



Institute for Credentialing Excellence

EXG2021

The Exchange is
Where We Will Tackle
Challenges Together



NOVEMBER
15-18
2021

NASHVILLE, TN

Integrating Decision-makers Into the Standard Setting Process: A How-to for Licensure and Certification

Corina M. Owens, PhD
Brett P. Foley, PhD
Alpine Testing Solutions, Inc.



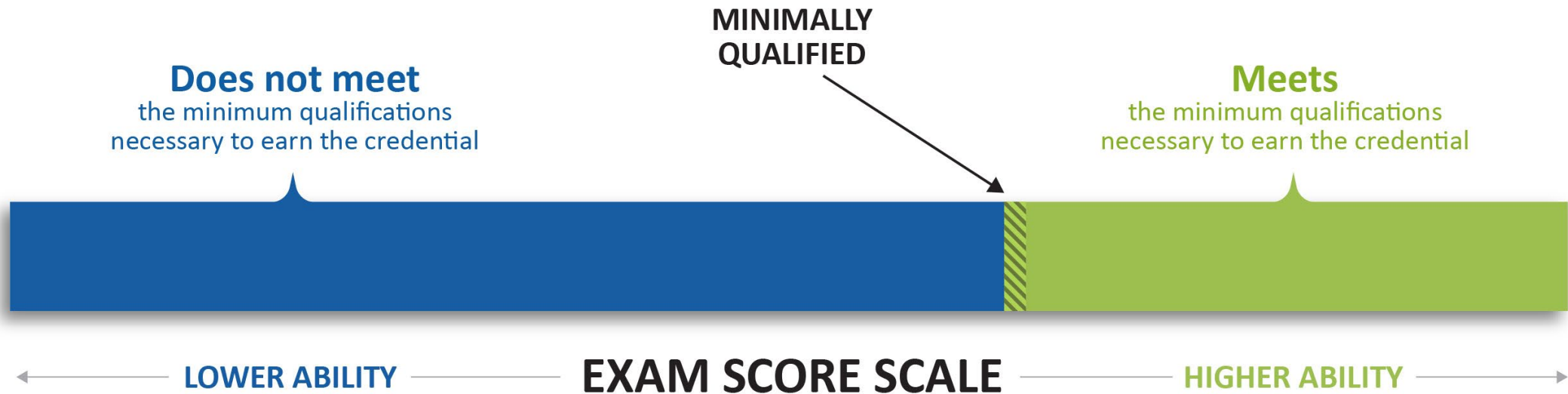
November 15,
2021

Introductions

Introductions

- Who are you?
- Where are you from?
- What do you do?
- What type of exam development experience do you have?
- Best snack/beverage you have snuck into a movie theater?

What is a standard setting study? Why conduct a standard setting?



- A process used to distinguish between candidates who DO NOT meet the minimal qualifications needed for the credential and candidates who DO meet the minimal qualifications

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)

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)

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

“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.

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)

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

“Good” cut scores...

- are supported by validity evidence
- have broad stakeholder buy-in
- are accepted as fair and reasonable

Framework for Systematically Integrating Policymakers

- Structure:
 - Responsibility Assignment Matrix
- Organization:
 - Entities assigned formal roles within overarching standard setting process

Example of Meaningful Involvement: Board of Directors

- 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**

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

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

Good Examples

	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	

	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	R	R		C	A	
Type(s) of Feedback Provided	R	R		C	A	
Standard Setting Method(s)	R	R		C	A	
Panel Composition	R	C		C	A	
Defensible Cut Score Range	I	A	R	R	I	
Final Cut Score	C	C		R	A	

		Who is Involved?				
		Strategic Organizational Staff (VP for Assessment)	Operational Staff (Director of Examination Program)	Psychometric Support	Subject Matter Experts	External Stakeholders
What is to be determined?	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		

Responsibility Role

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

Who is involved

Who is involved?

- Organization staff
 - Strategic Staff (VP for Assessment)
 - Operational Staff (Director of Examination Program)
- Psychometric support
- Subject matter experts

- Task Force
- Exam Committee
- Board of Directors
- External Stakeholders

**May want to call out specific individuals or job roles*

What is to be determined?	Who is involved?						Responsibility Role
	Organization Staff	Psychometric Support	Subject Matter Experts	Examination Committee	Board of Directors	External Stakeholders	
Cut Score Approval Process							R = Responsible/Recommender Recommend a solution and do the work to achieve the task.
Type(s) of Feedback Provided							A = Accountable/Approver Answerable for the correct completion of the task; final authority
Standard Setting Method(s)							C = Consulted Provide input on task
Panel Composition							I = Informed Kept up-to-date on progress and decisions made
Defensible Cut Score Range							
Final Cut Score							

