



NATIONAL ALTERNATE ASSESSMENT CENTER

Deficient, Proficient, Magnificent: Setting Achievement Standards for Alternate Assessments

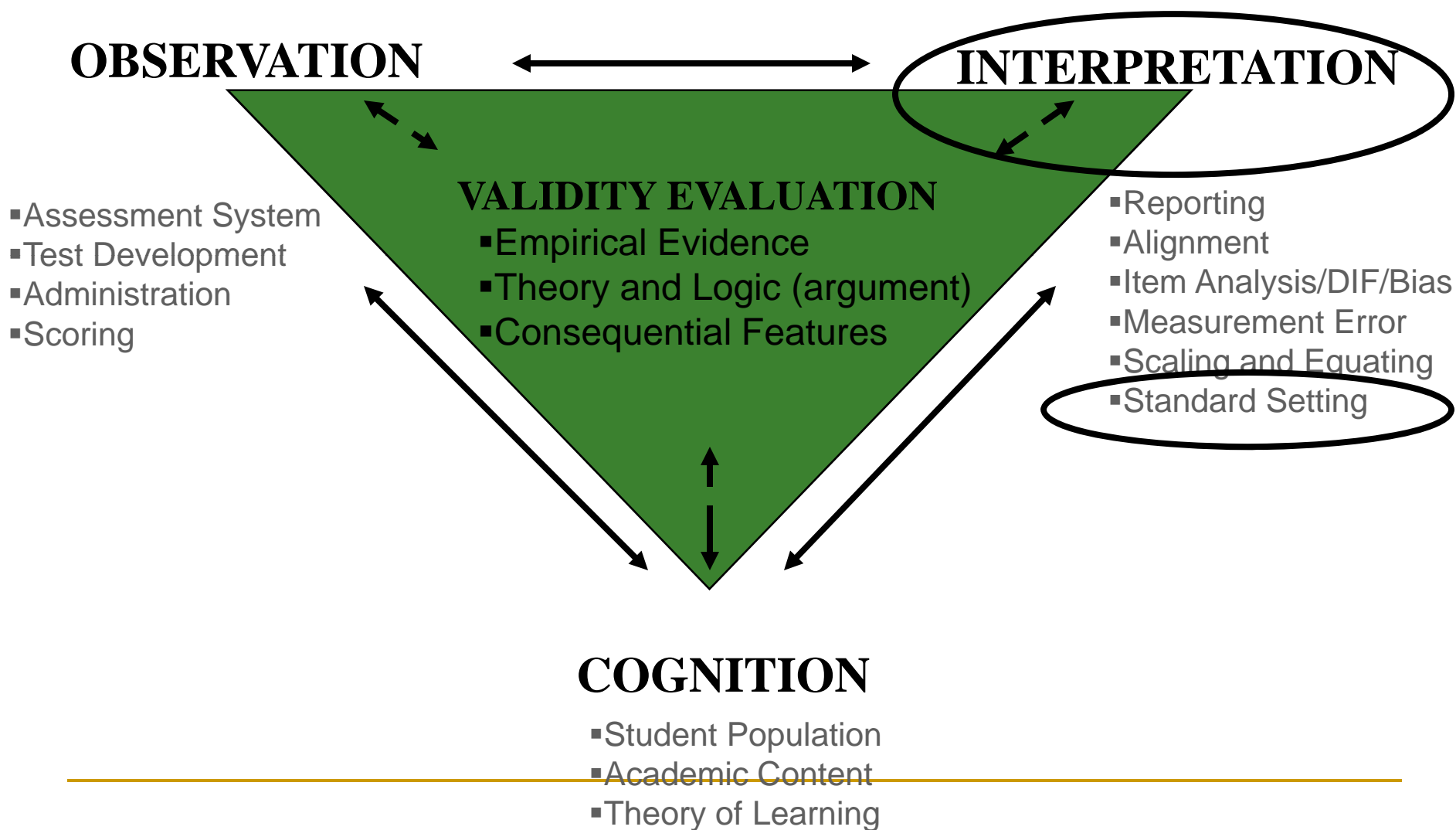
July 25, 2007

Jacqui Kearns, Director
National Alternate Assessment Center
University of Kentucky

Susan Kennedy, Acting Chief
Bureau of School & District
Improvement
Connecticut Department of
Education

The Assessment Triangle and Validity Evaluation

Marion & Pellegrino (2006)



What is standard setting?

