Developing a Valid and Reliable Formative and Summative Assessment System

Presented by

Dr. Jeff Barker and Kay Elder

To request permission for the redelivery of this presentation, please contact Dr. Jeff Barker or Kay Elder at the Georgia Center for Assessment.
Objectives

- With the passage of Senate Bill 364, districts have the opportunity to develop and administer formative/summative assessments in content areas and grade levels that do not have Milestone assessments.
  - Discuss the research-based methods required to develop valid and reliable formative and summative assessments.
    - Topics included are: learning targets, complexity and rigor, blueprints, and assessment/item specifications.
  - Review the components of the assessment development cycle.
    - The development of formative and summative assessments will assists districts and schools in measuring individual student achievement and identify areas where students demonstrate mastery and face challenges.
  - Discuss how assessments can be utilized to improve instructional practice and learning.
Take 2 minutes and discuss with the person next to you what type of information the Milestones assessments provide and what information is not provided.

For the information that is not provided, how do you collect that information?
Milestones Assessment

- Designed to provide information about how well students are mastering the state-adopted content standards in the core content areas.
- Designed to provide students with critical information about their own achievement and their readiness for their next level of learning – be it the next grade, the next course, or endeavor (college or career).

- Evaluates student learning at the end of the year by comparing it against a benchmark (scale scores and performance levels).
- Measures mastery of learning standards/domains over time. While this is not as diagnostic in nature, it could be used for reflection of instruction and future planning.
- Provides the big picture for your classroom, grade level, school, etc.
**Summative Assessment**

- Assessment **OF** learning
- It is an **EVENT** - not a process
- Provides **performance data**
- Designed to **assess standards/domains** for communicating performance
- Results are utilized to **sum up the learning** at the end of the teaching.
- Results usually **do not provide** you with **in-depth** instructional information.
- Results are utilized to impact teaching and learning for the **next** school year.
Discussion

- Take 2 minutes and discuss with the person next to you what formative assessment looks like in your district or school.
- How is it being used and how do you know it provides evidence of the intended purpose?
Formative Assessment

- Assessment FOR learning
- It is a PROCESS - not an instrument or event
- Can be formal or informal
- Provides ongoing instructional information and helps reflect on instruction
- Identifies “gaps” in what students know and should know
- Designed to assess learning targets and guide decisions about instruction and learning
- Provides active student feedback about mastery and progress
- Provides information for the planning, development and implementation of professional learning
Formative or Summative

Ask yourself three questions:

- What is the purpose?
  - Why are we developing the assessment(s)?
  - What is our ultimate goal?

- Who will use the information?
  - Considering your audience is important in developing your assessment and will impact your design.

- How will it be used?
  - The use of results will impact the test development process, blueprints, reporting design and most importantly instruction.
Formative is when the chef tastes the soup.
Summative is when the guests taste the soup.
Serving bland soup to guests is probably the best indicator that the cook was not proficient at tasting the soup during preparation.
On the other hand, too much reliance on tasting the soup may lead to inattention to the goals – the soup is cold when the guests arrive!
A common understanding of the purpose and expectations of assessment is critical when determining the use of assessment related to student learning, instruction and achievement.
Individually, draw the front of a penny.

Include as many details as you can without looking at one.

Do not compare with a partner until instructed.
ASSESSMENT DEVELOPMENT
Valid and Reliable

How do you assure assessments are both?
Determines whether an assessment measures something consistently

- How to maintain reliability for assessments
  - Standardized administration and scoring procedures to help maintain validity, reliability, and fairness.
  - Teachers give the same form of the assessment.
  - Testing windows are set within the district.
  - P-values (percent of students answering correct) are monitored and reviewed for consistency across the district.
Validity

Determines whether an assessment measures what it is designed to do.

- How to maintain validity for assessments
- Assessment development procedures approved by:
  - American Psychological Association (APA)
  - National Council on Measurement in Education (NCME)
  - American Educational Research Association (AERA)
  - Technical Advisory Committee (TAC) comprised of national experts in the field of testing and psychometrics
- Focus groups
Validity

- Aligned with instructional calendars
- Standardized item and test specifications
- Blueprints developed
- Depth of Knowledge incorporated to assure rigor

- Universal design incorporated to increase accessibility to all students
- Bias and sensitivity reviews
- Field testing
Assessment Development Process

1. Develop or Revise Curriculum
2. Develop Blueprint
3. Develop Item Specifications
4. Develop Field Test Items
5. Review Field Test Items
6. Item Review
7. Develop Forms
8. Review Forms
9. Administer Assessment
10. Data Review
# Development Process

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Review the curriculum, instructional calendars and pacing guides.</td>
</tr>
<tr>
<td>2.</td>
<td>The blueprint is developed and based upon the instructional calendars.</td>
</tr>
<tr>
<td>3.</td>
<td>Item specifications are created to provide consistency among test items.</td>
</tr>
<tr>
<td>4.</td>
<td>Field test items are developed based upon blueprint.</td>
</tr>
<tr>
<td>5.</td>
<td>Field test items are reviewed, approved and embedded into assessments.</td>
</tr>
<tr>
<td>6.</td>
<td>Approved items are reviewed once again. If items are changed or edited, they must be field tested again.</td>
</tr>
<tr>
<td>7.</td>
<td>Assessment forms are developed based upon blueprints and reviewed.</td>
</tr>
<tr>
<td>8.</td>
<td>Schools may review each assessment form and provide feedback.</td>
</tr>
<tr>
<td>9.</td>
<td>Schools administer the assessments.</td>
</tr>
<tr>
<td>10.</td>
<td>A data review is conducted to determine how items performed (p-value) on the assessment.</td>
</tr>
</tbody>
</table>
DEVELOPMENT CONSIDERATIONS
Rigor or Difficulty
Reflection

- Which are your assessments more like?
  - Jeopardy
  - Survivor
  - Celebrity Apprentice

- Do your assessments not only assess knowledge, but skills, application, reasoning and problem solving?