Cut Score Approval Process

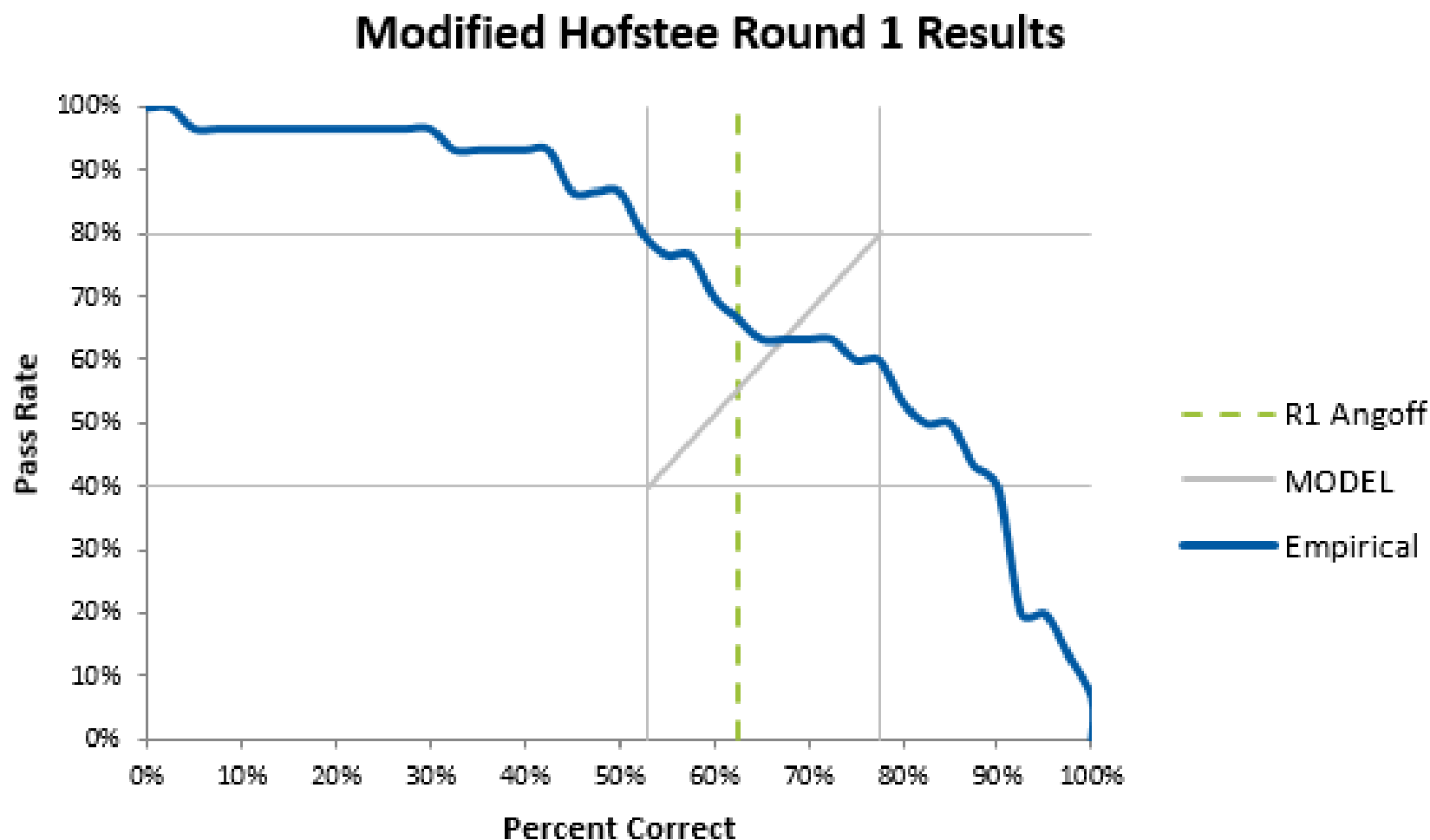
Standard Setting Method(s)

- Methods
 - Angoff (and variations)
 - Bookmark
 - Hofstee
 - Contrasting Groups
 - Borderline Groups
- Activities
 - SMEs take exam

Rating Form – Hofstee

Hofstee Ratings		
What is the minimum pass-rate that you would accept for the exam?	70%	(0% to 100%)
<i>The pass-rate is the percentage of candidates who pass the exam and as a result earn the certification.</i>		
What is the maximum pass-rate that you would accept for the exam?	80%	(0% to 100%)
<i>The pass-rate is the percentage of candidates who pass the exam and as a result earn the certification.</i>		
What is the minimum cut-score (expressed as a percentage) that you would accept for the exam?	55%	(0% to 100%)
<i>The cut-score is the score that a candidate must meet or exceed in order to pass the exam and as a result earn the certification.</i>		
What is the maximum cut-score (expressed as a percentage) that you would accept for the exam?	70%	(0% to 100%)
<i>The cut-score is the score that a candidate must meet or exceed in order to pass the exam and as a result earn the certification.</i>		