- “the task of deriving levels of performance on educational or professional assessments by which decisions or classifications of persons (and corresponding inferences) will be made”
 - Cizek, (1993a)
-

Why do we have to do it?

- Regulation §200.1
 - “For students... with the most significant cognitive disabilities, who take an alternate assessment, a State may, through a documented and validated standards-setting process, define alternate achievement standards...”
 - Peer Review Guidance Critical Element 2.1
 - *[States may] define alternate academic achievement standards, provided those standards (1) are aligned with the State’s academic content standards; (2) promote access to the general curriculum; and (3) reflect professional judgment of the highest achievement standards possible.*
-

Setting Alternate Achievement Standards - Overview

1. **Write clear performance level descriptions aligned with the state content standards:** The task of writing PLDs is similar to that for large-scale assessments. Should be written by people who understand the population and what can be achieved.
2. **Match the judgmental task of the standard-setting method to the test type:** Methods can focus on items, performance tasks, whole bodies of work, or students. Choose the methodology, or an adaptation, that best fits the assessment.
3. **Validate that the cut score best reflects the intention of the performance level descriptor:** It is important to conduct follow-up studies to ensure the “right” students are being placed in each performance level.
4. **Document the rationale and procedures used in setting the achievement standards.** Documentation is an important and often overlooked step in the process.

How to Develop Good PLDs

- Convene a panel of experts
 - Special education teachers
 - Others familiar with the content, curriculum, special education students, or assessment
 - Discuss the students
 - Review relevant research on special education learning and achievement
 - Review sample student work
 - Discuss the desired end product for the descriptor
-

Selecting content for the AA & Writing the PLD's

- **Alternate Assessment Advisory Committee**
 - w/ CSDE curriculum consultants and assessment consultants
 - Reviewed the **CT Curriculum Frameworks**
 - **Selected** Content Standards and Expected Performance items
 - to serve as the **basis for the CMT/CAPT Skills Checklist**
-

Selecting content

- **Three all day meetings** were held to establish the content and the placement of the Content Standards and Expected Performance statements
 - Used a “**dot vote**” procedure
 - identify the **Expected Performance within each Content Standard**
 - **considered the most important** for students with significant cognitive disabilities
-

Selecting Content

- To focus the selection and discussion, three questions were considered:

1. What endures? What skills and knowledge will students gain that last from one year to the next?

2. What is essential for progress to the next level of instruction? What skills are essential for future success?

3. What contributes to the understanding of other standards? What are the skills, that once mastered, give a student the ability to use reasoning and thinking skills to learn and understand other curriculum objectives?

“Dot Vote” Procedure

- Participants utilize a **specific number of votes (5)** to identify content they considered most important
 - Participants place **dots next to content** arrayed
 - Content is **rank ordered** according to number of times selected
 - If selected by **more than half** of the participants then discussed by the whole group
 - **Discussion leads to consensus** regarding the importance of the items
-

Finalizing Content

- Do the Content Standards and Expected Performance selected address the full range of skills included in the Content Strand? Is it placed correctly across grades?
 - Are the content Standards and Expected Performances selected sufficient for the assessment of grade level skills at each grade 3 to 8 and 10?
-

Finalizing Content

Procedure used to validate the **selection** and **placement** of content

- Independent expert review completed by the mathematics and Language arts curriculum consultants at the CSDE

ALIGNED BY DESIGN

Write Performance Level Descriptors

- What should the “proficient” student know and be able to do?

Example

Proficient Reading 3rd Grade

Provided with supports that individualize age/grade appropriate materials for the student’s of language development; The student demonstrates an independent and mostly accurate understanding of the ***specified*** reading skills/concepts in natural contexts. Occasional inaccuracies, which do not interfere with conceptual understanding, may be present.

These skills/concepts include:

- answering questions
- providing relevant details using a reading vocabulary,
- locating information,
- explaining a concept,
- solving problems.

