

Issues to Consider When Undertaking a Standard Setting Exercise

October 28, 2015



Presenters



**Liberty Munson,
PhD**
Psychometrician
Microsoft Learning
Experiences



Ada Woo, PhD
Director of
Measurement and
Testing
National Council of
State Boards of
Nursing



**Kurt F. Geisinger,
PhD**, Director,
Buros Center for
Testing, University
of Nebraska-
Lincoln



**Manny Strahle,
PhD, GISF**
Founder &
President
AERE

Session Format

- Display a question to the presenters
- Presenters respond
- Audience responds
- Display next question

Topics

- Raters and SMEs
 - Methodology
 - Facilitation
 - Business Operations
 - Accreditation
 - Evaluation

Raters and Subject Matter Experts (SMEs)

Raters and SMEs (MS)

How many SMEs should participate in the process?

- At least 5
- Representation of target audience
- Likelihood for legal challenges/legal environment
- Purpose/intended use

What evidence is available to support the number of SMEs to be used for a standard setting?

- Fowell, S. L, Fewtrell, R., & McLaughlin, P. J. (2006).
- Cizek, G. J. (1996).
- Norcini, J. J.,& Shea, J. A. (1997).

Raters and SMEs (Cont.) (KG)

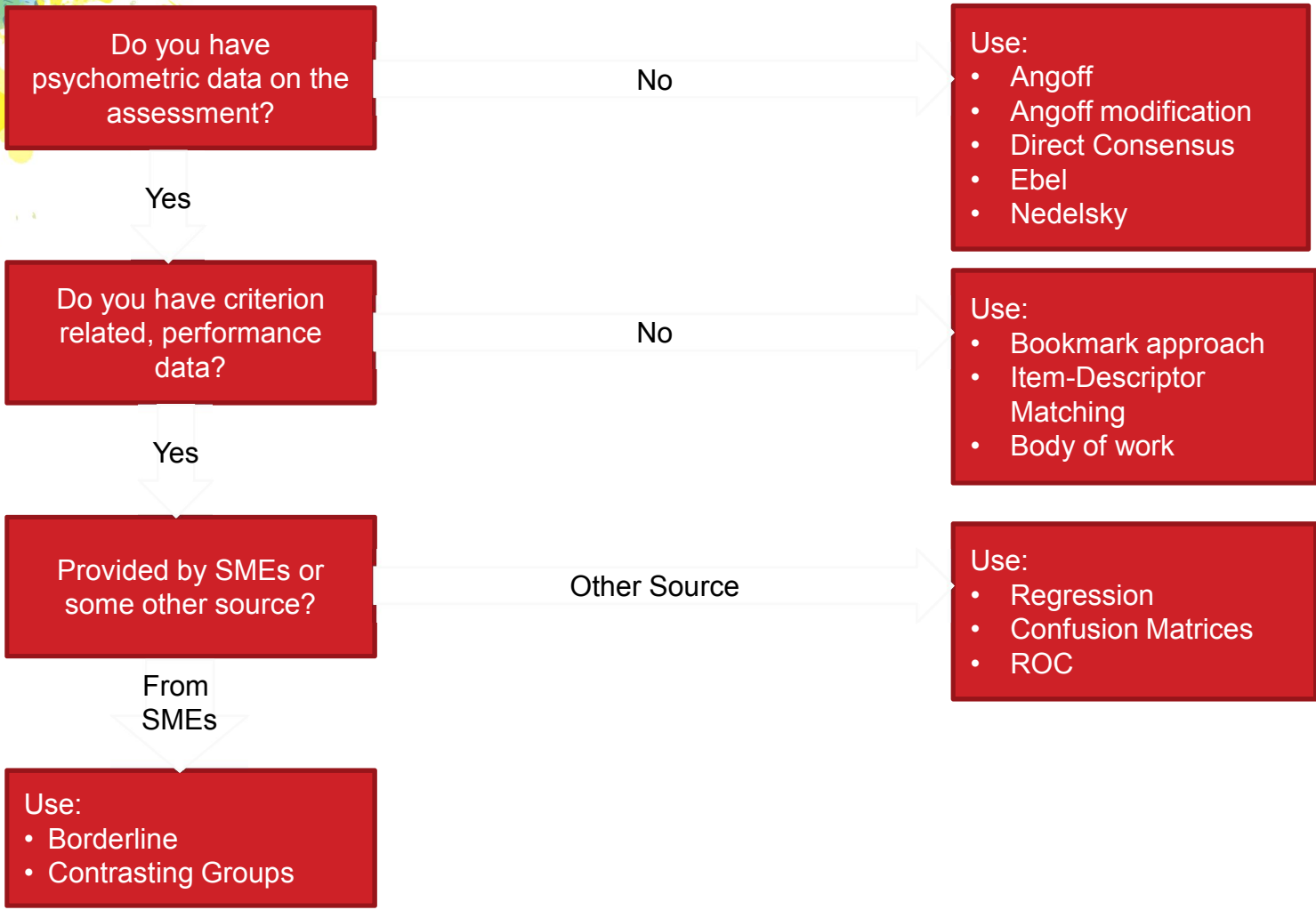
What factors should one use in selecting panelists for a standard-setting?

- Possess knowledge and skills needed to make appropriate judgments
- Their knowledge, skills, and expertise will not be questioned
- Able to think independently
- Represent all reasonable aspects of profession
- Understand that process is for selecting entry level professionals
- Cooperative
- Independent of test authors and test plan developers (if accreditation is a goal)

Methodology

Methodology (LM)

What methods do you use, why, and when? What are the decision points when selecting a method?