Hofstee Results:



Panel Composition

- May want to break out into two parts:
 - Panel Composition (structure, size, representation)
 - Panel Selection (selection of actual panelists)

Defensible Cut Score Range

Type(s) of Feedback Provided

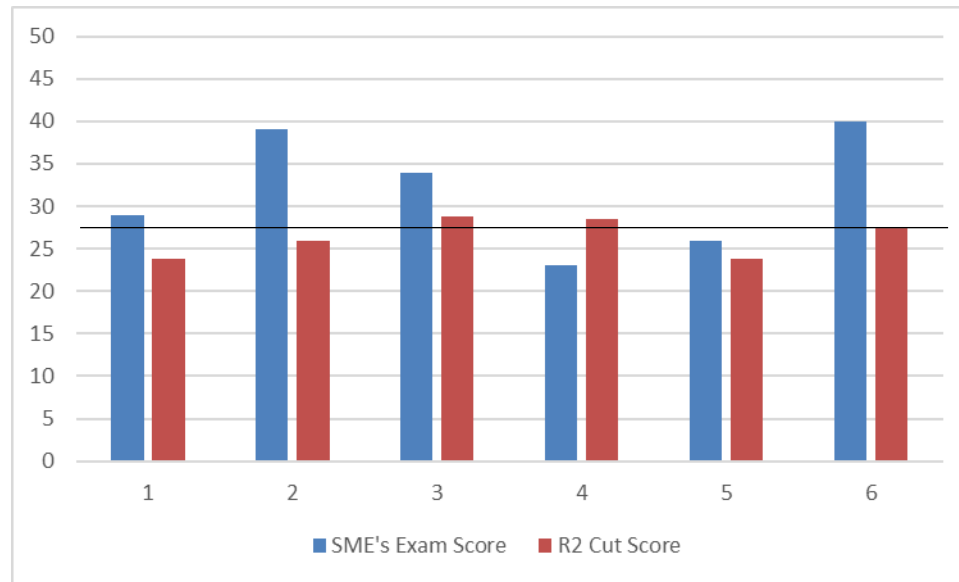
Two Main Feedback Types

- Feedback to SMEs
 - Shared during standard setting process
 - e.g., item p-values, between-round feedback
- Feedback to Decision Makers
 - Panel information
 - Panel Results (averages, variability)
 - Defensible cut score ranges
 - Impact data (e.g., projected pass rates)
 - Historic pass rates

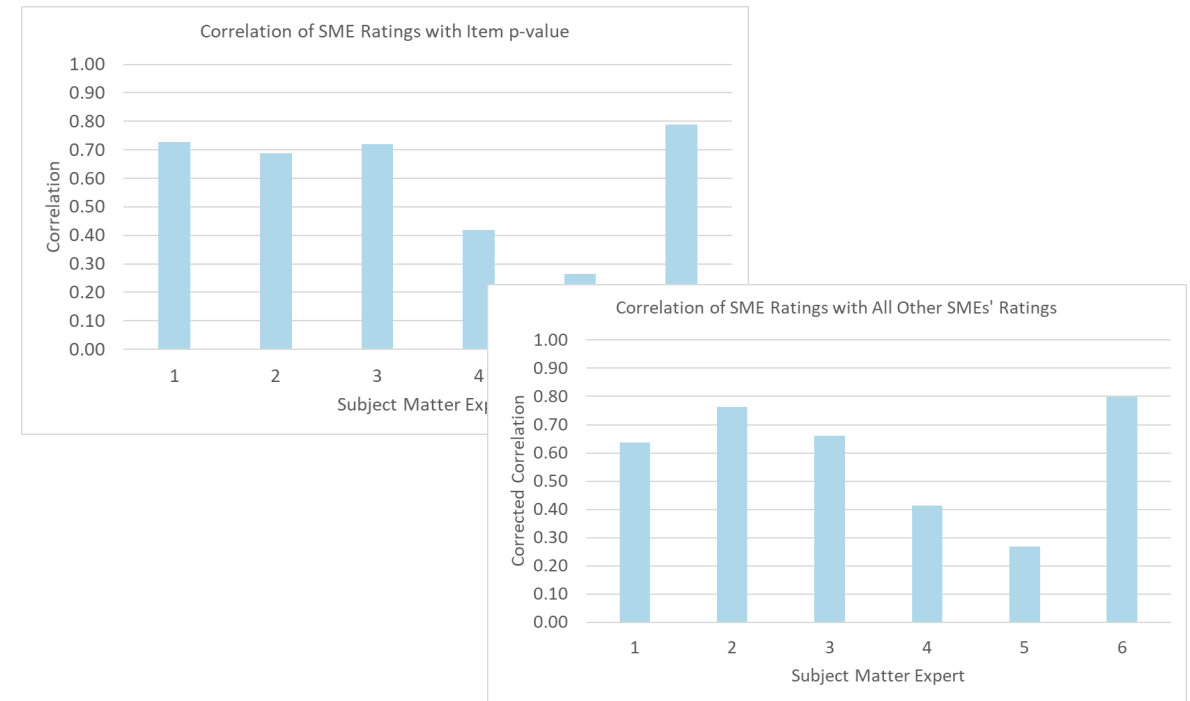
Example Policymaker Feedback: IT Certification Program

