And the Rest of the Story…

Part of the Interim Assessment Toolkit we don't have time to discuss at RILS

Center for Assessment

Presentation at the 2019 Reidy Interactive Lecture Series
September 27, 2019   Portsmouth, NH

The big picture

- Phase 1: Identifying Assessment Gaps and Needs

    What do you want to do educationally?
    What do you need in terms of assessment information?

- Phase 2: Identifying and Prioritizing Assessment Characteristics and Evidence of Assessment Quality

    What assessment will fulfill your needs?

The big picture… and main parts

- Phase 1: Identifying Assessment Gaps and Needs
    - Part 1 – Overall Vision
    - Part 2 – Assessment Vision
    - Part 3 – Assessment Lay of the Land
    - Part 4 – High Priority Need

- Phase 2: Identifying and Prioritizing Assessment Characteristics and Evidence of Assessment Quality
    - Part 1 – Clarification of Use
    - Part 2 – Defining Assessment Characteristics
    - Part 3 – Identifying and Evaluating Evidence of Technical Quality

Phase 2, Part 3: Identifying and Evaluating Evidence of Technical Quality

- Toolkit provides support for systematic gathering and evaluation of evidence
- Assessment Quality Claims
  - Test Design
  - Test Administration
  - Reporting – Achievement Status
  - Reporting – Prediction
  - Reporting – Diagnostic
  - Reporting – Growth
  - Reporting – Pedagogical

Part 3: Identifying and Evaluating Evidence of Technical Quality – Test Design

- Assessment Quality Claims
  - Test Design
    1. The intended purpose and uses of the assessment are clearly stated.
    2. Test items and passages align to the intended content standards or academic expectations.
    3. Test blueprints reflect an appropriate distribution of content, item types and cognitive demand within forms given the *target sampling domain*.
    4. Item development and review procedures and materials are in place to ensure all newly developed items are fair and meet technical quality standards.
    5. Test development and review procedures are in place to ensure forms meet the content and statistical quality requirements reflected in test blueprints and specifications.
    6. Test forms are designed to ensure scores can be compared across forms, occasions and students in the manner intended.
    7. There are appropriate accommodations in place to support the fair measurement of all students in the intended test-taking population.

Part 3: Identifying and Evaluating Evidence of Technical Quality – Test Design – Claim #3: Test blueprints

Center for Assessment

Interim Assessment Identification and Evaluation Process		
Claim	**Key Questions to Pose**	**Potential Sources of Evidence**
3. Test blueprints reflect an appropriate distribution of content, item types and cognitive demand within forms given the *target sampling domain.*	Is the target sampling domain (e.g., grade 3 standards; Number Sense) associated with the test blueprint clear?	Documentation summarizing the required characteristics of each operational test form with respect to content representation, cognitive demand (e.g., Depth of Knowledge), and item formats.
	Is the distribution of test content represented by the blueprint appropriate for making inferences about student performance in the target sampling domain?	
	Does each test form require an appropriate range/distribution of cognitive complexity?	
	Are the item types represented appropriate to make inferences about student performance in the target content domain?	
	Is the target domain, as defined by the test blueprint, appropriate to support my assessment information needs?	

The big picture… and main parts

- Phase 1: Identifying Assessment Gaps and Needs
  - Part 1 – Overall Vision
  - Part 2 – Assessment Vision
  - Part 3 – Assessment Lay of the Land
  - Part 4 – High Priority Need

- Phase 2: Identifying and Prioritizing Assessment Characteristics and Evidence of Assessment Quality
  - Part 1 – Clarification of Use
  - Part 2 – Defining Assessment Characteristics
  - Part 3 – Identifying and Evaluating Evidence of Technical Quality

- **Phase 3: A Use and Utility Evaluation Tool?**