Online Standard Setting Workshop

Charalambos Kollias
Brief introduction to Standard Setting
What is standard setting?

“In the field of achievement testing, standard setting is regarded as a process of decision making concerning the classification of the test/exam results in several successive, but limited number of levels of achievement.” (Kaftandjieva, 2010 p. 12)

A task
• to determine cut score on test corresponding to description of performance

High level of importance
• falling above or below cut score on test has importance consequences on test taker (Reckase, 2009)
A task

- to determine cutscore on test corresponding to description of performance High level of importance
- falling above or below cut score on test has importance consequences on test taker
Process and Validation for Setting Standards (Reckase, 2009)

Agency that Calls for the Standard
Policy Definition of Standard
Elaborated Description
Test Design
Translation of Policy and Elaborated Content to Value on Test Score Scale

Process and Validation for Setting Standards (Reckase, 2009)
Agency that Calls for the Standard & Policy Definition of Standard (Reckase, 2009)

Standards
• established/ set up

Agencies call for standards
• Standards
  – established/ set up

Policy definitions of standards
  – whether standard is difficult to reach (academic recognition) or whether standard is for the minimum requirements of all individuals
  – first setting of a standard
• Standards
  – established/ set up

Policy definitions of standards
  – whether standard is difficult to reach (academic recognition) or
  whether standard is for the minimum requirements of all individuals – first setting of a standard
Elaborated Description
Test Design (Reckase, 2009)

Elaboration of policy definitions
• Description of what someone exceeding standard can do

Standard operationalized as score on a test.

Test Design and Content
• Test design to be consistent with policy definition
  — Test to represent content specified in elaborated description.
Test Design and Content

• Test design to be consistent with policy definition
  – Test to represent content specified in elaborated description.
Translation of Policy and Elaborated Content to Value on Test Score Scale (Reckase, 2009)

Standard Setting
• Policy descriptions & elaborated descriptions translated into numerical language of test score

Translation requirements:
• people who are ...
  – fluent in language of policy definitions
  – elaborated descriptions
  – language of test score

Translation of Policy and Elaborated Content to Value on Test Score Scale (Reckase, 2009)

Standard Setting
• Policy descriptions & elaborated descriptions
translated into numerical language of test score

Translation requirements:

• people who are ... – fluent in language of policy definitions – elaborated descriptions – language of test score
Translation of Policy and Elaborated Content to Value on Test Score Scale (Reckase, 2009)

Problem:
• Difficult to find people with all capabilities

Solution:
• Processes to help with translation process
  — standard setting methods

Standard setting methods
• methods for facilitating translation from verbal descriptions to numerical descriptions

Participants
• Judges/panelists = translators
Solution:

• Processes to help with translation process
  – standard setting methods

• methods for facilitating translation from verbal descriptions to numerical descriptions

Participants

• Judges/panelists = translators
Standard setting methods

Test centered

• Subject matter experts (SMEs) review & discuss characteristics of students within or at “border” of achievement level, then carefully consider knowledge and skill demands of each item on test. Then make judgment regarding probability of examinee at borderline of achievement level getting the answer correct.

Examinee-centered

• Testing groups of examinees already classified in achievement levels and using their performance to establish cut scores.
probability of examinee at borderline of achievement level getting the answer correct.

**Examinee-centered**

- Testing groups of examinees already classified in achievement levels and using their performance to establish cut scores.
Facilitating the derivation of valid standards

1. Selecting qualified and sufficient number of panelists
2. Providing adequate training of panelists
3. Facilitating panelists’ discussion of ratings
4. Providing qualitative feedback on provisional judgments
5. Facilitating discussions on final cut score for consensus
6. Conducting a comprehensive survey of panelists’ impressions
7. Deriving cut-score intervals

(Sireci, Randall, & Zenisky, 2012)
on final cut score for consensus 6. Conducting a comprehensive survey of panelists’ impressions 7. Deriving cut-score intervals (Sireci, Randall, & Zenisky, 2012)