Types of Feedback – Initial Cut Score Decision

- Standard setting panelists scores as compared to Round 2 individual cut scores and

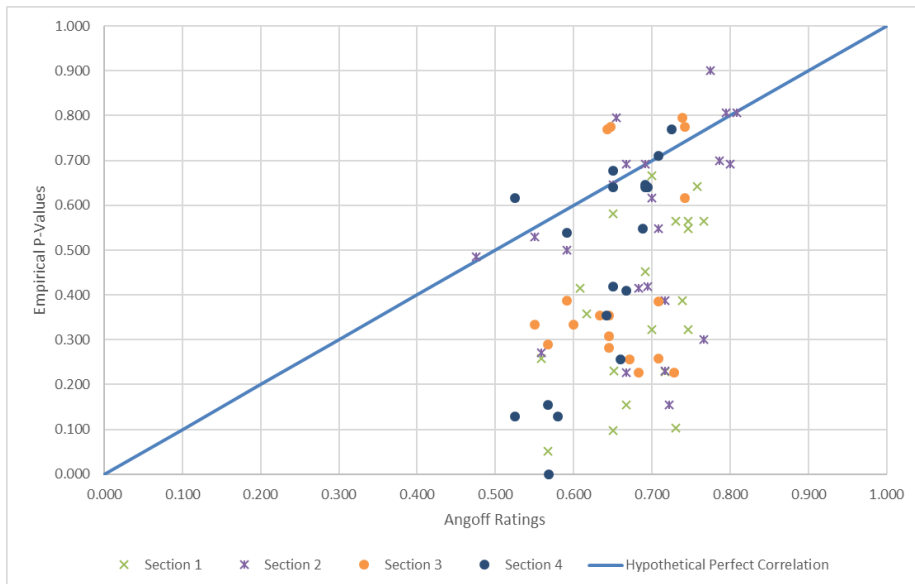


- Relationships Amongst Available Data



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

SS Panel Information

Standard Setting Panel	
Number of Panelists	
Gender	Firm Size
<ul style="list-style-type: none">MaleFemale	<ul style="list-style-type: none">Sole PractitionerSmall FirmMedium FirmLarge FirmNot Applicable
Race/Ethnicity	Region of Practice
<ul style="list-style-type: none">WhiteNon-white	
Role ¹	
<ul style="list-style-type: none">Recently LicensedSeasoned PractitionerMember Board MemberEducator	
¹ Counts do not add to because the educators were also member board members/RLAs	
Average Years of Experience	
Median Years of Experience	

	Number of Panelists	%
Panelists:		
Experience:		
3-6 Years		
7-11 Years		
>12 Years		
Mode of Practice:		
Private		
VA		
Education		
Sex:		
Male		
Female		
Grad Year:		
Earliest		
Most Recent		
Location:		
States Represented		
Training:		
Schools Represented		

Cut Scores by Panelist Demographics

Median Cut Score Recommendations by Panelist Background

	Number of Panelists	Median Cut Score Recommendation	
		# correct	% correct

Experience:

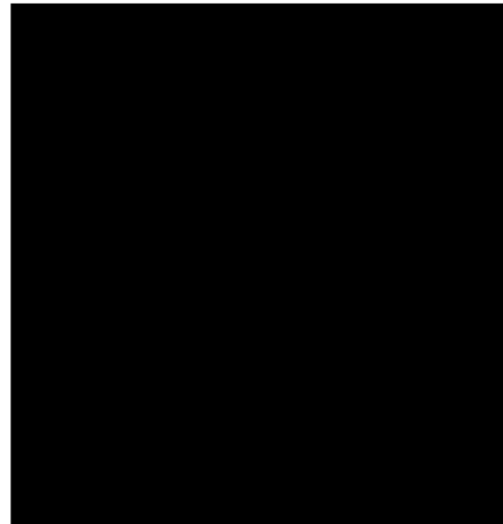
3-6 Years
7-11 Years
>12 Years

Mode of Practice:

