



Integrating Policymakers into the Standard Setting Process: Licensure and Certification Perspectives


Brett P. Foley, PhD
 Director, Professional Credentialing &
 Senior Psychometrician

Corina M. Owens, PhD
 Psychometrician

VALIDITY
 Fair, Reliable, Secure

CONDUCT
 STANDARD SETTING

1



What do the Standards Say about Cut Scores?

1. "The level of performance required for passing a credentialing test should depend on the knowledge and skills necessary for credential-worthy performance in the occupation or profession and **should not be adjusted to control the number or proportion of persons passing the test**" (AERA, APA, NCME, 2014, Standard 11.16, emphasis added)
2. Adjusting the cut score to regulate the number of accredited candidates entering the profession "...**raises serious problems for the technical quality of the test scores and threatens the validity of the interpretation of a passing score as indicating entry-level competence**" (AERA, APA, NCME, 2014, p. 177, emphasis added)

www.alpinetesting.com

2

What do the Standards Say about Cut Scores?



3. "The procedure and results of the standard-setting workshop should be clearly documented, including the method used to determine the recommended cut score(s), the **resulting cut score recommendations**, and an estimate of variability in panelists' recommendations. The **final cut score(s) adopted and used in practice** should also be clearly reported." (Buros, 2017, Standard 6.9, p. 10)
4. Certifying organizations must use criterion-referenced standard setting methods.... **Adjustments to the results of the standard setting process may be made if necessary** after the procedures have been completed. However, this should be done in a well-reasoned, methodical, and psychometrically sound fashion with justification provided for any adjustments." (ABSNC, 2016, p. 26).
5. "The following information must be retained in full detail by the organization: ...Results of standard setting studies, including who participated, training received, methodology(ies) used, **results of study versus actual standard applied** (and rationale for any deviations)." (ICE, 2011, p. 9)

www.alpinetesting.com

3

Why the Apparent Discrepancy?



- Standard setting is ultimately a policy decision, because...
 - Policymakers are the ones with the authority and responsibility to determine final cut scores
 - Policymakers have the right and responsibility to consider factors beyond the results of a standard setting study

www.alpinetesting.com

4

“But in our organization...”



- “We always just have the psychometrician tell us what the cut score should be.”
- “We always just use the average rating from the standard setting panel.”

- Delegation of decision making: Policy Decision!
- Adoption of a decision-making rule: Policy Decision!

- Work/Decisions can be delegated.
- Responsibility remains with policymakers.

5

Role of Policy Makers



1. “...the role of a standard-setting panel is to provide information, in the form of recommended performance standards, to the decision-making body. That group may then choose to make adjustments to the passing standards before implementing them.” (Hambleton & Pitoniak, 2006, p. 464)
2. “It is [policy bodies] that poses the authority and responsibility for setting standards; it is the role of standard setting panels to provide informed guidance to those responsible for the act of setting, approving, rejecting, adjusting or implementing any cut scores.” (Cizek, 2012, p. 6; emphasis in original)
3. “After considering all the relevant information, the policymakers will have to choose the operational cutscores. That choice is a policy decision, and the policymakers will have to consider the likely consequences for their decision and accept the responsibility for it.” (Zeiky, Perie, & Livingston, 2008, p. 163)
4. “Decision makers have a responsibility and a right to consider factors beyond the standard-setting panels’ recommendations when determining the final cut scores for assessments.” (Geisinger & McCormick, 2010, p. 44)

6

Premise of this Session



- Because...
 - Policymakers are the ones with the authority and responsibility to determine final cut scores, and...
 - Rationale should be provided when policymakers deviate from standard setting study recommendations

- It follows that policymakers should have...
 - a formal, systematic role in the standard setting process
 - sufficient information to guide their decisions
 - instruction in...
 - the standard setting process
 - interpretation of standard setting recommendations

www.alpinetesting.com

7

Framework for Systematically Integrating Policymakers



- Structure:
 - Responsibility Assignment Matrix

- Organization:
 - Entities assigned formal roles within overarching standard setting process

www.alpinetesting.com

8

Responsibility Assignment Matrix



		Who is Involved?					
		Organization Staff	Psychometric Support	Subject Matter Experts	Examination Committee	Board of Directors	External Stakeholders
What is to be determined?	Cut Score Approval Process						
	Type(s) of Feedback Provided						
	Standard Setting Method(s)						
	Panel Composition						
	Defensible Cut Score Range						
	Final Cut Score						

