

Key Considerations for Validating Alternate Assessment Systems

Scott Marion, Center for Assessment
Reidy Interactive Lecture Series
Portsmouth, NH
September 25-26, 2008



Knowing What Students Know and Knowing the Students...

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



Challenges of Alternate Assessments

- Small, very heterogeneous group of students
 - Creates difficulties for statistical analyses
- Flexibility in:
 - Assessment targets
 - Assessment events
 - Administration
- All create psychometric challenges, especially for comparability



Is it Psychometrics or Social Justice?

- Initially, we focused on consequences—yes they are part of validity!—because the intended consequences were a major rationale for including all students in standards-based education.
- However, we realized even before writing the NH EAG proposal that we needed to more completely evaluate the technical quality of alternate assessments.



We've been here before...

- *Remember performance-based assessments...*
- *Content frameworks are described, and specifications for the selection of items are provided for standardized achievement tests. Correlations with other tests and sometimes with teacher assessments of achievement may also be presented. Such information is relevant to judgments of validity but does not do justice to the concept (Linn, Baker, & Dunbar, 1991, p. 16).*

Many Possibilities

- Elsewhere we (Marion & Perie, 2008) have presented examples of evaluation questions and potential studies within familiar (e.g., joint standards) and less familiar (e.g., *Knowing What Students Know*, Ryan, 2002) frameworks for structuring *legitimate* validity evaluations.
- Further, the work of the NHEAG and NAAC have demonstrated that many familiar forms of analyses are possible even if they require some different thinking.



Think Differently...but not really

- Validating alternate assessments requires us to approach the challenge a bit differently than with general assessments, but the “different” is often more about how we conduct certain types of analyses, rather than the validation effort itself.
- Be clear about what we’re trying to do and be critical about what we can conclude
- Don’t bang round pegs into square holes...find a different tool