Private
VA
Education
VA, Private

Sex:

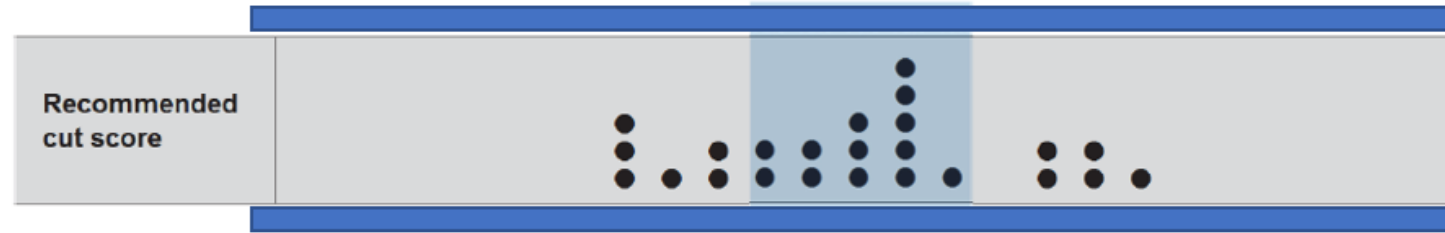
Male
Female



Panelist Recommendations, By Group

Distribution of Panelist Cut Score Recommendations

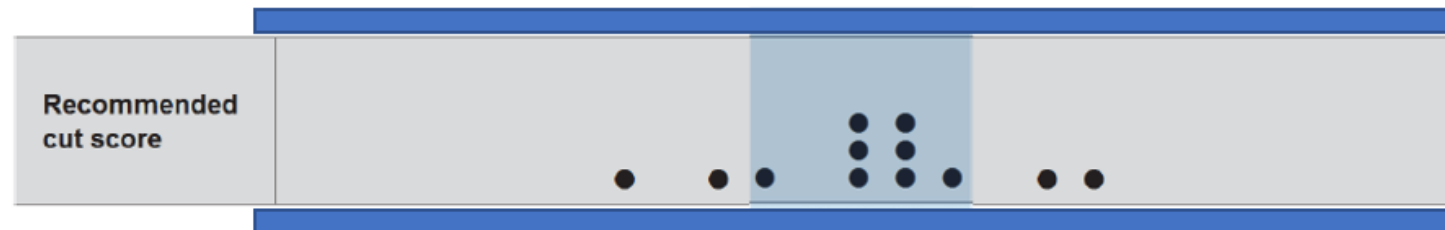
All panelists; rounded to nearest 5