- How do you know?
An assessment is only as good as the process by which it is created.

Is your staff trained in the following areas?

- Assessment Literacy
- Formative Assessment
- Item Writing Training
- Data Analysis
- Milestones
- Formative Instructional Practices (FIP)
- Assessment Development
- Building a Culture of Effective Instruction and Assessment
What if I have assessments developed?
How do I assure the assessments are valid and reliable and exhibit high quality?
An Assessment Audit will help you determine whether your assessment system has a balance of assessments and if they represent quality assessments.

- Does it meet the assessment needs of all learners?
- Is it accessible to all students?
- Is it aligned to the rigor in the curriculum and instruction?
- Are the assessments and the results produced valid and reliable?

Some of the areas to evaluate to help answer these questions include:

- Item Alignment to Standards and/or Performance Targets
- Evaluation of Depth of Knowledge (DOK) Classifications
- Universal Design
- Adherence to Style Guide
- Bias/Sensitivity
- Item/task analysis: p-value, point biserial, KR 20, discrimination index, etc.
Guiding Questions

Does your assessment system:

- Reduce the testing burden?
- Protect instructional time?
- Provide accurate and valuable information to positively impact student achievement?
- Evaluate instructional practices for quality and effectiveness?
- Provide high-quality assessments so that educators and families can better understand student learning needs?
- Provide information about professional development needs and engaging educators in thinking critically about instruction and assessment and the use of assessment data?
HOW DO I KNOW?
WHERE DO I START?
If you have assessments, review the alignment of the purpose, audience and use.

Conduct an assessment audit and implement changes

Incorporate what you learn from the audit into the culture

If you are developing assessments:
  - Determine the purpose
  - Determine how it will be used
  - Determine who the audience will be
  - Discuss formative instructional practice and assessment

Develop a short and long term plan – assessment development is not an easy or quick process

Set achievable goals

Determine how to measure your goals

Celebrate small wins!
How Do I Know?

- Review the validity and reliability components for your development plan
- Conduct professional learning for item writing and assessment development
- Determine if the results are having an impact on instruction, learning and achievement
CLASSROOM ASSESSMENTS
As you begin to develop or improve your assessment systems, carry that knowledge and process into the culture.

Many times we forget to model the same practices in the classroom.

Our observations have been that instruction is most often not aligned to the assessment.

Classroom assessments are key to improving instructional practices and student learning/achievement.
Five Keys to Quality

- **Key 1**: Designed to serve the *specific information needs of intended user(s)*
- **Key 2**: Based on clearly articulated and appropriate *achievement targets*
- **Key 3**: *Accurately measures* student learning and achievement
- **Key 4**: Yield results that are *effectively communicated* to their intended users
- **Key 5**: *Involve students* in self-assessment, goal setting, tracking, reflecting on, and sharing their learning.
Five Keys to Quality

Key 1
- Who will use the information?
- How will they use it?
- What information, in what detail, is required?

Key 2
- Are learning targets clear to teachers?
- What kinds of achievement are to be assessed?
- Are these learning targets the focus of instruction?

Key 3
- Do the assessment methods match the learning targets?
- Are items, tasks, and scoring rubrics of high quality?
- Does the assessment control for bias?
Five Keys to Quality

Key 4
- Can assessment results be used to guide instruction?
- Do formative assessments function as effective feedback?
- Is achievement tracked by the learning target and reported by the standard?
- Do grades communicate achievement accurately?

Key 5
- Do assessment practices meet students’ information needs?
- Are learning targets clear to students?
- Will the assessment yield information that students can use to self-assess and set goals?
- Are students tracking and communicating their evolving learning?
Do all educators:

- Understand that assessment can serve a variety of purposes
- Begin with clear and well-defined learning targets
- Develop assessments – formal and informal – with quality control criteria
- Analyze and provide results in a timely and understandable fashion
- Engage students in productive self-assessments to support learning
Classroom Assessment From...

- Classroom “tests” disconnected from the focus of instruction
- Assessments using only selected response formats
- Assessments only assessing low level thinking (recall)
- “Mystery” assessments, where students don’t know in advance what they are accountable for learning
- All assessments and assignments, including practice, “count” for a grade
- Students are passive participants in the assessment process
- Students not finding out until the “graded” event their strengths and areas of challenge
CLASSROOM ASSESSMENT TO...

START A CONVERSATION IN YOUR SCHOOLS TO BEGIN TO ENVISION THE “TO…”
Classroom assessments are aligned to the focus of instruction.
Assessments using a variety of formats and item types.
Assessments assess all levels of thinking.
Students know in advance what they are accountable for learning — no surprises!
When students are in the process of learning, grades are not the focus. This encourages students to take risks in their learning and not fear failure in the learning process.
Students are active participants in the assessment process
Students have a clear understanding of the learning expectations and what mastery looks like.
Students understand what they mastered and what they need to work on and how to get there.
Collaboration and partnerships among schools, districts and RESAs improve the quality of your assessments and can be highly cost effective.

Remember, there is no quick fix - it is a process.

Seek out experts to assist and guide you.

Keep students as the focus!

“As much as you put into it is as much as you get out of it” – Nuno Bettencourt
The success of a well-designed system of assessments is dependent on the knowledge and skills of the school leaders and teachers who are responsible for bringing the system to life in schools.

-National Center for the Improvement of Educational Assessment
Questions?