PLDs to cut scores

- PLDs can be written at any time in the development process
 - We often argue that PLDs should be written before designing the assessment
 - Cut scores are typically set after the assessment has been administered for the first time
-

Match the judgmental task of the standard setting method to the test type

■ Test Type

Standard Setting Method

■ Portfolio

Body of Work, Dominant Profile, Policy capturing

■ Performance Tasks

Modified Angoff, Yes/No, Bookmark. Performance Profile

■ Rating Scales

Modified Angoff, Bookmark, Yes/no

- See Cizek (2001)

Standard Setting Workshop

- Choose standard setting method
 - Match method to assessment characteristics
- Convene panel of stakeholders
 - Need to be experts in special education, general curriculum, and assessment
- Review content standards and assessment
- Discuss PLDs and borderline student
- Train on and practice methodology
- Run 2-3 rounds with individual judgment, feedback, and group discussion
- Provide impact data
 - Consider breaking out data by disability type
- Summarize results

4th Grade Reading Standard Cut Score

Example

Grade Reading

Scoring Dimensions	A	B	C	Total
<u>Dominant Profile Consensus</u>				
Beginning	10	10	10	32
Proficient	14	14	14	45
Magnificent	19	18	18	55
<u>Consensus after Body of Work</u>				
Beginning	9	10	11	31
Proficient	14	14	14	43
Magnificent	19	18	18	55
<u>Final Consensus (after impact data feedback)</u>				
Beginning	9	10	11	31
Proficient	14	14	14	43
Magnificent	19	18	18	55

Impact Data Example

Magnificent	1%
Proficient	11.5%
Improving	44.3%
Beginning	42.5%

After the Workshop

- Smooth the data
 - After the workshop ends, state DOE might want to adjust the cut scores to align the impact data across grades or subjects
 - Some argue that smoothed cut scores should stay within error bands (consider SEJs and SEMs)
 - Adopt the cut scores
 - Standard setting panel makes recommendation but policy board adopts cut scores
 - Policy board should be given information about what happened during the meeting, any adjustments made afterwards, measures of variance, and impact data
-

Validation

- Cut scores do not have validity but the interpretation and uses of the cut scores do
 - Just as there are no absolute criteria against which specific cut scores can be evaluated, there are no perfect criteria for evaluating standard setting studies (Kane, 1994, 2001)
 - But... even though there are no absolute criteria you still must provide evidence that the cut scores are reasonable and appropriate
-

Consequential Validity

- What effect is the use of the performance levels having on student opportunity to learn?
 - Teacher professional development
 - Pedagogical changes
 - Student exposure to richer curriculum
- Are there any unintended consequences from using the performance levels?
 - Certain content standards being de-emphasized
 - Certain students being neglected
 - Teacher retention decreasing

Validation Over Time

- Follow students over time
 - Is improvement captured by the performance levels?
 - Do the classifications over time make sense? (e.g., results that show a student to be proficient in grade 3, basic in grade 4, below basic in grade 5 and proficient in grade 6 would not be considered reasonable)
- Survey teachers/other stakeholders over time
 - Are performance levels capturing true achievement?
 - Do they see possibility for improvement for the students?
 - What suggestions do they have for modifying PLDs and/or cut scores

Documentation

- From our joint *Standards* (AERA, APA, & NCME)
 - Document the PLDs, selection of panelists, training provided, ratings, and variance measures (Standard 1.7)
 - Document the rationale and procedures for the methodology used (Standard 4.19)
-

Conclusion

- Setting alternate achievement standards should follow the logic of the entire alternate assessment program
 - Each step should be made in consideration of the other steps and the goals of the program
 - Write clear & appropriate PLDs
 - Choose an appropriate methodology and implement it using best professional practices
 - Validate the cut scores and classifications over time
 - Document every step of the process
-

References and Resources

- Cizek, G. (2001). Setting Performance Standards: Concepts, Methods, & Perspectives. New Jersey: Lawrence Erlbaum.
 - Perie, M. (2006). Standard Setting: Procedures, Validation, and Documentation. Alternate Assessment Seminars.
www.naacpartners.org or www.nciea.org
-