Group 1; rounded to nearest 5

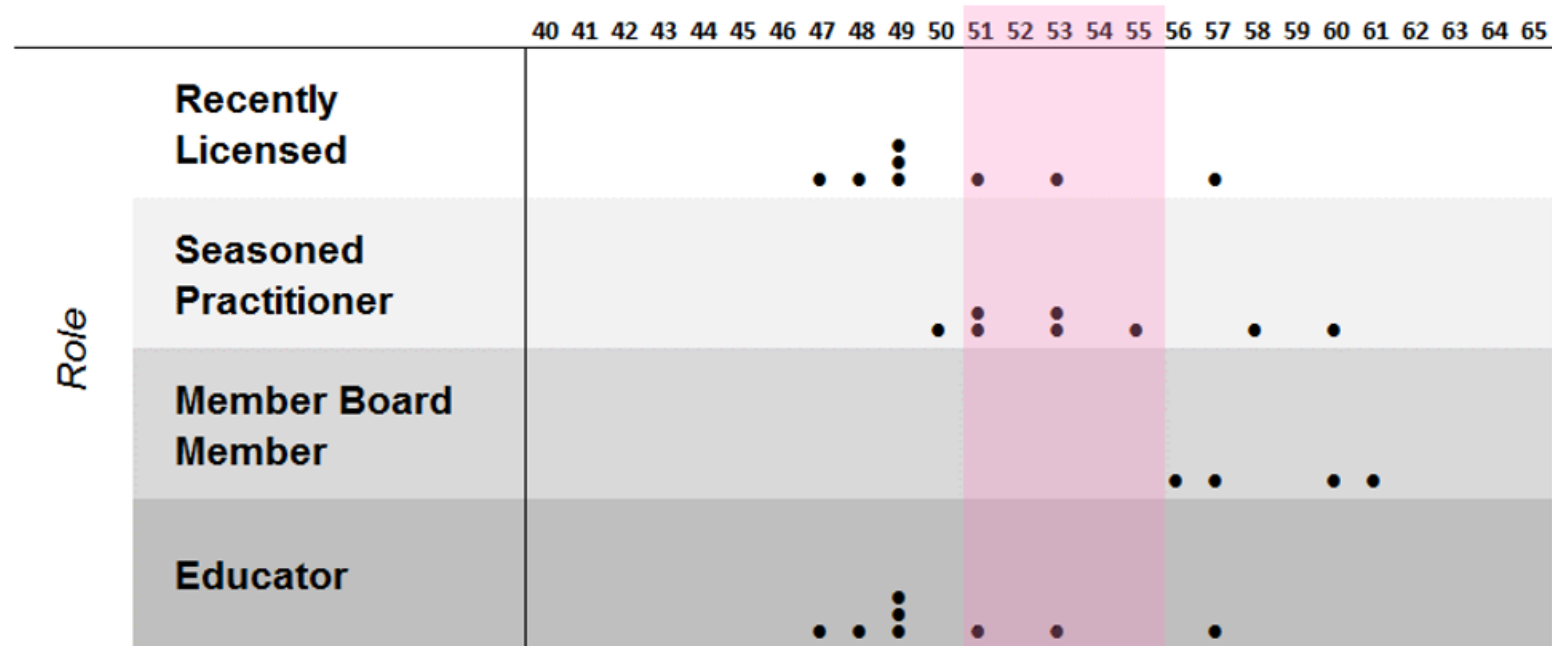


Group 2; rounded to nearest 5



Panel Recommendations

Distribution of Panelist Cut Score Recommendations



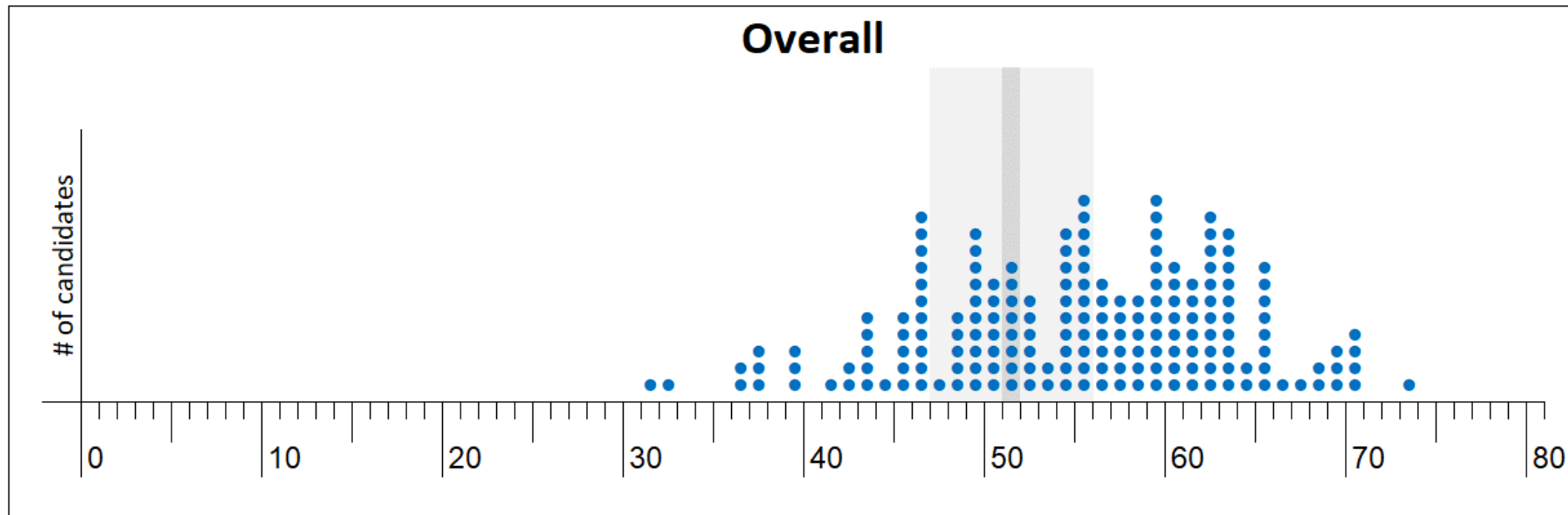
Recently Licensed = less than 5 years of licensure

