Waiting Room

Today’s webinar will begin shortly.

REMINDERS:

- Dial 800-503-2899 and enter the passcode 6496612# to hear the audio portion of the presentation

- Download today’s materials from the sign-in page:
  - Webinar Series Part 8 PowerPoint slides
Communicating Results
Webinar Series Part 8
<table>
<thead>
<tr>
<th>#</th>
<th>Title</th>
<th>Date</th>
<th>Length</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction: DDMs and Assessment Literacy</td>
<td>3/14</td>
<td>60 minutes</td>
<td>4-5pm</td>
</tr>
<tr>
<td>2</td>
<td>Basics of Assessment</td>
<td>4/4</td>
<td>90 minutes</td>
<td>4-5:30pm</td>
</tr>
<tr>
<td>3</td>
<td>Assessment Options</td>
<td>4/25</td>
<td>60 minutes</td>
<td>4-5pm</td>
</tr>
<tr>
<td>4</td>
<td>TA and Networking Session I</td>
<td>7/11</td>
<td>3 hours</td>
<td>9am-12pm</td>
</tr>
<tr>
<td>5</td>
<td>Determining the Best Approach to DDMs</td>
<td>7/18</td>
<td>60 minutes</td>
<td>4-5pm</td>
</tr>
<tr>
<td>6</td>
<td>Measuring Student Growth and Piloting DDMs</td>
<td>8/15</td>
<td>60 minutes</td>
<td>4-5pm</td>
</tr>
<tr>
<td>7</td>
<td>Integrating Assessments into Educator Evaluation: Developing Business Rules and Engaging Staff</td>
<td>10/24</td>
<td>60 minutes</td>
<td>4-5pm</td>
</tr>
<tr>
<td>8</td>
<td>Ramping up for Next Year: Strategies for Using Current Assessments in DDMs.</td>
<td>12/5</td>
<td>60 minutes</td>
<td>4-5pm</td>
</tr>
<tr>
<td>9</td>
<td>TA and Networking Session III</td>
<td>12/12</td>
<td>3 hours</td>
<td>2:30pm-5:30pm</td>
</tr>
<tr>
<td>10</td>
<td>Communicating Results</td>
<td>1/23</td>
<td>60 minutes</td>
<td>4-5pm</td>
</tr>
<tr>
<td>11</td>
<td>Sustainability</td>
<td>2/27</td>
<td>60 minutes</td>
<td>4-5pm</td>
</tr>
</tbody>
</table>
Audience & Purpose

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

★ After today participants will understand:
★ Approaches to Scoring and Setting Parameters
★ Approaches to using results from DDMs to inform practice and evaluation.
Key Messages for Stakeholders

Four Key Messages about Student Impact Ratings

Use Multiple Measures
DDMs and SGP are part of comprehensive evaluation system

Focus on Students
The focus of measuring student impact is improving student learning.

Build Capacity
Developing DDMs builds knowledge about assessment and data use.

Engage Educators
Educators have expertise developing and evaluating assessments.
DDMs in a Backwards Design Model*

★ DDMs require us to define clearly what we want students to be able to do following instruction.

★ After we have clearly defined what we want students to be able to do, then we plan curriculum and lessons to get us to that goal

★ Integration

★ Mass Frameworks: What should students learn?
★ DDMs: How do students demonstrate that learning?
★ Curriculum Mapping: How do we get there?

* Understanding by Design 2.0 © 2011 Grant Wiggins and Jay McTighe
DDMs in a Backwards Design Model

★ Engage Educators in the processes of determining scoring protocols and setting parameters for high, moderate, and low student growth

★ If educators are involved in the these processes, DDM results will be more meaningful.
Scoring & Setting Parameters Brief

Upcoming Resource
How do I determine if a student has demonstrated high, moderate, or low growth?
What do I need to determine high, moderate, or low student growth?

1. Clear directions for scoring individual student work.

2. Clear directions for determining a student’s growth.

3. Parameters for high, moderate, and low student growth.

These components may be done independently, or combined.
1. Clear directions for scoring individual student work.
Who scores the DDM?

- **Outside Organizations**
  - Commercial assessments
  - Automated methods
  - Paid raters (e.g., college students, retired teachers)

- **Teams of Teachers** (e.g., all 5th grade teachers)
  - Team members rate each other’s students’ responses
  - Multiple raters score each response

- **Individual Teachers**
  - Random auditing (rechecking)
Scoring Guides are Important

- Support explicitly stating the aspects of student work that are essential.
- Can be shared with students (and parents) for the sake of transparency.

