Who Gets to Pass? Setting Defensible Passing Scores to Protect the Public

Pittsburgh, Pennsylvania

Presenters:
Thomas Langenfeld: ACT, Inc
Grady Barnhill: Commission on Dietetic Registration
Jeff Kenney: National Council for Interior Design Qualification

Promoting Regulatory Excellence

Goals of Today’s Program on Passing Scores:

- Through hands-on participation, we will learn the similarities and differences between three methods for determining a cut score.
- By participating, we will experience the cognitive requirements of each method.
- Have fun.

Problems Inherent in all Pass Score Studies

- Application of Human Judgment - all methods entail a subjective process
- Panel Effect - different panels may produce different outcomes
- Method Effect - different methods may produce different outcomes

◆ What are we to do?
Common Requirements of Each Procedure

- Common Starting Point: Judges must fully understand the purpose and meaning of the credential.
- The test content must be carefully tied to requirements and meaning of the credential.
- Concept of qualification (competence): Judges must understand the concept of minimum qualification (competence).

Group Participation

- Hands-on participatory program
  - Fictional Test
  - Fictional Credential
- We will apply three passing score methods
  - Modified Angoff Method
  - Bookmark Method
  - Hofstee Method

Fictional Test and Credential

- Purpose of Test - to certify state regulators in interpreting testing concepts to assist them in providing information to the public.
- Test: Interpreting Test Concepts
Modified Angoff Procedure & Bookmark Procedure

- Concept of Minimum Competency (aka: the Just Qualified Candidate, the Minimally Qualified Candidate, the Barely Proficient Candidate)
  - What knowledge would you expect?
  - What skills would you expect?
  - How does minimum competency relate to the task list?
- How good is good enough?

Working with Judges on Applying the concept of Minimum Competency

- Focus on the Just Qualified Candidate
  - Emphasize - this is the person who is just barely able to qualify.
  - The Just Qualified Candidate - what do you expect him or her to know? To be able to do?

Passing Standard (JTS 14.17)

- The level of performance required for passing a credentialing test should depend on the knowledge and skills necessary for acceptable performance in the occupation or profession and should not be adjusted to regulate the number or proportion of persons passing the test.
  
  Just say no to Normative!
Passing Point Considerations

- Restriction of Trade vs. Public Protection
- Public Protection vs. Access to Care
- Adequate/inadequate vs. good/superior

Angoff Item Ratings

- How many out of 100 Just Qualified Examinees would you expect to answer each question correctly?
  - Rate in increments of five.
  - Comment: Even on extremely easy items, some examinees will misread, mismark, or make some other mistake in answering.
  - Comment: Even on extremely difficult items, some examinees may guess the answer correctly.

Modified Angoff Process

1) Judges go through items, one-by-one, and rate them
2) Discussion
3) Round 2: p-values provided; judges rate again
4) Discussion
5) Finalize ratings
**Angoff Results**

<table>
<thead>
<tr>
<th>Items</th>
<th>Diff</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.70</td>
<td>60</td>
<td>45</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>66.43</td>
</tr>
<tr>
<td>2</td>
<td>0.47</td>
<td>50</td>
<td>60</td>
<td>70</td>
<td>70</td>
<td>60</td>
<td>65.96</td>
</tr>
<tr>
<td>3</td>
<td>0.63</td>
<td>75</td>
<td>55</td>
<td>60</td>
<td>60</td>
<td>50</td>
<td>66.07</td>
</tr>
<tr>
<td>4</td>
<td>0.56</td>
<td>70</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>50</td>
<td>67.86</td>
</tr>
<tr>
<td>5</td>
<td>0.42</td>
<td>45</td>
<td>55</td>
<td>50</td>
<td>65</td>
<td>40</td>
<td>65.36</td>
</tr>
<tr>
<td>6</td>
<td>0.77</td>
<td>80</td>
<td>75</td>
<td>65</td>
<td>70</td>
<td>70</td>
<td>73.00</td>
</tr>
<tr>
<td>Findings</td>
<td>63.33</td>
<td>60.83</td>
<td>65.00</td>
<td>68.33</td>
<td>57.50</td>
<td>53.00</td>
<td>41.11</td>
</tr>
</tbody>
</table>

**Bookmark Procedure**

- Book of Test Items: *Ordered Item Booklet (OIB)*
- In the OIB, items are arranged in order of difficulty from the least difficult to the most difficult. (Item Difficulty calibrated through an IRT model.)
- Because the Bookmark Procedure incorporates IRT, it possesses greater flexibility than other methods (i.e., it can be applied to a test using both multiple choice and constructed response items).
- Recommended passing score is on an IRT scale.

**Bookmark Directions for Judges**

- Begin by reading the first item in the OIB
  - Answer the question: **Is it likely that a just qualified candidate will correctly answer the item?**
  - 50% or 67% Likely
  - Judges answer “Yes” or “No”
- Proceed through the OIB answering “yes” or “no” for each item.
- When judges respond to an the item with “no,” **stop** and place the bookmark following the last “yes” item.
### Bookmark Results

#### Summary Data Table

<table>
<thead>
<tr>
<th>Raters</th>
<th>Bookmark Placement</th>
<th>IRT Theta Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>-0.84</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>-0.22</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>-0.84</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>0.05</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>-0.22</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td><strong>-0.41</strong></td>
</tr>
</tbody>
</table>

### Hofstee Method – Compromise Method

**Question #1**

Based on my review of this set of test questions, I think that the percentage of correctly answered questions required to achieve certification should not go below ___a___% correct, or above ___b___% correct. If I were to pick a single number as the requirement for passing, it would be ___c___% correct.

### Hofstee Method

**Question No. 2**

Based on my experience with this field of practice, I would not expect the overall failure rate for state regulators taking the Interpreting Test Concepts test to go below ___a___% fail, or be above ___b___% fail. If I were to select a single number as the expected failure rate for this group, it would be ___c___%.
Hofstee Procedure
Calculating the Pass Score

1. Calculate the average of each of the six responses.
2. Two critical points:
   - \( \text{Fail}_{\text{min}} \), \( \text{Knowledge}_{\text{max}} \)
   - \( \text{Fail}_{\text{max}} \), \( \text{Knowledge}_{\text{min}} \)
3. Plot points and draw line across distribution.

Hofstee Example

Findings from the three procedures

- Findings for our Test
  - Modified Angoff:
  - Bookmark:
  - Hofstee:
Assumptions of All Three Procedures Made Regarding the Judges

1. Raters know the examinee population
2. Raters know the test content
3. Raters know the qualifications of the Just Qualified Examinees

Similarities and Differences Between the Methods

- Cognitive Requirements of Modified Angoff
- Cognitive Requirements of Bookmark
- Cognitive Requirements of Hofstee

Utilizing the Findings

- Setting the Final Pass Score
- Incorporating Results from Multiple Methods which may not provide the same passing standard
- Considering the consequences of False Positives and False Negatives (Consequential Validity)
- Ultimate responsibility for the performance standard resides with the organization granting the credential
Who Gets to Pass? Setting Defensible Passing Scores to Protect the Public

Speaker Contact Information

Thomas Langenfeld
ACT, Inc.
Iowa City, IA 52243
Tel: (319) 337-1125
Email: thomas.langenfeld@act.org

Jeff Kenney
NCIDQ
1602 L Street NW, Suite 200
Washington, DC 20036
Tel: (202)721-0220
Email: jeff@ncidq.org

Grady Barnhill
Commission on Dietetic Registration
Chicago, IL 60606
Tel: (312) 899-4768
Email: gbarnhill@eatright.org