Knowing What Students With Severe Cognitive Disabilities Know: Constructing the Assessment Triangle for Alternate Assessments

Scott Marion, Center for Assessment
Jim Pellegrino, Univ. of IL, Chicago
Martha Thurlow, NCEO
Gaye Fedorchak, NH DOE

Two Collaborating Projects

- The New Hampshire Enhanced Assessment Initiative
  - State Partners--NH, MA, CO, NM
  - Organizational Partners—NCEO, Measured Progress, Center for Assessment
  - Funding agent—Elementary and Secondary Ed—Title I
- National Collaborative Center on Standards and Assessment Development
  - State Partners--CO, IA, KS, KY, MD, MA, MI, NH, NM, NC, and SC
  - Organizational Partners—UK, NCEO, UNC-Charlotte, CAST
  - Funding Agent--OSEP

Major Project Goals

- To create a prototype for an alternate assessment technical manual.
- To gain a deeper understanding of the "assessment triangle" for students participating in alternate assessments:
  - To build models of domain proficiency for students with the most severe cognitive disabilities.
  - To offer insights into the most appropriate means of these assessing students.
  - To learn how to make the most defensible inferences regarding what these students know and are able to do.

The Alternate Assessment Technical Manual: Using a Validity Framework

Scott Marion
Center for Assessment
CCSSO Large Scale Assessment Conference
June 19, 2005

Sounds Straightforward?

- After all, most states have technical manuals to document the quality of their general assessment systems.
- Unfortunately, the quality of existing technical documentation is lacking. This is not necessarily the fault of contractors or state personnel. Often a budget issue.

Validity Should be Central

- We argue that the purpose of the technical manual is to provide data to support or refute the validity of the inferences from the alternate assessments at both the student and program level.
- But, it is not so easy…

Validity of Performance Assessments

- Linn, Baker, & Dunbar (1991) argued that evaluating complex performance-based assessments using conventional correlation-based procedures tended to under-represent the actual validity and quality of these assessments. They outlined a set of criteria, based on Messick's (1989) conceptualization of validity that could be used to more appropriately evaluate the quality of these newer (at that time) forms of assessment

The Challenge of Alternate Assessments

- Documenting the technical qualities of performance is not easy because the measurement field has been slow to move away from traditional correlational indicators
- This challenge is compounded for alternate assessments because of the:
  - heterogeneity of the group of students being assessed
  - relatively small numbers of students/tests
  - often (depending on state) idiosyncratic assessment experiences

Expanding Technical Quality

- Following Linn, et al. (1991), we support the need to expand our conception of technical quality to better evaluate alternate assessment programs.
- We will draw on the work of Cronbach, Messick, and Shepard and build our evaluation of technical quality around a unified conception of validity.
  - For example, if the alternate assessment program leads to positive instructional improvements for students with disabilities, it can be argued that these consequences support the validity of the program.