Responsibility Role

- R = Responsible/Recommender**
Recommend a solution and do the work to achieve the task.
- A = Accountable/Approver**
Answerable for the correct completion of the task; final authority.
- C = Consulted**
Provide input on task.
- I = Informed**
Kept up-to-date on progress and decisions made.

www.alpinetesting.com

9

Bad Example



		Who is Involved?					
		Organization Staff	Psychometric Support	Subject Matter Experts	Examination Committee	Board of Directors	External Stakeholders
What is to be determined?	Cut Score Approval Process		A				
	Type(s) of Feedback Provided		A				
	Standard Setting Method(s)	I	A				
	Panel Composition	I	A				
	Defensible Cut Score Range		A	R			
	Final Cut Score	I	A				

Responsibility Role

- R = Responsible/Recommender**
Recommend a solution and do the work to achieve the task.
- A = Accountable/Approver**
Answerable for the correct completion of the task; final authority.
- C = Consulted**
Provide input on task.
- I = Informed**
Kept up-to-date on progress and decisions made.

www.alpinetesting.com

10

Examples of Applying The Framework



- Professional IT Certification Program
- Medical Licensure Program
- Example materials



www.alpinetesting.com

11



Example Framework: IT Certification Program




VALIDITY
Fair, Reliable, Secure

www.alpinetesting.com

12

12

IT Certification Example



	Strategic Organizational Staff (VP for Assessment)	Operational Staff (Director of Examination Program)	Psychometric Support	Subject Matter Experts	External Stakeholders
Cut Score Approval Process	A	R	C		
Type(s) of Feedback Provided		A	R	I	
Standard Setting Method(s)		A	R	I	I
Panel Composition		A	C	I	I
Defensible Cut Score Range		I	A	R	
Final Cut Score		A	C		
Calculate Evaluative Statistics based on Cut Score		A	R		
Review Results regarding Keeping or Revising Cut Score		A	R	C	
Policy Decisions regarding Cut Score Revisions	A	I	C		

R = Responsible/Recommender
Recommend a solution and do the work to achieve the task


A = Accountable/Approver
Answerable for the correct completion

C = Consulted
Provide input on task

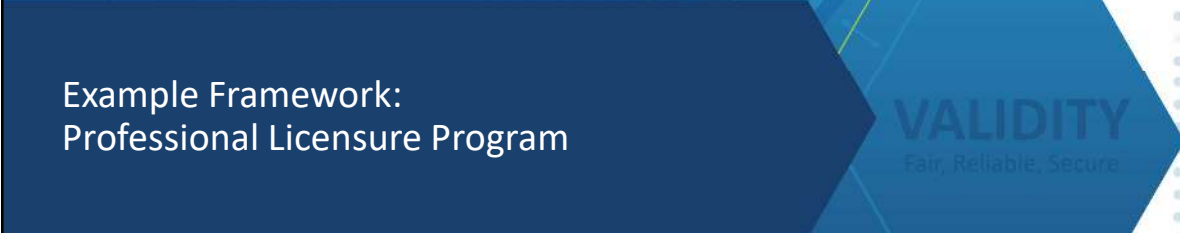
I = Informed
Kept up-to-date on progress and decisions made

www.alpinetesting.com 13

13



Example Framework: Professional Licensure Program



VALIDITY

Fair, Reliable, Secure

www.alpinetesting.com 14

14

Licensure Program Example



	Who is Involved?					
	Organization Staff	Psychometric Support	Subject Matter Experts	Examination Committee	Board of Directors	External Stakeholders
What is to be determined? Cut Score Approval Process	C	R			A	
Type(s) of Feedback Provided	C	R			A	
Standard Setting Method(s)	A	R			I	I
Panel Composition	A	C			I	I
Defensible Cut Score Range	I	A	R		R	
Final Cut Score	C	C			A	

Responsibility Role

- R = Responsible/Recommender**
Recommend a solution and do the work to achieve the task
- A = Accountable/Approver**
Answerable for the correct completion of the task; final authority
- C = Consulted**
Provide input on task
- I = Informed**
Kept up-to-date on progress and decisions made

www.alpinetesting.com

15

Board of Directors Involvement




- Pre-Standard Setting
 - **Participate in the discussion and approve design**
- Standard Setting
 - **Observe**
- Post-Standard Setting
 - **Review** Data
 - **Ask questions** and **discuss** variables
 - **Approve** final cut score

