Validity

Definitions
Validity: “the degree to which evidence and theory support the interpretation of test scores entailed by proposed uses of tests”¹ (typically an argument with qualitative and quantitative evidence)

“It is not the test that is validated and it is not the test scores that are validated. It is the claim and decisions based on the test results that are validated.”²

Reliability: the consistency of measurements when the testing procedure is repeated on a population of individuals or groups¹ (typically a mathematical calculation)

Validity evidence helps answer questions like these . . .
• Is it legitimate to use the scores on the national Medical Licensing Examination to determine if a person can legally practice as a physician in Hawai‘i?
• Is it valid to use elementary students’ scores on the Hawai‘i State Assessment to determine if an elementary teacher is effective?
• Is it cogent to use SAT scores as part of the college admissions process?
• Is it valid to use GRE Subject Test scores to determine whether a student should graduate?

Validation involves . . .
• Evaluating the proposed interpretations and uses of test scores
• An argument that explicitly states inferences and assumptions inherent in the proposed interpretation and uses
• An evaluation of the coherence of the argument and the plausibility of its inferences and assumptions

Two things to consider
1) The content of the test/task should be relevant and representative.
2) The (intended and unintended) consequences associated with the planned use of scores are positive. This includes social consequences.

continued on back

Your Turn

Given your fall 2013 assessment project, think about how you might demonstrate that the planned interpretation and use of test/task results are valid. Use these guiding questions to help you think about validity.

1. What is one SLO that you will assess for your fall project?

2. What will students do to demonstrate their achievement of the SLO?
   E.g., test, research report, oral presentation, etc.

3. How will you check that the content of the task/test is
   (a) relevant to and
   (b) representative of the knowledge, skills, and/or dispositions stated in the SLO?
   E.g., experts review and evaluate (experts include you and also your peers); pilot the task/test and review student results; review the literature and compare to your content, etc.

4. What are appropriate interpretations and uses of the results/scores?
   E.g., student ability in writing, student preparedness for workplace demands, student readiness for graduation, effectiveness of a pedagogical technique, program effectiveness in preparing students for writing tasks, etc.
   E.g., use to guide program curriculum decisions, use to change program admissions standards, etc.

5. What might be the consequences associated with the interpretation and use of the results/scores?
   E.g., course content becomes more aligned with the SLO, narrowing of course content, single out individual faculty members, create common expectations of student performance, ethnic group differences exposed, increased participation of adjunct faculty in discussions of student learning, etc.

**Your answers to the questions above are the beginning of a validity argument**