

# Validity evaluations of states' AA-AAS: Arguments and initial findings

---

Scott Marion, Center for Assessment  
CCSSO National Conference on Student  
Assessment (Session 116)  
June 22, 2009



# Kane's argument-based framework

---

- "...assumes that the proposed interpretations and uses will be explicitly stated as an argument, or network of inferences and supporting assumptions, leading from observations to the conclusions and decisions. Validation involves an appraisal of the coherence of this argument and of the plausibility of its inferences and assumptions (Kane, 2006, p. 17)."



# Two Types of Arguments

---

- *An interpretative argument specifies the proposed interpretations and uses of test results by laying out the network of inferences and assumptions leading from the observed performances to the conclusions and decisions based on the performances*
- *The validity argument provides an evaluation of the interpretative argument (Kane, 2006)*



# The Interpretative Argument as Theory

---

- ❑ Similar to the role that theory plays in science, the interpretative argument provides a framework for interpretation and use of test scores
- ❑ Like theory, the interpretative argument guides the data collection and methods and most importantly, theories are falsifiable as we critically evaluate the evidence and arguments
- ❑ The states' arguments in this session have certainly help organize and prioritize studies



# From Shepard (1993, p. 429)...

- What does the testing practice claim to do;
- What are the arguments for and against the intended aims of the test; and
- What does the test do in the system other than what it claims, for good or bad?
- Test validation is basically the process of offering assertions about a test or a testing program and then collecting data and posing logical arguments to refute those assertions
  - If the assertions cannot be refuted, we can say that they are tentatively supported (and that's the best we can do!)



# Criteria for Evaluating Interpretative Arguments

---

- ❑ Clarity—should be clearly stated as a framework for validation. Inferences and warrants specified in enough detail to make proposed claims explicit.
- ❑ Coherence—assuming the individual inferences are plausible, the network of inferences leading from the observations to conclusions and decisions make sense
- ❑ Plausibility—particularly of assumptions, are judged in terms of all the evidence for and against them



# Values and Consequences

---

- ❑ Evaluating a decision procedure requires an evaluation of values and consequences
- ❑ “To evaluate a testing program as an instrument of policy [e.g., AA-AAS under NCLB], it is necessary to evaluate its consequences” (Kane, 2006, p.53)
- ❑ Therefore, values inherent in the testing program must be made explicit and the consequences of the decisions as a result of test scores must be evaluated.



# Validity Evaluation

---

- ❑ As part of this project, we have asked the states to create a validity argument that is bigger than the specific expectations of this project
- ❑ Validity evaluation plan reflecting prioritization
- ❑ Undertake at least three studies
- ❑ Begin synthesizing study results to evaluate the validity argument





# The Arguments and Investigations

---

- ❑ As you listen to these three state leaders/evaluators, listen for the coherence among the goals, arguments, and research studies
- ❑ The argument helps the sum become more than the collection of studies



# Evaluating the Validity Argument

---

- ❑ The methods and results of each study can be judged on its own merits
- ❑ However, synthesizing the results across studies to begin evaluating the interpretative argument can be much more informative than considering each study separately
- ❑ The state participants are just finishing up their first studies, but as you listen, think about how you might go about synthesizing and integrating the various studies into a coherent evaluation



- 
- We will now hear from Ellen Forte (the lead evaluator for both DC and Puerto Rico), Melissa Fincher and Janet Stuck from Georgia and Connecticut, respectively, who are leading their state's efforts to develop and evaluate their validity arguments