Benefit:

- More input from various sources
- Increases understanding, buy-in, and advocacy
- Creates balance in the decision-making process

www.alpinetesting.com

16



CONDUCT STANDARD-SETTING


VALIDITY
Fair, Reliable, Secure

www.alpinetesting.com

17

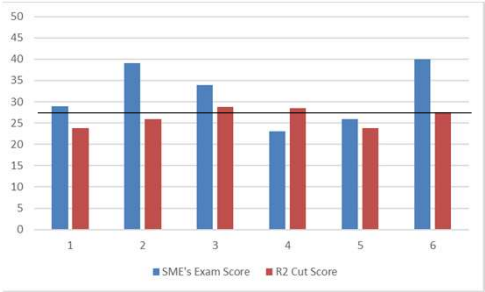
Example Policymaker Feedback: IT Certification Program

17

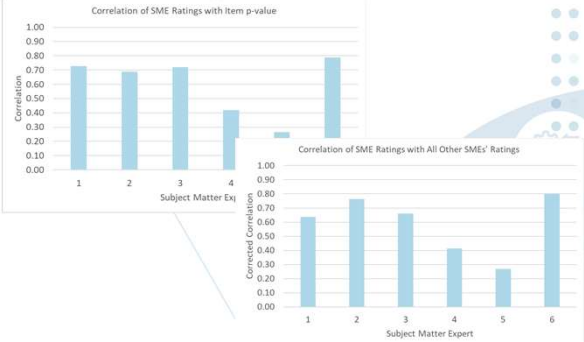


Types of Feedback – Initial Cut Score Decision

- Standard setting panelists scores as compared to Round 2 individual cut scores and mean
- Relationships Amongst Available Data



Subject	SME's Exam Score	R2 Cut Score
1	28	24
2	38	26
3	32	29
4	23	28
5	26	24
6	39	27



Subject Matter Ex	Correlation
1	0.72
2	0.68
3	0.70
4	0.42

Subject Matter Expert	Corrected Correlation
1	0.65
2	0.75
3	0.65
4	0.42
5	0.28
6	0.78

18

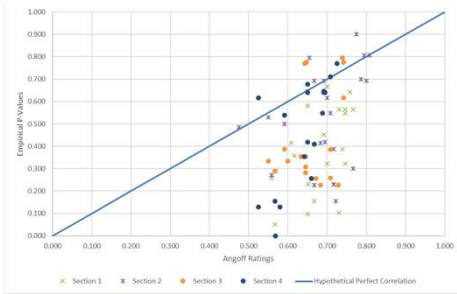
18

Types of Feedback – Single Form Assembly with No Beta



- Cut Score Verification Analysis

- Comparison between SS ratings and Empirical Results
 - r across full form and within section
 - Distribution of Angoff ratings and empirical results
- Illustration of chosen cut score plus error band with associated pass rates



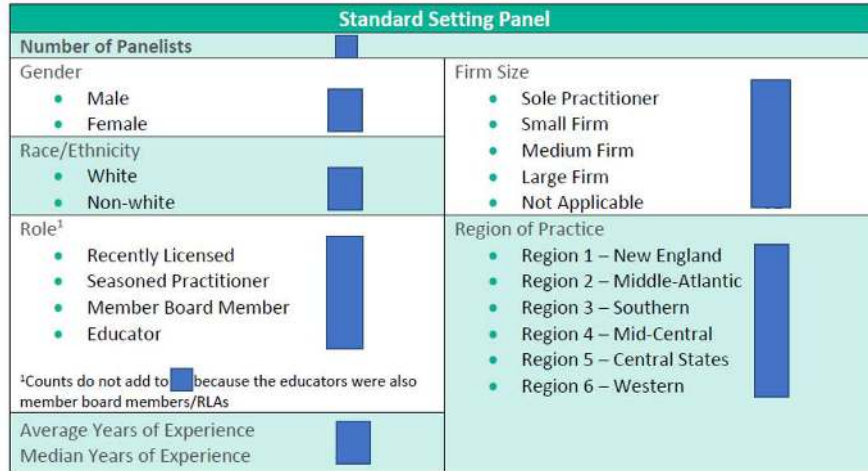
Range of p-value/ratings	Count		
	Angoff	P-Value	In Common
0.00 - 0.20	0	14	0
0.21 - 0.40	0	32	0
0.41 - 0.60	25	29	7
0.61 - 0.80	77	27	24
0.81 - 1.00	6	6	2