Methodology (Cont.) (MS)

Is it possible to only use the compromise method such as the Beuk and Hofstee without another method such as the Angoff or Bookmark? **YES, BUT...**

- I would not recommend for higher stake exams
- More suitable for **lower stakes exams**
- Cizek: Ensure you are following an acceptable process and method
- Why not consider non-quantitative compromise techniques?

Methodology (Cont.) (AW)

What happens if the passing standard is too high or too low? (AW)

- Defend it! What? Defend It!
- Rerun it
- Adjust it
- Use SEs/CSEs consider using smaller intervals
- Conduct compromise method or alternative method
- Exclude outlier SMEs if their results are extremely different if you have enough (but provide a rationale) or if they do not have adequate knowledge

Methodology (Cont.) (LM)

How do you determine the final cut score? What factors are taken into consideration? How important are programmatic passing rate goals?

What kinds of adjustments to a proposed standard (from a panel) should one consider, and why?

- Results of standard setting process +/- 3 SEs
- Input from SMEs on where in range is most appropriate
- Expected passing rates/Programmatic goals
- Purpose/Intended use
- Validation using judgments of specific individuals

Methodology (Cont.) (KG)

Should you provide feedback to panelists during the standard setting process? If yes, what types of feedback could you provide?

- Yes, if available
- Pass rate impact (consequence) and comparison of individual panelist's ratings (rater location)

Reference: Reckase, M. D. (2001)

How should an organization disseminate standard setting results?

- Press release
- Newsletter
- Organization website
- Professional journals
- Methodology papers

Which is worse – Type One or Type Two errors in classification?

- Depends upon the context
 - » A situation where individuals are needed may be a factor
 - » Impact of unqualified practitioners on users of services is also a critical factor
- More often than not, Type One errors are worse

Methodology (Cont.) (KG)

If tests are equated one year to the next, is a separate standard setting study still needed?

Generally not, but it depends on whether the nature of the test and the position has changed. (Of course, if the test has changed substantially, forms should not be equated.) But the question of how many years one could go without re-doing the standard-setting study is a great question!

Facilitation

What should one do when panelists complain about one or more test questions?

- Note the issue with the question and review with panelists
- Take notes and refer to the test developers
 - » Do not let the standard-setting panel dwell on it
- If the question is unfair, remove it, and rate it as 100 or consider imputing the missing data

Facilitation (Cont.) (AW)

What happens if the subject matter experts in the standard setting workshop cannot come to a consensus?

- Continue to facilitate conversations to gather everybody's perspective
 - » Use circular questioning, which is asking everybody their opinion (Carl Whitaker)
- OK if consensus is still not achieved
 - » If many items don't reach acceptable consensus, consider reviewing methodology (e.g., definition of minimally qualified)

Facilitation (Cont.) (MS)

What are some facilitation issues that should be considered, and why?

- Managing dynamics
 - » Resistance, power, and fears
 - hidden agendas, disengagement, disruptive SMEs
- Managing expectations
 - » Basic training, time, goals
- Managing logistics and outcomes
 - » Lack of consensus, unqualified SMEs, power differentials among the SMEs

Facilitation (Cont.) (AW)

Can standard setting be conducted virtually? If yes, what makes a successful virtual standard setting workshop?

- Not recommended for first standard setting workshop
- Established programs and processes
- SMEs should be familiar with the exam, organization, process
- Keep shorter than in-person standard setting workshop
- Challenges with rating databases and keeping raters engaged (e.g., SMEs may leave the workshop)
- Be mindful of item security
- Consider a hybrid methodology

Business Operations

Business Operations (AW)

What is the cost of conducting a standard setting workshop?

- Cost of SMEs (e.g., compensation, travel, accommodation, food, etc.)
- Staff time (e.g., materials, database preparation, facilitation and reporting)
- Technology (e.g., WebEx, internet, computers)
- Consultation (e.g., psychometricians and legal)

Business Operations (Cont.) (AW)

Should an organization disseminate standard-setting results? If so, what information should be made public?

- Accreditation requires sharing process and outcomes
- Share scaled cut score (at Microsoft all exams scaled to passing score = 700) rather than raw cut score
 - » Makes communication complicated, though, b/c candidates assume 700 = 70%
- Legal professionals have recommended simply Pass/Fail as the report and nothing else
- What should the organization do to facilitate performance feedback?

Business Operations (Cont.) (MS)

What happens if your stakeholders don't agree with the recommended cut score results? How do you manage it?

- Find out why your stakeholders are disagreeing
- Provide documentation of standard setting process, including raters' credentials
- Provide pass rate predictions when appropriate
- Discuss with stakeholder groups
- Conduct a policy panel to evaluate cut score

Accreditation

Accreditation (MS)

What should I be aware of for accreditation such as ISO/IEC 17024 or NCCA?

- Process rather than outcome focused
 - » Recommended because rigorous requirements result in more better exam-related decisions (even if you choose not to be accredited)
 - » Common and frequently used processes
 - Angoff
 - External committee that is unbiased/impartial of influencing cut score decision
 - Report especially including the demos of SMEs

Evaluation

Evaluation (KG)

How can one best validate the appropriateness of a standard once it is set?

- Does it match expected levels of competency in those passing?
- Does the proportion passing approximate the expected proportion?
- What kinds of complaints are received?

Evaluation (Cont.) (KG)

What should one do when there are complaints about specific questions after the meeting?

- Take notes
- Assure complainant that their views will be thoroughly considered
- Make note of any re-wording or other minor changes
- Do not let the meeting devolve into a session evaluating items rather than setting a proposed standard

thank
you!

Liberty Munson – Liberty.Munson@microsoft.com

Ada Woo – awoo@ncbsn.org

Kurt Geisinger – kgeisinger@buros.org

Manny Straehle – manny@aerexperts.com