

Standard Setting Course Description
Summer School in St. Petersburg, Russia, July 29 – August 4, 2018

Instructor: Dr. Mary J. Pitoniak

Course Title: Setting Standards on Educational Assessments

Description

This course is designed to provide participants with information about the wide range of considerations relevant to setting standards on educational assessments. These include how to choose a standard setting method, which methods are currently being used, and how to know if the cut scores set for an assessment yield valid interpretations within the context of a particular testing program.

Beginning with a historical overview, the session will provide a context regarding how decisions about standard setting are made today. The fundamentals of standard setting will be presented, including required steps for all methods. Methodologies currently being used most often in setting performance standards will be reviewed. Information on vertically moderated standards and adjusting committee-recommended cut scores will also be discussed.

Hands-on practice time will be given to allow participants to thoroughly understand the cognitive tasks involved in making the judgments for two of the most commonly used methods, Bookmark (Lewis, Mitzel, Mercado, & Schulz, 2012) and modified Angoff (Angoff, 1971; Hambleton & Pitoniak, 2006). This exercise will also prepare participants to plan and run Bookmark and modified Angoff standard setting workshops. This session is intended for anyone who needs to understand how to run a standard setting session and the complexities involved.

Finally, significant time will be devoted to studying the validity of standard setting procedures and the resulting cut scores. Using Kane's (2001) framework, the session will explore three sources of evidence: procedural, internal, and external. The use of the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014) for evaluating standard setting studies will also be described.

Prerequisite Skills or Knowledge Needed for Course Participation

This course requires familiarity with basic measurement concepts and vocabulary, such as reliability, validity, and fairness. At least basic knowledge of classical test theory and Item Response Theory (IRT) is needed for understanding of different types of standard setting methods. The target course participants are educational researchers including advanced graduate students, early career scholars, senior educational researchers and practitioners who are interested in improving their understanding and knowledge of standard setting.

Course Format

The format of the course will be part lecture and part hands on opportunities for students to participate in standard setting activities. Students are encouraged to bring with them descriptions of standards set on educational assessments in their districts/countries in order to evaluate the methods used to set them and the validity of the interpretations being made.

Learning Objectives

Upon completion of the course, participants will understand:

- Terminology used in standard setting
- Different types of standards that are used (e.g., relative vs. absolute, criterion-referenced vs. norm-referenced) and the advantages of each
- Role of judgment in standard setting and how it affects the design and implementation of standard-setting studies
- Meaning and use of classification errors
- Roles of different stakeholders in standard-setting process
- Categorizations of standard-setting methods
- General steps for standard setting studies
- Key aspects of training panelists in standard setting
- General features of many standard-setting methods
- In-depth features of modified Angoff and Bookmark methods
- Concept of definition of borderline (minimally competent) candidates
- Framework for evaluation validity of standard-setting studies
- Relevance of *Standards for Educational and Psychological Testing* to standard setting

Participants will be able to:

- Choose among standard-setting methods
- Set targets for panelist recruitment
- Develop performance-level descriptions
- Run standard setting study using modified Angoff and Bookmark methods
- Document and evaluate the validity of the results of a standard setting study

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 508–597). Washington, DC: American Council on Education.

Hambleton, R. K., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 433–470). Westport, CT: American Council on Education/Praeger.

Kane, M. (2001). So much remains the same: Conception and status of validation in setting standards. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 53–88). Mahwah, NJ: Erlbaum.

Lewis, D. M., Mitzel, H. C., Mercado, R. L, & Schulz, E.M. (2012). The Bookmark standard setting procedure. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (pp. 225–254). New York: Routledge.