Setting Cut Scores for ACCUPLACER

Deanna L. Morgan, The College Board
ACCUPLACER Conference
June 24-26, 2010

DEMORGAN@COLLEGEBOARD.ORG
Goals of this Session

• Define standard setting
• Identify factors for consideration
• Setting cut scores for multiple choice tests
• Setting cut scores for free response tests
What do we mean by “Setting Cut Scores” or “Standard Setting”?

- Defining the place on the score scale which represents sufficient knowledge and skills for placement into an upper level course.

<table>
<thead>
<tr>
<th>Lower-Level Developmental Course Placement</th>
<th>Upper-Level Developmental Course Placement</th>
<th>College-Level Course Placement</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>Test Scale</td>
<td>120</td>
</tr>
<tr>
<td>Upper-Level Developmental Cut Score</td>
<td></td>
<td>College-Level Cut Score</td>
</tr>
</tbody>
</table>
Policy Considerations

• A series of decisions must be made prior to a standard setting.

• To the extent possible decisions should be made and documented prior to the start of the standard setting to prevent the appearance of bias in the process.

• For example, specifying how data will be gathered and cut scores computed ahead of time will minimize complaints that decisions about data, such as rounding or dropping the most extreme scores, were made in reaction to the outcome rather than in a thoughtful and planned manner.
Facilitator Selection

- The Facilitator has the responsibility for leading and overseeing the standard setting process.

- Should be trained and experienced in standard setting, should be external to the institution to the extent possible, should remain impartial and direct the process without interjecting opinion as to the “correct placement” for the cut score.
Criteria for Selecting Panelists

- Representative distribution of demographic characteristics, e.g., gender, race/ethnicity, geographic location in system (if multi-site campus)
- Recent experience teaching in the subject matter area and at the course level or levels into which the cut scores will be used to place students
- Distribution of experience level from beginning instructor to veteran.
Performance Level Descriptors

- Describe what students at the borderline of each course level should reasonably know and be able to do
- Relate directly to the course requirements/prerequisites
- Distinguish clearly from one level (developmental course) to the next (college-level course)
How to Select a Method

- No one best method
- Using different methods will most likely result in different cut scores
- Some methods are better suited to certain types of assessments and decisions than others
- Consider precedent for legal defensibility of the method and your risk of legal challenge
Modified Bookmark Approach

• A very popular method that has been used frequently for setting cut scores in various educational settings around the country. Well vetted with measurement professionals and have withstood challenges in court.

• Requires access to test items and item difficulty statistics. This information is available from the College Board by contacting your HEAM (SAM) representative or going online to:

• Complete Appendix A of the “Guidelines for the Release of Data” form at the above link.

• Please allow at least 6 weeks from when you need the data to get all paperwork completed, including signatures, and for the data to be prepared and shipped to you.
Bookmark Methodology

• Test items should be ordered by statistical difficulty value (IRT b-value) from easiest to most difficult and a complete ordered set provided to each panelist.

• Using the ordered item set and the performance level descriptors, panelists should work individually to identify the location in the item set that they feel best separates groups of examinees into categories and then literally place a bookmark at that location in the ordered item set.

• Before beginning the task of placing the bookmark panelists should develop the performance level descriptors that will guide the process, take the exam without benefit of having the correct answers in front of them, and undergo training on the process.

• This placement task takes place over three rounds of discussion and deliberation.
Bookmark Placement

What students should know and be able to do
Bookmark Placement and Discussion

• At least two, preferably three, rounds of judgment and discussion should occur during the process.

• Discussion of the items and any particularly tricky or notable characteristics of the items should occur following panelists taking the test. This discussion should focus on item content and skills needed and not on where the cut score should be placed.

• Placement of bookmarks should be the individual panelist’s decision and not a group decision. Only the discussions that occur between placement rounds are group activities.
Considerations for the Bookmark Method

- Multiple cut scores may be set at one time
- Feedback from panelists is generally positive
- Panelists can consider distance between cut scores
- Requires analyzed student data and the actual test items to proceed
Modified Body of Work Method

- Requires access to samples of student work – in this case essays at each score point. This information is available from the College Board and included in the document provided at the ACCUPLACER resource tab for the convenience of institutions who need to set cut scores prior to collecting data on their students.

