

Validity Argument

Course Prefix, Number & Name of Performance Assessment:

Reviewers:

Date:

Instructions: Based on the results of Validity Inquiry Process and evaluation of data collected, develop a comprehensive Validity Argument. The Validity Argument should address 1) the instrument's purpose and intended interpretation and use of data, 2) quality of instrument and scoring guide, and 3) reliability of data collected. For the Quality and Reliability sections, include references to evidence and findings gathered through the review process.

1) Instrument's Purpose and Intended Interpretation and Use of Data:

- 1) Articulate the purpose of the performance assessment in terms of what it is intended to measure.
- 2) Describe how the results will be used (e.g., provide formative feedback, decisions regarding continuation in program, and/or program evaluation).
- 3) Note how the data collected through this performance assessment connects (i.e., triangulates) with data collected through other assessment instruments currently in use.

2) Quality of Instrument and Scoring Guide:

- **Domain Coverage (Content Analysis B1):** Describe the content experts or professional organization who developed the learning outcomes or standards that the performance assessment and/or the program or unit assessment plan addresses.
- **Generalizability (Content Analysis B4):** Discuss the performance assessments that make up the program or unit assessment plan in terms of how comprehensive they are in addressing the domain of knowledge (e.g., richness of the problems, projects, or scenarios incorporated into this performance assessments).
- **Content Quality (Content Analysis B2):** Describe the qualifications and expertise of the individuals involved in the design of the performance assessment.
- **Content Quality (Content Analysis B3):** Discuss how this performance assessment fits into the larger collection of assessment instruments (i.e., assessment plan) in terms representing a balance between complex, extensive performance tasks and shorter structure assignments.
- **Domain Coverage (Validity Inquiry Form Q1):** Discuss to what extent the performance assessment, based on the assignment instructions, adequately represents the domain of knowledge being evaluated and the outcome(s)/standard(s) aligned to it.
- **Content Quality (Validity Inquiry Form Q2):** Describe how this performance assessment incorporates an evaluation of process or application skills as well as content knowledge.

- **Cognitive Complexity (Validity Inquiry Form Q3):** Provide an analysis of how the performance assessment expectations promote higher order thinking skills and application of knowledge based on the Rigor/Relevance Framework or another approach to analyzing cognitive complexity.
- **Meaningfulness (Validity Inquiry Form Q4 and Student Survey):** Describe the content expert view of the meaningfulness and authenticity of the performance assessment in terms of the problem, project, or scenario presented. Discuss the results of the Student Survey in terms of their view of the meaningfulness and authenticity of the performance assessment.
- **Consequences (Validity Inquiry Form Q5):** Describe the consequences of the performance assessment in terms of effect on learners and their future learning, percent of overall grade, and/or high stakes aspects of the performance assessment in relation to when it is administered in the program of study. Also, discuss the method for establishing the pass-fail or cut score and the extent to which the benefits of implementing the assessment outweighs any unintended adverse consequences.
- **Fairness (Validity Inquiry Form Q5):** Discuss aspects of fairness of the performance assessment in relation to how all students are provided the same opportunity to gain the knowledge and skills necessary to complete the assessment, the format is familiar to students, and the time required to complete the assessment is reasonable.
- **Fairness (Metarubric, Criteria, Scale, Descriptions, and Overall Qualities):** Provide an overall description of the rubric or scoring guide in relation to the criteria, scale, descriptions, and overall qualities. Discuss performance assessment instructions and scoring guide in terms of students' ability to utilize the descriptions as feedback and accurately interpret the score.

3) Reliability of Data Collected:

- **Cost and Efficiency (Content Analysis B5 and Validity Inquiry Form Q7):** Articulate the benefits of the performance assessment in relation to the practicality of implementation. Document the data collection procedures and process for distributing results to students and evaluators.
- **Training of Evaluators:** Describe the process for training evaluators to implement the performance assessment (i.e., calibration training). Discuss tests and results of inter-rater agreement and systematic process for re-certification of evaluators.
- **Scoring Guide (Metarubric, Use of Rubric):** Note observations of evaluators based on actual implementation of the scoring guide or rubric. Discuss the application of the criteria and levels in relation to student work products including steps taken to ensure there is no missing information or instructional preparation causing students to perform poorly on any aspect of the assessment.
- **Data Analysis:** Provide an analysis of the effectiveness of the data reporting and the use of the results for each of the scoring criteria as well as results for criteria groupings that represent a

subscale, if applicable. Also provide an analysis of the results of the assessment as a whole. In the analysis, discuss the capacity of the instrument to discriminate or differentiate among student performance. In addition, discuss the analysis of results in relation to individual evaluators. In reviewing scores of individual evaluators, discuss evidence related to the use of the entire rating scale or consistent concentration only one or more elements. Include the results of any subsequent checks for inter-rater agreement, especially checks prompted by findings from the review of results.

References

- Baldwin, D., Fowles, M., & Livingston, S. (2005). Guidelines for constructed-response and other performance assessments. Princeton, NJ: Educational Testing Services.
- Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education, 37*(9), 830-837.
- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction* (6th Edition). White Plains, NY: Longman Publishers.
- International Center for Leadership in Education. (n.d.). Rigor/relevance framework. Retrieved from <http://www.leadered.com/our-philosophy/rigor-relevance-framework.php>.
- Kane, M. (2013). The argument-based approach to validation. *School Psychology Review, 42*(4), 448-457.
- Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15-21.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13-23.
- Pieper, S. L. (2012, May 21). Evaluating descriptive rubrics checklist. Retrieved from http://www2.nau.edu/~d-elearn/events/tracks.php?EVENT_ID=165
- Stevens, D. D., & Levi, A. J. (2005). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback and promote student learning*. Sterling, VA: Stylus Publishing, LLC.