Continuous Improvement Tip:
Plan a review of your scoring guides after the first year of implementation
Scoring Guides Must be Clear

★ Scoring Guide Example:

- 2 points for a correct answer with student work shown correctly
- 1 point for an incorrect answer with student work shown correctly

★ Issue:

★ Not clear around how to score a student with a correct answer with no student work shown or with student work shown incorrectly.
★ Not clear what “shown correctly” means.
Scoring Guides Must be Clear

★ Scoring Guide Example - Improved:

- 2 points for a correct answer with either a chart or table showing how the student set up the problem.
- 1 point for an incorrect answer, but the work demonstrates setting up the problem with a table or picture. Supporting work may include incorrect numbers or other mistakes.
- 1 point for a correct answer and there is no supporting work or if student work is not organized in a table or chart.
- 0 point for no correct answer, and work is not organized in a table or chart.

★ The scoring guide could be further improved by incorporating anchor examples.
## Using Rubrics

<table>
<thead>
<tr>
<th>Rubric Style</th>
<th>Analytic</th>
<th>Holistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student work is assessed by clearly defined criteria along multiple dimensions (i.e., each rows of the rubric assesses a different criterion).</td>
<td>Student work is assessed as a whole product based on an overall impression.</td>
<td></td>
</tr>
</tbody>
</table>
# Using Rubrics

<table>
<thead>
<tr>
<th>Measurement Goal</th>
<th>Achievement</th>
<th>Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Achievement</strong></td>
<td>An <em>individual</em> example of student work is assessed to determine the level of performance or <em>achievement</em> that student has demonstrated.</td>
<td><strong>Growth</strong></td>
</tr>
<tr>
<td><strong>Growth</strong></td>
<td><em>Multiple</em> examples of student work have been assessed to determine the level of <em>growth</em> a student has demonstrated.</td>
<td></td>
</tr>
</tbody>
</table>
## Example Rubrics

<table>
<thead>
<tr>
<th>Rubric Style</th>
<th>Analytic</th>
<th>Holistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Achievement</strong></td>
<td>MCU (p. 30)</td>
<td>MCAS</td>
</tr>
<tr>
<td><strong>Growth</strong></td>
<td>Technical Guide B (pg 15)</td>
<td>TN Fine Arts Growth Measure System (pg 8)</td>
</tr>
</tbody>
</table>
## Growth Rubrics

### Analytic

<table>
<thead>
<tr>
<th>Low Growth</th>
<th>Moderate Growth</th>
<th>High Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 or 3</td>
<td>4-7</td>
<td>8 or more</td>
</tr>
</tbody>
</table>

Number of writing mechanics, such as punctuation, capitalization, misspelled word, where student has corrected the mistake in future writing.

<table>
<thead>
<tr>
<th>Low Growth</th>
<th>Moderate Growth</th>
<th>High Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2 or more</td>
</tr>
</tbody>
</table>

Number of examples of improvement of language usage and sentence formation, such as word order, subject-verb agreement, or run-on sentences where student has corrected the mistake in future writing.

### Holistic

<table>
<thead>
<tr>
<th>Low Growth</th>
<th>Moderate Growth</th>
<th>High Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Little to no improvement in following writing conventions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Average improvement in following writing conventions

High improvement in following writing conventions
## Considerations

<table>
<thead>
<tr>
<th>Rubric Style</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analytic</strong></td>
<td><strong>Holistic</strong></td>
</tr>
<tr>
<td>• Easier to achieve high levels of interrater reliability</td>
<td>• The use of professional judgment supports conclusions</td>
</tr>
<tr>
<td>• Provides formative feedback to students around expectations</td>
<td></td>
</tr>
</tbody>
</table>
Considerations

<table>
<thead>
<tr>
<th>Measurement Goal</th>
<th>Achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Familiar to Educators</td>
</tr>
<tr>
<td></td>
<td>• Easier to design</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>• No additional steps are needed to determine if a student has demonstrated high, moderate, or low growth</td>
</tr>
<tr>
<td>• Use of professional judgment supports conclusions around growth</td>
</tr>
</tbody>
</table>
Achievement vs. Growth

⭐ Achievement Rubrics

- Score Student Work
- Achievement Rubric
- Determine Growth
- Set Parameters

⭐ Holistic Growth Rubrics

- Score Student Work
- Growth Rubric
- Determine Growth
- Set Parameters
2. Clear directions for determining a student’s growth.
Approaches to Measuring Student Growth

- Pre-Test/Post Test
- Repeated Measures
- Holistic Evaluation
- Post-Test Only

Learn more
Webinar 5  http://www.doe.mass.edu/edeval/ddm/webinar.html
Pre-Test/Post Test

★ **Issue: **Floor & Ceiling Effects

★ **Two Tiered Test: Ceiling**
  ★ All students take Qs 1-20 for pretest.
  ★ Students who get first 10 Qs correct take Qs 11-30 for posttest

★ **Adding additional easier items: Floor**

★ **Issue: **Is “Gain” synonymous with “Growth”
  ★ Adding additional moderately challenging items to pre and post.
Repeated Measures

In the example on the left, we see steady growth. In the example on the right, we see a dramatic growth in students between December and March, and very little growth between September and December and March and June.
Holistic Evaluation