Score	Frequency Form A	Reverse Cumulative Frequency Percent (Pass Rate)	
		Form A	
Minus 2 S.E.	35	17	72%
Minus 1 S.E.	36	27	69%
	37	18	63%
Chosen Cut	38	29	59%
Plus 1 S.E.	39	31	54%
	40	30	47%
Plus 2 S.E.	41	21	41%

Example Policymaker Feedback: Professional Licensure Program

www.alpinetesting.com 20

SS Panel Information



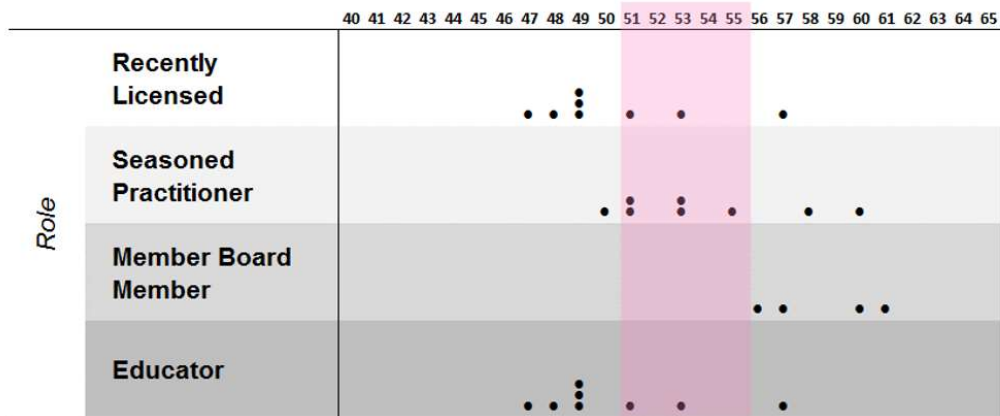
www.alpinetesting.com

21

Panel Recommendations by Stakeholder Type



Distribution of Panelist Cut Score Recommendations



www.alpinetesting.com

22

Impact Data with Recommended Ranges



Pass Rates by Demographic Group						
Cut Score	Demographic Group					
	All	Male	Female	White	Non-White	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	
XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	

