Standard Setting

Inclusive Assessment Seminar

Marianne Perie

NHEAI

New Hampshire Enhanced Assessment Initiative:
Technical Documentation for Alternate Assessments

National Alternate Assessment Center

The Assessment Triangle and Validity Evaluation

(Marion, Quenemoen, & Kearns, 2006)

OBSERVATION ⟷ **INTERPRETATION**
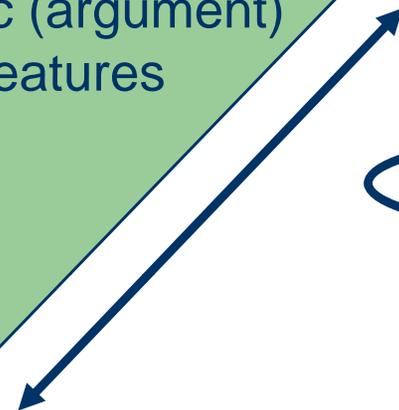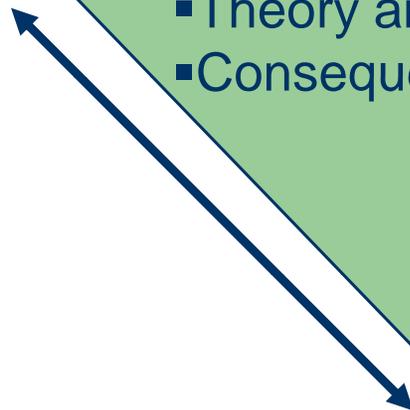
VALIDITY EVALUATION
- Empirical Evidence
- Theory and Logic (argument)
- Consequential Features

- Assessment System
- Test Development
- Administration
- Scoring

- Reporting
- Alignment
- Item Analysis/DIF/Bias
- Measurement Error
- Scaling and Equating
- Standard Setting

COGNITION
- Student Population
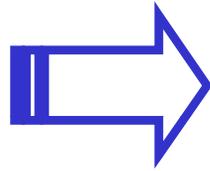- Academic Content
- Theory of Learning

Regulation on alternate achievement standards §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…"
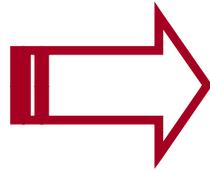
ACHIEVEMENT STANDARDS

4 Components

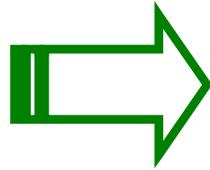Performance levels ➡ **Labels** each level of achievement
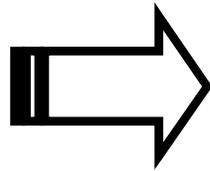
Performance descriptors ➡ **Describes** each level of performance

Exemplars ➡ **Sample student work** at each level of performance

Cut Scores ➡ Scores that **separate** the different levels of performance

Alternate vs. General Assessment

- Much is the same:
  - Write distinct, detailed performance level descriptors (PLDs) that link directly to the content standards
  - Use a documented and validated standard setting methodology
  - Convene a panel of educators and stakeholders who are familiar with the population of students tested and represent the diversity of the state
  - Train them in the content standards, PLDs, and methodology
  - Collect individual judgments and use group discussion across rounds
  - Provide normative feedback to the committee to show the relationship between the cut scores and results
  - Aggregate judgments to determine final panel-recommended cut scores
  - Convene appropriate policymaker(s), apprise them of the recommendations, and have them legally adopt the cut scores
  - Document the PLDs, rationale and procedures for the methodology, selection of panelists, training provided, ratings, and variance measures

5

Alternate vs. General Assessment

- Only a few details differ:
  - PLDs can be written to a grade span rather than a grade level if adjoining grades are sufficiently similar.
  - PLDs range from entry points to exit points but must represent the highest standard possible for this population.
  - Panels are more likely to include parents and other stakeholders involved in this special population.
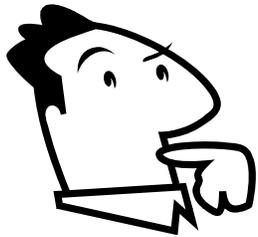
Key Considerations

1. **Writing clear, descriptive performance level descriptions aligned with the state content standards**: The task of writing PLDs is similar to that for large-scale assessments. They must be written by people who understand the population and what can be achieved.

2. **Matching 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. **Validating 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.

PLDs

- Requirements
  - Represent highest standard possible for this population
  - Capture essential skills
  - Align with state content standards
  - Clearly differentiate among levels
  - Progress logically across levels (e.g., is Proficient appropriately higher than Basic)
  - Progress logically across grade levels (e.g., is grade 5 proficient sufficiently more advanced than grade 3 proficient)
  - Represent knowledge and skills that can be evaluated by the assessment (e.g., don't discuss independence in the PLD if your assessment doesn't measure independence)
- Options
  - Write PLDs for a grade span rather than for a grade level

Basic Steps for Writing PLDs

1. Convene a panel of stakeholders for each subject and grade span
2. Provide information about the students and assessment
3. Share relevant literature about what students with disabilities can learn and do
4. Share sample student work
5. Discuss requirements for PLDs
6. Provide sample PLDs
7. Ask panel to draft PLDs as a group; work for consensus
8. Compare PLDs across panels (subject/grade span) and revise as needed
9. Final revisions and adoption by policymakers

9

Choosing a Method

- Match the judgmental task of the method to the types of data gathered through your assessment
  - (E.g., Angoff or Bookmark work better with assessments with multiple items while Body of Work or a profile approach work better with a portfolio or holistic assessment)
- See handout summarizing methods and corresponding assessment types

Adopting a Cut Score

- Consider smoothing your results across grades or across subjects so that your impact data make sense.

- Cut scores should be adopted by a policy body (e.g., state board) who is provided with relevant information.
  - Panel-recommended cut scores
  - Smoothed cut scores
  - Variance measures (SEJ/SEM)
  - Impact data (could be broken out by disability)
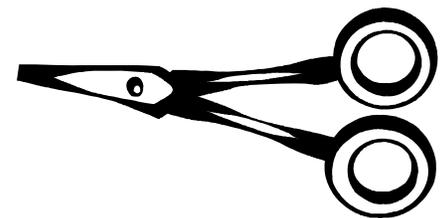
Validating the Performance Standards

- Cut scores may be considered provisional until they have been used and validated.

- Validation strategies:
  - Teacher judgments on students (contrasting groups)
  - Longitudinal data on students' progression across grade-levels (Stay in performance level? Move up/down? How much movement?)
  - External measures (other evaluations, grades)

- Monitor effects of using performance levels over time

Checklist for Performance Standards

- Understandable and useful for stakeholders
- Clearly differentiate among levels
- Grounded in student work but not tied to status quo
- Built by consensus
- Focused on learning

CCSSO Handbook, p. 16