

# Implications of Policy Demands on Validity Evaluations: Discussant Comments

Scott Marion

Center for Assessment

June 15, 2008

CCSSO: Nat'l Conference on Student Assessment (Session 49)

Orlando, FL



# The Role of Policy Statements in Validity Arguments

- ▶ Separate from or integral to validity arguments?
- ▶ A starting point or used throughout the validation effort?
- ▶ Appropriate as external criterion?
- ▶ A few comments about each of the presentations first...



# Hoffman—Validity in the *Context*...

- ▶ Gene correctly notes that a test, program, etc must be validated for each of its purposes (but lets not forget *uses* too)
- ▶ Gene argues that accepting or rejecting purposes must precede validation, but how should we treat “unauthorized” uses in validation (e.g., using state school accountability tests for student promotion decisions)?
  - I suspect the field is split here, but this is all the more reason why test developers and sponsors need to be clear about the purposes for which their assessments/programs ARE valid—this requires evidence!



# Hoffman--*Utility*

- ▶ *Gene: Utility is not a simple function of test validity. That is, assessments could be flawed, but if their implementation were stimulating positive changes in educational practices, then in spite of the testing flaws, the program could have high utility.*
- ▶ Scott: However, I am struggling to distinguish this notion of utility from the more common conception of consequential evidence
- ▶ *Gene: However, if the assessments are substantially flawed, we could expect the system to eventually fail*
- ▶ Scott: Which is why consequential and other forms of evidence must be considered as part of a comprehensive validity evaluation



# Hoffman—*Local value & communication*

- ▶ I think Gene's notion of local value is a useful distinction, but I still argue that utility and local value can (and should) be subsumed under construct validation
- ▶ Since validity is about the defensibility of inferences, "publicly digestible communications" should also be part of the larger validity argument



# Martineau

- ▶ “Policy-based validation” is really a clever way to deal with one of the major challenges in any validation effort, that of the *criterion problem*.
- ▶ As Bob Linn and others like to say, “there is no gold standard” against which we can judge our results
- ▶ This doesn’t provide a gold standard, but can provide a sound approach for addressing the criterion problem



# Martineau

- ▶ Joe's two examples provided an important contrast
- ▶ In the ELP case, Joe and his colleagues were able to operationalize the policy phase "unrestricted access" into quite tangible probability statements
  - I could quibble with how this was operationalized, but this was certainly defensible
- ▶ The high school test example required a more tenuous operationalization. This is not a fault of Joe's team, but due more to the fuzziness of the concepts and that lack of explicit intent
  - This examples makes clear the need for us to help policy makers make these statements as clear and unambiguous as possible
  - Anybody who's tried to discern legislative intent know how much fun that can be



# Martineau

- ▶ In this standards-based assessment system, much of the meaning we derive from student scores (i.e., inference) comes from the performance categorization
- ▶ Joe's study points out the importance of carefully attending to the validity of our standard setting results when evaluating our overall assessment systems
  - Standard setting, legislation, and sausage



# Summary

- ▶ Separate from or integral to validity arguments?
  - I'm am pretty clear that these policy statement are not separate from validation and are crucial for determining the intended outcomes
- ▶ A starting point or used throughout the validation effort?
  - I might have said starting point earlier, but these presentations provided a convincing rationale for their use throughout the validity program
- ▶ Appropriate as external criterion?
  - Not by themselves, but need to be operationalize as Joe noted
- ▶ Unintended negative consequences
  - Policy statements rarely alert us to negative consequences, but it is the validity evaluator's job to search for and bring these negative consequences to light