Exam Committee:
Acceptable Cut Scores

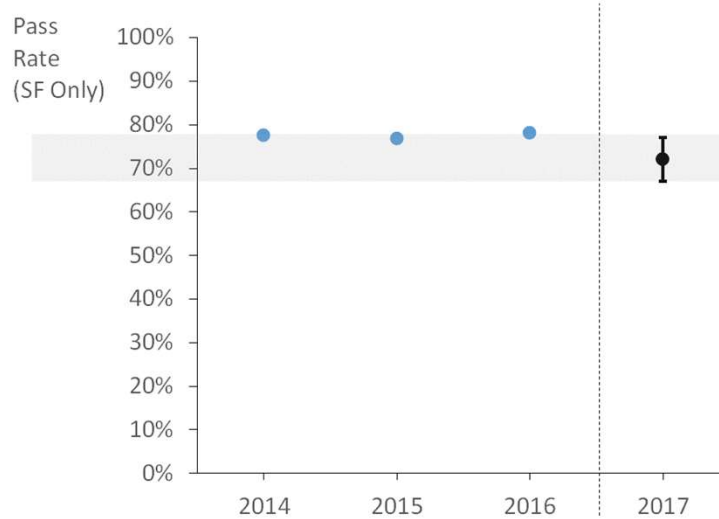
Recommended
Defensible Range

Exam Committee:
Acceptable
Pass Rates

www.alpinetesting.com

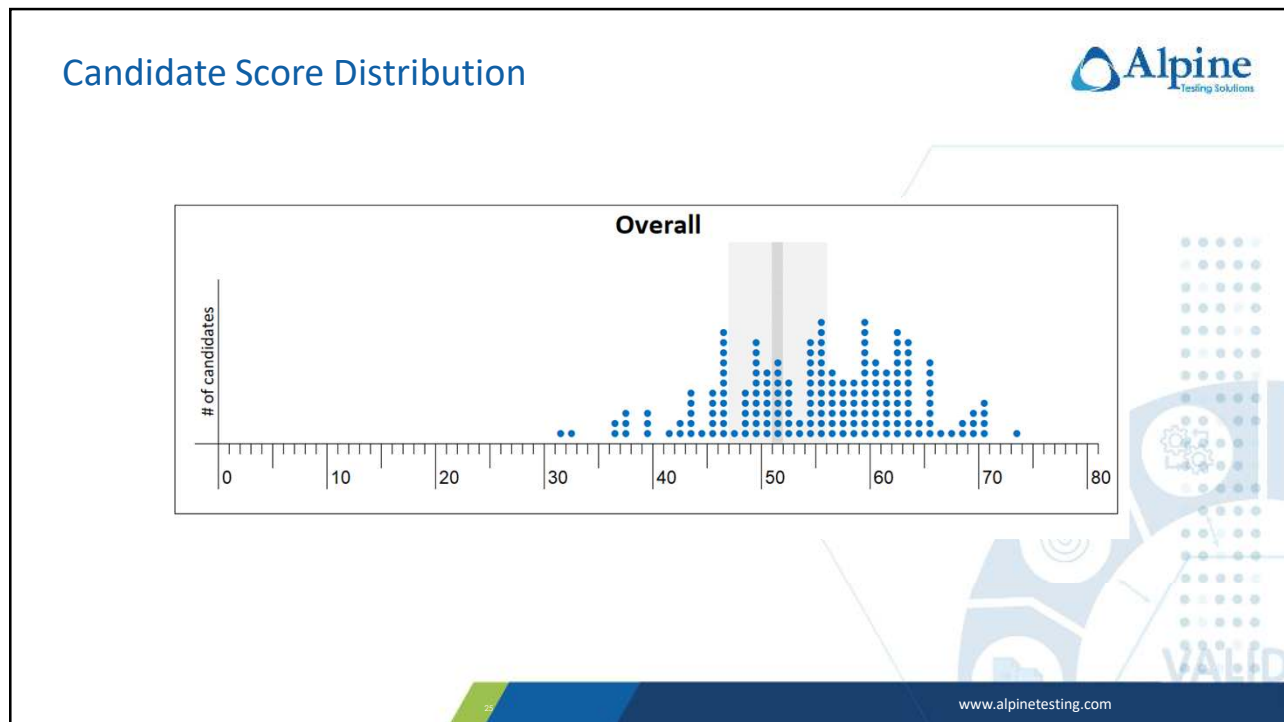
23

Historic Pass Rate Comparison



www.alpinetesting.com

24



25

Concluding remarks



- Responsibility Assignment Matrix should be...
 - Developed prior to operational standard setting activities
 - Used in formal documentation
- Formal inclusion of policymakers may increase...
 - Understanding of standard setting process
 - Representativeness of involved stakeholders
 - Stakeholder buy-in
 - Validity of classification decisions

www.alpinetesting.com

26

Concluding remarks



- Questions?
- Ideas?

www.alpinetesting.com

27

References



- Accreditation Board for Specialty Nursing Certification. (2016). *Accreditation standards*. Mt Laurel, NJ: ABSNC.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Association.
- Buros Center for Testing. (2017). *Standards for accreditation of testing programs*. Lincoln, NE: Board of Regents of the University of Nebraska and Author.
- Cizek, G. J. (2012). An introduction to contemporary standard setting: Concepts, characteristics, and contexts. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 3-14). New York, NY: Routledge.
- Geisinger, K. F., & McCormick, C. M. (2010). Adopting cut scores: Post-standard-setting panel considerations for decision makers. *Educational Measurement: Issues and Practice*, 29(1), 38-44.
- Hambleton, R. K., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 433-470). Westport, CT: ACE/Praeger.
- Institute for Credentialing Excellence. (2011). *Benchmarking best practices*. Washington, DC: Author.
- Zieky, M. J., Perie, M., & Livingston, S. A. (2008). *Cutscores: A manual for setting standards of performance on educational and occupational tests*. Princeton, NJ: Educational Testing Service.

www.alpinetesting.com

28