Seasoned Practitioner = 5 or more years of licensure

Member Board Member = currently serves on an NCARB member board

Educator = teaches at an NAAB accredited school

Candidate Score Distribution



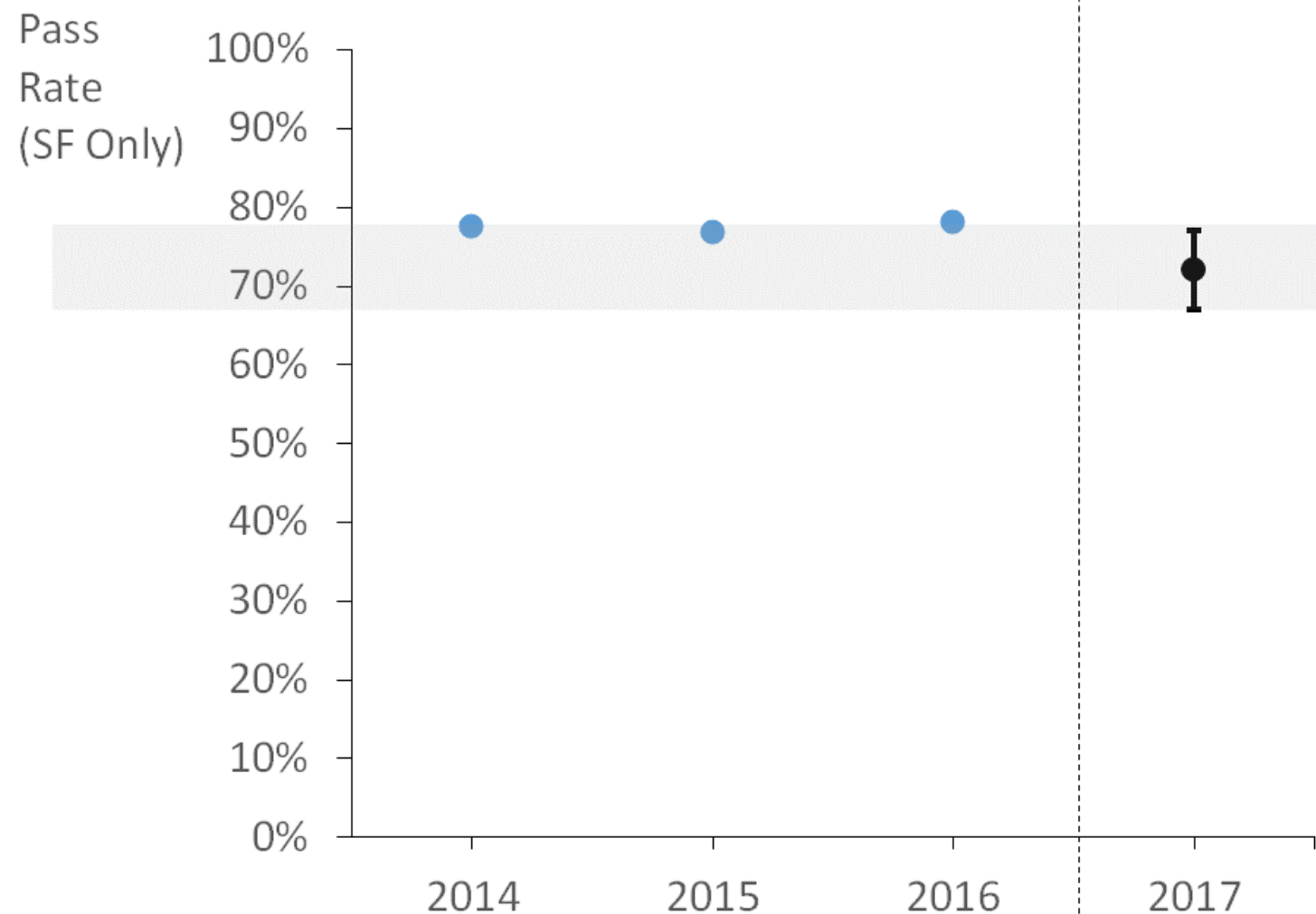
Impact Data

Standard Setting Results		Pass Rate SF only	Pass rate ALL candidates
Form A		Overall (N = 1645)	Overall (N = 2056)
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
Defensible Range	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%
	XXX	XX%	XX%