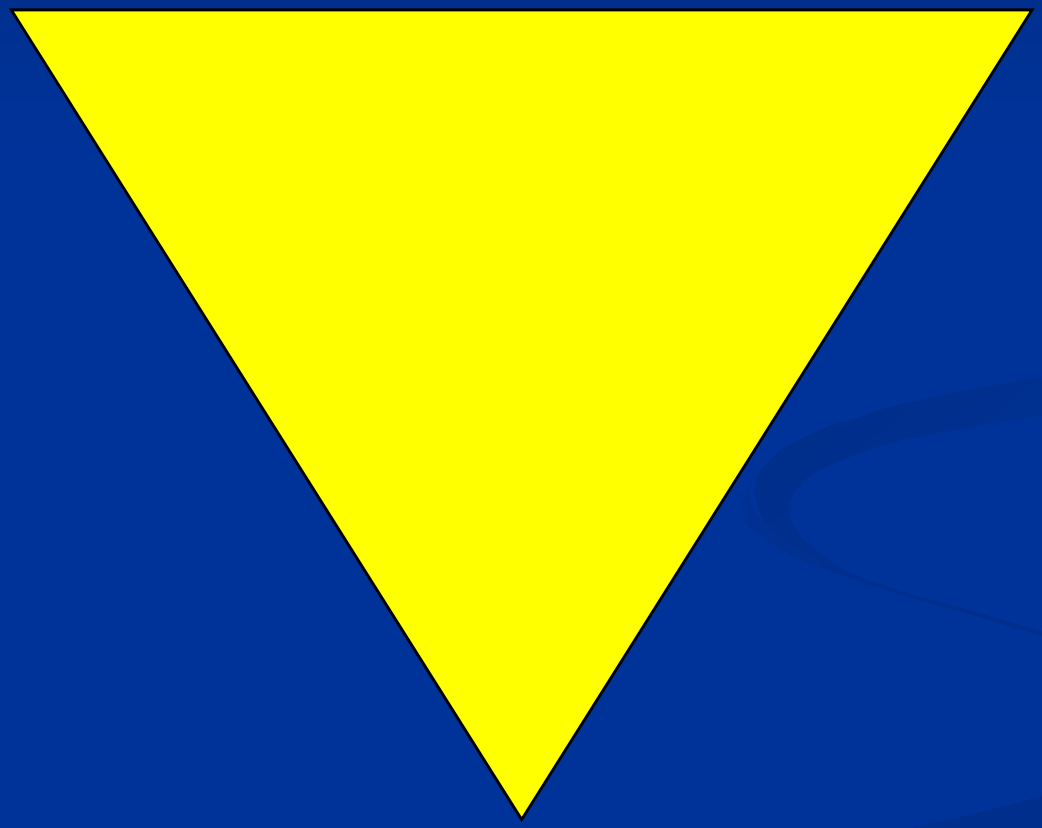
KWSK

- Fortunately, we are undertaking this work after the publication of *Knowing What Students Know: The science and design of educational assessment* (NRC, 2001), which synthesized a tremendous body of learning and measurement research and set an ambitious direction for the development of more valid assessments. *Knowing What Students Know* (KWSK) builds off of Mislevy's (1996) notion of assessment as a "process of reasoning from evidence" (p. 39).

The Assessment Triangle

Observation

Interpretation



Cognition

A Heuristic

- We intend to use the assessment triangle as a heuristic to organize the validity evaluation.
- The triangle immediately reveals an important piece of missing information:
  - We are lacking models of cognition that can be applied generally to students with the most severe cognitive disabilities.

A Starting Point

- Therefore, the first goal of this project is to build a working model of increasing domain competence for students with the most severe cognitive disabilities.
- Once models of cognition are created, we will have a basis to begin evaluating the degree to which the measurement and scoring/interpretation models represent the model(s) of learning for these students.

Models of Cognition

- If I haven't convinced you yet about the importance of the articulation among cognition, assessments, and interpretative models, don't worry—I brought the reinforcements. Jim Pellegrino will walk you through the relationship of the components of the assessment triangle and validity.
- Prior to arriving at models of domain proficiency, we have to first have a thorough understanding of both the students and the content. Martha Thurlow will provide insights into both topics.

Draft Technical Manual TOC

- Section I—Overview, Background, and Key Components of the Validity Evaluation
- Section II—Test Development, Administration, Scoring, and Reporting
- Section III—Technical Criteria
- Section IV—Consequential aspects of the assessment system
- Section V—The Validity Evaluation

Section I—Overview, Background, and Key Components of the Validity Evaluation

- **Overview of the Assessment System**
- **What is the content?**
- **Who are the students?**
- **Introduction of the Validity Framework and Argument**

Section II—Test Development, Administration, Scoring, and Reporting

- **Test Development**
- **Administration & Training**
- **Scoring**
- **Reporting**

Section III—Technical Criteria

- Alignment
- Item Analysis and DIF/bias
- Generalizability
- Decision consistency and accuracy
- Scaling and Equating
- Standard Setting

Section IV—Consequential aspects of the assessment system

- **Effects on students learning opportunities**
- **Effects on teacher professional growth**
- **Programmatic effects on schools and districts**

Section V—The Validity Evaluation

- **Revisiting the validity evaluation questions**
- **Synthesizing and weighing the various sources of evidence**
- **An overall judgment of the validity of the AA-AAS system**

Responsibility

- The responsibility for collecting and analyzing these data does not rest solely with the contractor.
- It should be a joint effort between the state, the contractor, and others (e.g., university partners).
- There is no expectation that the full manual be produced each year, but it is crucial that there be a plan for systematic data collection.

Theory to Practice

- While it is always fun to talk in this rarified air of theory and rich ideas, it does not do the field and the students much good.
- This project is based on several rounds of pilot testing in 4 partner states (NH, MA, NM, & CO).
- Gaye Fedorchak, Director of Special Education for the NH DOE will provide her pragmatic views of the challenges and opportunities of this project.

For more information

- smarion@nciea.org
- pellegjw@uic.edu
- thurl001@umn.edu
- GFedorchak@ed.state.nh.us

- Presentations will be posted at:
  - www.nciea.org
  - www.nceo.org