⭐ Previously discussed use of growth rubrics
Post-Test Only

Use of Pre Calculus Grade to make predictions around AP Calculus Score

<table>
<thead>
<tr>
<th>Previous Grade</th>
<th>Low Growth</th>
<th>Moderate Growth</th>
<th>High Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1</td>
<td>2-3</td>
<td>4-5</td>
</tr>
<tr>
<td>B</td>
<td>1-2</td>
<td>3</td>
<td>4-5</td>
</tr>
<tr>
<td>A</td>
<td>1-3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Other Approaches: Post-Test Only

★ When to use Post-Test Only
  ★ When pre-test doesn’t provide information
  ★ When post-test is important (Capstone, Research)

★ Key Question: Does this approach foster meaningful conversations later?
Parameters for high, moderate, and low student growth.
Setting Parameters

★ Qualitative Approach
   ★ Engage Teachers
   ★ How much growth is high, moderate, or low?

★ Quantitative Approach
   ★ Historical data
Checking Parameters

- **Quantitative Approach**
  - Enough variability in scores?
  - Problems with bias?

- **Validating Results**
  - Choose random students, are they correctly categorized?
Example: Music

<table>
<thead>
<tr>
<th></th>
<th>C Major</th>
<th>D Major</th>
<th>C Minor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Expression</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Students are rated on a three point rubric for accuracy and expression for each scale. (3 is high, 1 is low)*

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>No errors</td>
<td>1-3 errors</td>
<td>4 or more errors or failure to finish</td>
</tr>
<tr>
<td><strong>Expression</strong></td>
<td>Consistent tempo and smooth transition between notes.</td>
<td>Some inconsistency in tempo or dynamics</td>
<td>Inconsistency in tempo or dynamics</td>
</tr>
</tbody>
</table>
Music Example Qualitative Approach

- Pre-Test = 9, Post-Test = 11
  - Learned D Major Scale, OR made progress on D Major scale and expression

- Pre-Test = 9, Post-Test = 14
  - Learned D Major and C Minor Scale OR just D Major with major improvement in expression.

<table>
<thead>
<tr>
<th>C Major</th>
<th>D Major</th>
<th>C Minor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Expression</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Students are rated on a three point rubric for accuracy and expression for each scale. (3 is high, 1 is low)
Music
Example
Qualitative Approach

★ Pre-Test = 11, Post-Test = 13
  ★ Learned C Minor Scale, **OR** made progress on accuracy and expression in first two scales.

★ Pre-Test = 11, Post-Test = 16
  ★ Learned C Minor Scale **AND** made progress on accuracy and expression in scales.

<table>
<thead>
<tr>
<th></th>
<th>C Major</th>
<th>D Major</th>
<th>C Minor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Expression</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Students are rated on a three point rubric for accuracy and expression for each scale. (3 is high, 1 is low)
Music Example Qualitative Approach

★ 2 Points of Growth
★ Is learning D major scale similar to learning the C minor scale?

★ 5 Points of Growth
★ Is learning D major and C minor similar to learning C minor and making improvements in expression and accuracy?

<table>
<thead>
<tr>
<th></th>
<th>C Major</th>
<th>D Major</th>
<th>C Minor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Expression</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Students are rated on a three point rubric for accuracy and expression for each scale. (3 is high, 1 is low)
Music Example Qualitative Approach

★ 2 Points of Growth
★★ Is learning D major scale or learning the C minor scale low growth? Is it moderate growth?

★ 5 Points of Growth
★★ Is learning D major and C minor or learning C minor and making improvements in expression and accuracy moderate growth? Is it high growth?

<table>
<thead>
<tr>
<th></th>
<th>C Major</th>
<th>D Major</th>
<th>C Minor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>Expression</strong></td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Students are rated on a three point rubric for accuracy and expression for each scale. (3 is high, 1 is low)
Music Example

Quantitative Approach

★ Assume we agree to cut scores for high, moderate, and low growth.

★ How would these results from across the district inform future parameter setting?

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Moderate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>16</td>
<td>87</td>
<td>44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Moderate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>12</td>
<td>135</td>
</tr>
</tbody>
</table>
Tracey is a student who was rated as having high growth.

Investigate:
- Johnny’s Effort
- Teachers’ perception of Johnny’s musical growth
- Other evidence of Johnny’s musical growth
Communicating Results
Validity & Impact Rating

★ Remember: Validity is around how an assessment is used, and not a property of an assessment

★ A Rating of Low Impact = requires further inquiry.

★ Results from Inquiry
  ★ Improvement to Curriculum
  ★ Improvement to Measures
  ★ Improvement to Teacher’s Practice
Register for Webinar Series Part 9

★ Part 9: Sustainability
★ Date: February 27, 2014
  Time: 4-5pm EST (60 minutes)
★ Register: https://air-event500.webex.com/air-event500/onstage/g.php?d=598148178&t=a
Questions

⭐ Contact

⭐ Craig Waterman at cwaterman@doe.mass.edu
⭐ Ron Noble at r noble@doe.mass.edu

⭐ Tell us how we did: