Colorado Measures of Academic Success
Science and Social Studies High School

Next Step – Setting Cutscores

February 2015

Joyce Zurkowski, Executive Director of Assessment
Jennifer Beimers, Psychometrician, Pearson
Overview

- Item and Test Development Process
- Process for Determining Cutscore Recommendations
New Content Standards

New Assessments

New Performance Levels (cutscores)
Why Do We Set Cutscores?

To support the interpretation of student results
Item and Test Development Process
Item and Test Development Process

- **Colorado Educator Item Development**
- **Pearson Item Development**
- **CDE & Pearson Fact Checking & Revising**
- **CDE Verifying & Revising**
- **Colorado Educator Item Review (Content and Bias/Sensitivity)**
- **CDE & Pearson Editing**
Item and Test Development Process

- Field Test
- Colorado Educator Rangefinding
- Item Scoring
- Colorado Educator Data Review
- First Operational Assessment
Item and Test Development Process

- Item Scoring
- Colorado Educator Performance Level Setting
- Performance Level Descriptors & Cut Scores Adoption
- Scoring and Reporting
- Local Distribution of Results
Scores

- PERFORMANCE LEVELS
  - Overall scale score
  - Standard scale scores
  - Selected response and constructed response scale scores
  - Prepared graduate competency (PGC) and grade level expectation (GLE) percent correct
Performance Level Labels

- Distinguished Command
- Strong Command
- Moderate Command
- Limited Command

Performance level descriptors are finalized during this process.
<table>
<thead>
<tr>
<th>Performance Levels and Policy Claims</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distinguished Command</strong></td>
</tr>
<tr>
<td>Students performing at this level demonstrate a distinguished command of the concepts, skills, and practices embodied by the Colorado Academic Standards for Science or Social Studies assessed at their grade level. They are academically well prepared to engage successfully in further studies in this content area.</td>
</tr>
<tr>
<td><strong>Strong Command</strong></td>
</tr>
<tr>
<td>Students performing at this level demonstrate a strong command of the concepts, skills, and practices embodied by the Colorado Academic Standards for Science or Social Studies assessed at their grade level. They are academically prepared to engage successfully in further studies in this content area.</td>
</tr>
<tr>
<td><strong>Moderate Command</strong></td>
</tr>
<tr>
<td>Students performing at this level demonstrate a moderate command of the concepts, skills, and practices embodied by the Colorado Academic Standards for Science or Social Studies assessed at their grade level. They will likely need academic support to engage successfully in further studies in this content area.</td>
</tr>
<tr>
<td><strong>Limited Command</strong></td>
</tr>
<tr>
<td>Students performing at this level demonstrate a limited command of the concepts, skills, and practices embodied by the Colorado Academic Standards for Science or Social Studies assessed at their grade level. They will likely need extensive academic support to engage successfully in further studies in this content area.</td>
</tr>
</tbody>
</table>
High School Performance Level Descriptors (PLDs)

- Drafted by CDE and Pearson Content Specialists
- Posted and disseminated for public comment from January 9th through January 28th
- Finalized during the standard setting process
Who will Adopt the New Performance Levels?

- The State Board of Education
  - Relying on the recommendations of the educator panels
The approximately 25 educator panelists are:

- Experts in the concepts and skills reflected in the Colorado Academic Standards (CAS) and in student development of the grade-specific concept and skills
- Knowledgeable about the range of characteristics in the population tested
- Interested in the results of the cutscore setting process and consequences associated with the results

Selected through open recruitment process by Pearson and CDE.
Who will Facilitate the Meetings?

- **Pearson and CDE assessment staff:**
  - Psychometricians lead the meeting
  - Data analysts collect ratings, run analyses and generate feedback reports
  - Content experts answer content-related questions panelists may have
What Method will be Used?

- The Bookmark Method: one of the most widely used content-based procedures to set cutscores for large-scale state assessments
  - Goal of the method: Identify the cutscore associated with the “just barely” threshold level of a given performance level
    - Four performance levels
      - Limited Command
      - Moderate Command
      - Strong Command
      - Distinguished Command
  - Uses an online Ordered Item Booklet (OIB)
  - Focus on threshold students: assumes students who have been instructed in the Colorado Academic Standards, using locally determined methods and curriculum

On track toward college and career readiness
Ordered Item Booklet

- Represents the range of concepts and skills assessed
- Consists of:
  - Operational items and embedded field test items
  - Various item types
- Tool by which cutscore recommendations are made

Order is based on 3-parameter model.
Focus on Threshold Students

- Limited Command
- Moderate Command
- Strong Command
- Distinguished Command

Threshold Students

Recommended Cutscores

- Low Achievement
- High Achievement
What is the Process for Setting Cutscores?

**Day 1:**
- General Session
- Two Content Specific Panels
  - Introductions
  - Review of Performance Level Descriptors (See Handouts)
  - Develop descriptions of threshold students
  - Review items
  - Training in Bookmarking process, including practice
  - Complete readiness survey
  - Round 1 ratings
What is the Process for Setting Cutscores?

- **Day 2:**
  - Discuss Round 1 recommendations using:
    - Feedback report produced overnight
    - Item difficulty report (p-values)
  - Round 2 ratings
  - Complete readiness survey
  - Review Round 2 ratings using:
    - Impact and external data
  - Round 3 ratings
  - Review final recommendations
  - Complete questionnaire
November: Administration

December: Materials processed
          Responses scored

January: Student demographic data review
        Panelist selection
        Creation of cutscore meeting materials

February: Cutscore panels convene

March: Recommendations brought to the Board

April/May: Scores and reports in districts
How will Participation Rates be taken into Consideration?

- Total number of eligible 12th graders: 56,000
- Number of 12th graders with scores: 45,500
- Sample: 38,000

- The resulting sample is within acceptable range on key demographic variables
Recommendations will be Brought to the Board in March

The Board will be asked to adopt:

• Descriptors for the performance levels
• Cutscores