### Sample Analytic Rubric

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>The student response demonstrates an exemplary understanding of the Number Sense and Operations concepts involved in selecting and using appropriate operations to solve problems with rational numbers. The student multiplies a fraction by a whole number, multiplies a whole number by a mixed number, and then divides a whole number by a fraction.</td>
</tr>
<tr>
<td>3</td>
<td>The student response demonstrates a good understanding of the Number Sense and Operations concepts involved in selecting and using appropriate operations to solve problems with rational numbers. Although there is significant evidence that the student was able to recognize and apply the concepts involved, some aspect of the response is flawed. As a result the response merits 3 points.</td>
</tr>
<tr>
<td>2</td>
<td>The student response demonstrates fair understanding of the Number Sense and Operations concepts involved in selecting and using appropriate operations to solve problems with rational numbers. While some aspects of the task are completed correctly, others are not. The mixed evidence provided by the student merits 2 points.</td>
</tr>
<tr>
<td>1</td>
<td>The student response demonstrates only minimal understanding of the Number Sense and Operations concepts involved in selecting and using appropriate operations to solve problems with rational numbers.</td>
</tr>
<tr>
<td>0</td>
<td>The student response contains insufficient evidence of an understanding of the Number Sense and Operations concepts involved in selecting and using appropriate operations to solve problems with rational numbers to merit any points.</td>
</tr>
</tbody>
</table>

**MCAS Test Question**  
Grade 8  
2012  
Question 10
Calibrated Scoring Paper

Scoring Guide - Score Point 4

a. Stacy ran \(2 \frac{3}{4}\) miles.
\[11 \times \frac{3}{4} = \frac{33}{4} = 2 \frac{3}{4}\]

b. Stacy's average speed = 5.5 mph.
\[\frac{1}{2} \times 2 \frac{3}{4} = \frac{1}{4} \times \frac{2}{1} = \frac{22}{4}\]
\[\times 2\]
\[\times \frac{3}{2}\]
\[1 : 5 \frac{1}{2}\]

c. Laps = 12 laps
\[\times 1\text{ mile per 10 min.} \times \frac{3}{2}\]
\[\times \text{miles per 30 min}\]
\[3 \text{ miles} \div \frac{1}{4}\]
\[3 \times \frac{4}{1} = \frac{12}{1}\]
Reporting, Interpretation and Use

★ Reports provide information and conclusions from and based on the assessments
  ★ May provide information in the form of text, graphs, images, etc…

★ A goal of any assessment is to yield information that will help educators make appropriate decisions or driving meaningful change

★ Assessments are most useful when reports tied to objectives and are easy to understand and use
# Simple Score Report

<table>
<thead>
<tr>
<th>Student</th>
<th>PreTest</th>
<th>PostTest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Item1</td>
<td>Item2</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Documentation

- Technical Manual
- Administration Manual
Reflection #3

In summary, assessments are composed of:

- Table of Test Specifications
- Administration Protocol
- Instrument
- Scoring Method
- Documentation

Reflect for a moment on one of the assessments you’re considering for use as a DDM. Identify the components you have in place and those you’d want to develop.
Agenda

★ District-Determined Measures
★ Assessment Overview
★ Types of Assessments
★ Alignment and Rigor
★ Assessment Components
★ Assessment Quality
Assessment Quality

- Reliability
- Validity
- Fairness and Non-Bias
- Item Quality
- Feasibility
Reliability

★ Degree of consistency in measurement
★ We want to have confidence that scores are stable
★ Example: Weighing yourself on a scale
Reliability

- Four typical approaches
  - Internal consistency
  - Test-retest
  - Alternate forms or split-half
  - Inter-rater agreement

- Reliability coefficients are estimated using statistical formulas
  - We cannot “see” reliability
  - Ranges from 0 (no reliability) to 1 (perfect reliability)
Validity

- Validity refers to the *validity of inferences* made about assessments or based on assessment data.
- Gives you confidence that what you say about student assessment scores and therefore about students is justified.
- Example: Weighing yourself on two different kinds of scales.
Validity Based on Content

- For existing measures, districts review content in instrument and judge whether it matches curriculum (review of alignment and rigor)

<table>
<thead>
<tr>
<th>Cognitive Taxonomy</th>
<th>Lower-Order</th>
<th>Higher-Order</th>
<th>Subtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics.G.SRT.3.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LO #1: Students will correctly apply Pythagorean Theorem when prompted.</td>
<td>25%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LO: #2 Students determine when to correctly apply trigonometric ratio models.</td>
<td>25%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SubTotal</td>
<td></td>
<td></td>
<td>50%</td>
</tr>
</tbody>
</table>

**True/False:** Will a set of skis that are 6’ (six feet) high fit into an empty closet with the following dimensions?

<table>
<thead>
<tr>
<th>Dimensions:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Feet</td>
<td>Wide</td>
</tr>
<tr>
<td>3 Feet</td>
<td>Deep</td>
</tr>
<tr>
<td>5 Feet</td>
<td>High</td>
</tr>
</tbody>
</table>
Validity Based on Relationships

★ Assessment should show:

★ Moderate to strong and positive correlations with similar instruments/outcomes
★ Low positive or even negative correlations with dissimilar instruments/outcomes

★ Correlation = A statistical technique that is used to measure and describe the strength and direction of the relationship between two variables
★ Range from -1 to +1
Consequential Validity

- Realization of benefits
  - Student learning
  - Teacher improvement
- Minimization of negative consequences
  - Poor student or teacher attitudes toward the assessment or assessments generally
  - Limiting instruction only to the content covered in the instrument
  - Improper use of scores
Fairness and Non-Bias

★ Fairness
★ All examinees have equal opportunity to demonstrate knowledge on assessment

★ Non-Bias
★ Students with similar ability receive similar scores, regardless of group membership
Item Quality

★ Item quality is a key to assessment quality
★ We typically look at three things:

★ Difficulty
  ★ Ensure a range of difficulty (e.g., easy, medium, hard) in items
  ★ Average difficulty aligns to assessment purpose and target population

★ Discrimination
  ★ Ensure that people who got higher scores on the instrument overall tend to get higher scores on that item

★ Guessing
  ★ Reduce guessing by writing good response options for selected response items (e.g., multiple-choice items)
Feasibility

★ Cost
★ Technology
  ★ E.g., Paper and pencil, online, adaptive
★ Assessment length
★ Reports
  ★ E.g., access to, ability to interpret
★ Accommodations and accessibility
Reflection #4

★ Based on all that we discussed today, identify some basic “do’s and don’ts” about assessment that you need to consider when selecting or building a DDM.
Resources

★ Association of Test Publishers
  ★ http://www.testpublishers.org/

★ Buros Mental Measurements Yearbook and Tests in Print
  ★ http://buros.org/mental-measurements-yearbook
  ★ http://buros.org/tests-print

★ APA, NCME, and AERA Standards for Educational and Psychological Testing

★ National Council on Assessment in Education
  ★ http://ncme.org/
Register for Webinar 3
Assessment Options

★ April 25th from 4:00 PM – 5:00 PM

★ Identify one or more potential DDMs, create a map of assessment gaps appropriate for accountability purposes, develop a preliminary action plan to bring to the Technical Assistance and Networking Session

★ Click here to register: https://air-event500.webex.com/air-event500/onstage/g.php?d=590644030&t=a
Questions

Contact Ron Noble at r noble@doe.mass.edu

Feedback

Tell us how we did: