

Validating Accountability Systems

Brian Gong

Center for Assessment

Presentation at the

2008 Reidy Interactive Lecture Series (RILS)

September 25, 2008 Portsmouth, NH

Validity, Validation, Validating

- Changing views of validity. Validity is:
- Validity is a *state of the thing*: “This test is valid because... [it measures what it is intended to measure].”
- Validity is a *state of the interpretation of the thing*: “This test supports warranted interpretations.”



Validity, Validation, Validating (2)

- Validity is both interpretation and use:
“Evaluative judgment of the adequacy and appropriateness of inferences *and actions*”
(Messick, 1989)
- Validation arguments/evidence: Interpretive argument: Build a chain of reasoning from the test construction process to the desired claims.
Validity argument: Gather theoretical and empirical support for the “truthfulness” of the claims and to establish appropriate boundaries. (Kane, 2004)



Validity, Validation, Validating (3)

- Process of validating (unending):

Ultimately, an evaluative judgment is a *tentative* causal conclusion based on *partial* evidence drawn from an *uncontrolled* study of schools and districts. (Braun, 2008)

- Unending because of constraints
- Unending because of nature of scientific enterprise



So... Process of Validating is

- Interpretive argument: Build a chain of reasoning from the test construction process to the desired claims.
- Validity argument: Gather theoretical and empirical support for the “truthfulness” of the claims and to establish appropriate boundaries [limitations: “but not if...”]. (Kane, 2004)



Claims and “Chain of reasoning”

- Assessment different than Accountability
 - So... validating the assessment does not validate the accountability system (incomplete & different criteria)
- Accountability
 - Purpose
 - Design (requirements)
 - Theory of Action (interpretations, uses, and consequences)



Systemic Validity

Assessment practices and systems of accountability are systemically valid if they generate useful information and constructive responses that support one or more policy goals (Access, Quality, Equity, Efficiency) within an education system, without causing undue deterioration with respect to other goals.

(adapted from Braun and Kanjee, 2006)



Interpretive Argument

- Components of accountability system
 - Purpose
 - Design
 - Test
 - Design, Development, Materials preparation, Administration, Data processing, Scoring, Scaling, Equating, Reporting
 - Performance Standards
 - Indicators



Interpretive Argument (2)

- Components of accountability system
 - Purpose
 - Design
 - Test
 - Performance Standards
 - Indicators
 - Decision rules
 - Judgments/categories
 - Consequences (First tier, i.e., mandated as design)



Interpretive Argument (3)

- Components of accountability system
 - Purpose
 - Design
 - Implementation
 - Theory of action “details”
 - Before and between test and accountability judgments
 - After accountability judgments & consequences and before next assessment and judgment cycle



Validity argument

- Validity argument: Gather theoretical and empirical support for the “truthfulness” of the claims and to establish appropriate boundaries [limitations: “but not if...”]
- Applied to components:
 - Purpose
 - Design
 - Implementation
 - Theory of action “details”



Validity argument (2)

- Does the system:
 - “generate useful information and constructive responses”
 - “without causing undue deterioration with respect to other goals”
- How to construct this type of validity argument?
- Questions about validating: How detailed? Who’s responsible? How much? What about elements of “theory of action” that is under other people’s control (causal assumptions)?



For more information:

Center for Assessment

www.nciea.org



Brian Gong

bgong@nciea.org