- It is preferable to use essay samples collected from your own students when possible. At least 5 essays at every score point and up to 15 per score point are needed. It is especially important to have “extra” essays for score points across the middle of the score scale. Essays should be in the original written state and not corrected in any way for grammar, spelling, formatting, etc.

- This method is generally well received by faculty as it closely mirrors work they are familiar with – evaluating student work.
Essay Sorting and Discussion

• Three rounds of judgment and discussion should occur during the process.

• If possible, have panelists complete an essay on the prompt being used under typical student administration conditions at the beginning of the session or prior to attending the session.

• Discussion of the prompt, scoring rubric, and essays should occur following panelists taking the test. This discussion should focus on skills needed to be successful in the courses and not on the score point where the cut score should be placed.

• Placement of essays into performance categories should be the individual panelist’s decision and not a group decision. Only the discussions that occur between placement rounds are group activities.
Round 1 – Rangefinding

English 098
- Essay #3
- Essay #8
- Essay #9

English 099
- Essay #1
- Essay #5

English 101
- Essay #4
- Essay #7
- Essay #6
- Essay #2
## Feedback Example for Multiple Cut Scores

<table>
<thead>
<tr>
<th>Essay ID Number</th>
<th>Lower Level Developmental Course</th>
<th>Upper Level Developmental Course</th>
<th>College-Level Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>20</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td></td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>28</td>
<td></td>
<td>2</td>
<td>23</td>
</tr>
</tbody>
</table>
Round 3 – Pinpointing

• New essay sets are distributed to panelists at the start of Round 3. The essay sets have been selected to only represent score points that fall into the determined range as identified following Round 2 Rangefinding and confirmed by panelists in the Round 2 discussion.

• Panelists may request that the range be extended if they feel it is too narrow.

• Using the identified range new sets of essays with 2 to 5 essays per score point within the range are pulled and used in Round 3. Therefore, it is important to have on hand additional essays beyond the Rangefinding sets that can be used to form the Pinpointing sets.
Round 3 – Pinpointing (cont.)

- During Pinpointing the sorting task is the same as that used for Rangefinding but the range of score points in the set is smaller.

- For multiple cut scores separate Pinpointing sets should be used for each cut score unless the cut scores overlap in such a way that it is difficult to separate into unique sets of papers.

- Round 3 Pinpointing is the last round of sorting and the final cut score recommendation will be made based on that data and following the same procedure used to find the median and define the range for Pinpointing following Round 2.
Evaluation and Documentation

• During the process keep a copy of all materials that are distributed or generated during the standard setting meeting for reference if any challenge should be made about the cut score placement.

• Panelists should complete evaluation forms throughout the process to inform the facilitator of their level of understanding and to provide documentation of any concerns that may have been present during the time of the meeting.

• A decision should be made prior to the meeting whether results of the last round will be shared with panelists.
The Final Decision on Cut Scores

- The Authoritative Body (one person or more than one with the authority to make the final decision) should have all the information they want or need to make an informed decision.

- At a minimum the following should be provided:
  - Overall results
  - Variance in judgments and in measurement
  - Impact on subpopulations in the system
  - Documentation of the process (with as much detail as is necessary)
ACES Validity Handbook

- Online reference guide that answers the basic question, “What is validity?” and highlights why validity is an important issue for colleges and universities as well as assists with the planning and execution of validity studies.

- It is designed to serve as a general reference for validity and includes information about validity beyond what is specifically applicable to ACES.

- It includes specific information about the types of validity studies and their design that are available through the ACES system and helps interpret ACES study results.

- To request an ACES study: http://www.collegeboard.com/highered/apr/aces/aces.html
Final Reminder

- Facilitators do not set cut scores, they simply guide the process and provide the authoritative body with information on the psychometric issues surrounding cut score placements under consideration.

- Panelists do not set cut scores, they recommend cut score locations

- Cut scores are approved and set by the authoritative body at the institution which makes the decision to accept the panel recommendations as they are or to adjust the recommendations for policy reasons. The authoritative body sets the final cut score location(s).
For Additional Information on Setting Cut Scores


• For more information on Standard Setting see Cizek (2001).

• ACCUPLACER Online Resources Tab provides access to documents on setting cut scores for ACCUPLACER.