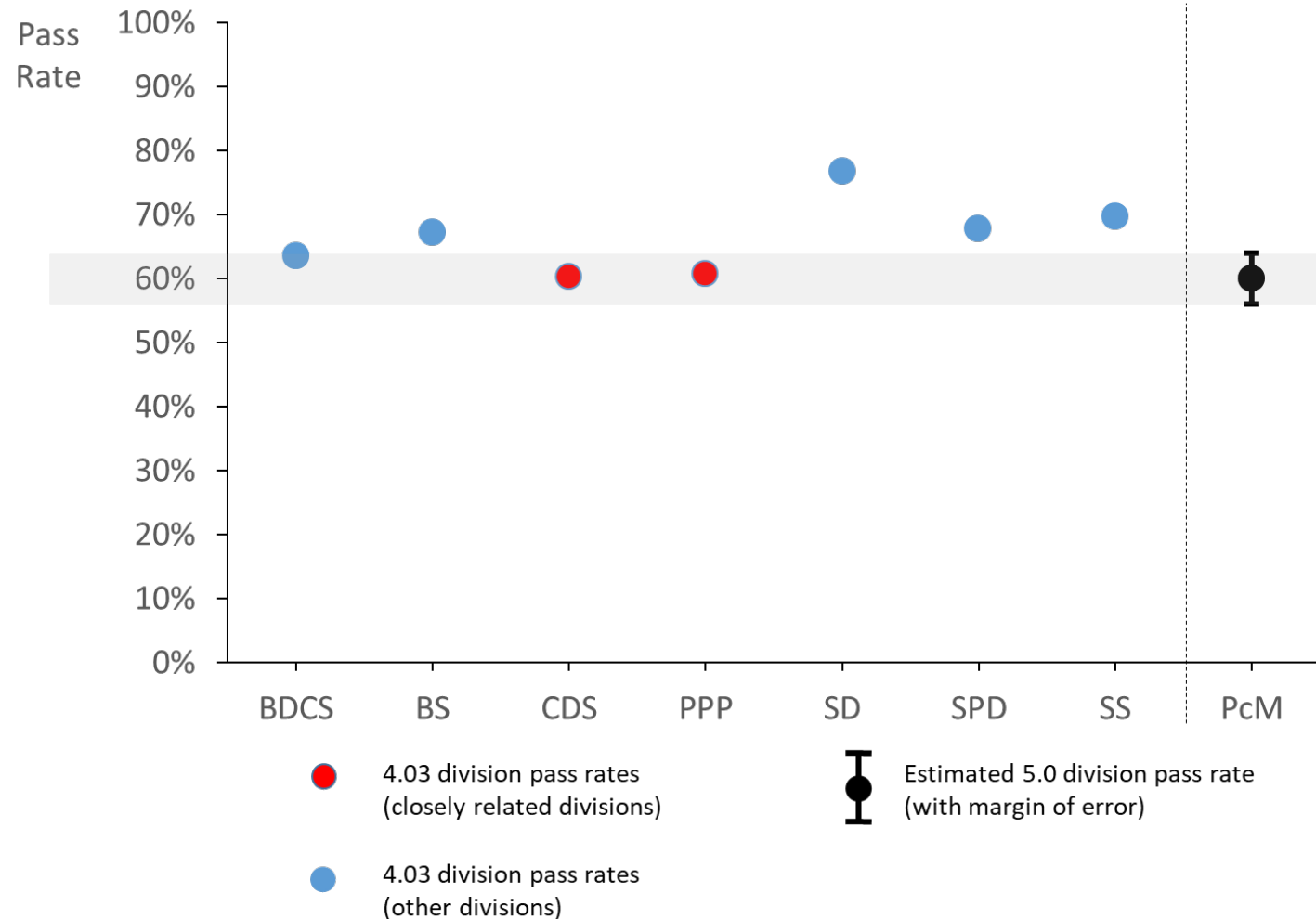
Impact Data with Recommended Ranges

		Pass Rates by Demographic Group						
		Cut Score	Demographic Group					
			All	Male	Female	White		Non-White
Exam Committee: Acceptable Cut Scores		XX	XX.X%	XX.X%	XX.X%	XX.X%	XX.X%	Exam Committee: Acceptable Pass Rates
		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%	
	Recommended Defensible Range	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%		

Historic Pass Rate Comparison



Comparison with Historic Pass Rates



Final Cut Score

Additional Steps

Examples of Additional Steps

- Calculate evaluative statistics based on cut score
- Review results regarding keeping or revising cut score
- Policy decisions regarding cut score revisions

You may not have thought
of...

Bonus ideas worth considering

- Readiness forms
- SME removal policy
- SME disqualification

Standard Setting Committee AGREEMENT TO PROCEED

Test: _____

This is to verify that I understand how to participate in the setting of cut scores using the standard setting method(s).

- I completed the training and participated in the practice cut score setting.
- Any questions I had about setting cut scores using the standard setting method(s) have been answered to my satisfaction.
- I understand the purpose of the committee.
- I understand what is meant by “minimally qualified candidate (MQC).”
- I understand how to make the cut score judgments.

By signing this form, I state that I am ready to proceed with the process of setting cut scores using the standard setting method(s).

Please print your name: _____

Signature: _____

Date: _____

If you are NOT ready to proceed, please print your name above but do NOT sign the form.
Please indicate below the aspects of setting cut scores by the standard setting method(s) that are not clear to you.

Concluding Remarks

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

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.

Questions?

Brett Foley, PhD, Alpine Testing Solutions, Inc.

brett.foley@alpinetesting.com

Corina Owens, PhD, Alpine Testing Solutions, Inc

corina.owens@alpinetesting.com

