Overview

Professional practices and measures of student learning are equally represented in determining a principal's final effectiveness rating in the Colorado State Model Evaluation System. Each of the two components represents 50 percent of an educator's final evaluation rating.

How Scoring Works

The State Model Evaluation System determines final effectiveness ratings by adding professional practice scores with measures of student learning scores. The process of combining measures starts by adding together the scores for each standard on the rubric, generating the professional practices score. Then, the score for measures of student learning is determined using multiple measures. Those final scores are added together to create a single effectiveness, or index, score, which then falls into the cut points of one of the following categories: ineffective, partially effective, effective, and highly effective.

Scoring Changes

The way in which scores and ratings are calculated has shifted slightly following feedback from educators, school leaders, and district personnel who felt the former scoring system failed to align with their professional values and impeded coaching conversations and to align the scoring process to the teacher and SSP systems. The following sections detail the scoring process.

Note that no changes were made to the progressive scoring at the professional practice level, meaning that an educator must still get all of the professional practices found in Level 1 (formerly known as Basic) and all of the professional practices found in Level 2 (formerly known as Partially Proficient) in order to receive a rating of Level 2 Practices in a specific element.

Professional Practices Scoring

Feedback from focus groups and district leaders indicated a preference for requiring a preponderance of evidence in order to earn a rating at the standard level and for establishing rigorous expectations for achieving the highest rating. The following table details the changes made in relation to this feedback.
Scoring Shifts in the State Model Evaluation System

<table>
<thead>
<tr>
<th>Scoring at the Standard Level</th>
<th>Former Scoring System</th>
<th>Revised Scoring System</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scoring at the Standard Level</td>
<td>Half of the elements for any given standard must be scored at the higher of two ratings in order to earn the higher rating.</td>
<td>There must be a “preponderance of evidence” in order for an educator to receive the higher of two ratings within a standard. If there are 6 elements in a standard, then the educator must receive 4 out 6 elements at the higher rating in order to earn the higher rating.</td>
<td>In the former system, scores did not necessarily provide an accurate representation of an educator’s ability because educators only needed 50 percent (not a majority) of elements rated at the higher rating to earn the higher rating for the standard. Thus, the revised system requires a majority of elements be at the higher rating to earn the higher rating for the standard.</td>
</tr>
<tr>
<td>Overall Professional Practices Score</td>
<td>The cut points were set based on the educator earning at least the minimum score for a rating in order to earn the rating.</td>
<td>The cut points were set based on the educator earning the midpoint or higher of the two consecutive ratings in each standard.*</td>
<td>The former scoring system did not accurately define proficient (according to stakeholder values) because an educator could score partially proficient on a majority of standards and still receive a score high enough to gain a proficient rating. The revised scoring system responds to this by making requirements to earn a proficient score more rigorous to better reflect stakeholders’ idea of proficient.</td>
</tr>
</tbody>
</table>

*The rating ranges were established based on the standards being weighted equally. LEAs may choose their own custom weights in order to emphasize specific standards based on local values and context.

Changes in Rating Scale for Professional Practice

The following tables demonstrate how the aforementioned changes have altered the cut scores for professional practice between the former and revised scales for principals.

Former Scale: 

Revised Scale:
Measures of Student Learning Scoring

No changes have been made to the MSL scoring system. The table below details the measures of student learning ratings for both the former and revised scoring systems.

<table>
<thead>
<tr>
<th>Measures of Student Learning Rating</th>
<th>0-3 Point Range per Rating (Rounded to the nearest hundredth)</th>
<th>Scale Converted to 0-540 (Rounded to the nearest whole number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Much Less Than Expected</td>
<td>0 - .49</td>
<td>0 - 134</td>
</tr>
<tr>
<td>Less Than Expected</td>
<td>.50 - 1.49</td>
<td>135 – 269</td>
</tr>
<tr>
<td>Expected</td>
<td>1.50 - 2.49</td>
<td>270 – 404</td>
</tr>
<tr>
<td>More Than Expected</td>
<td>2.50 - 3.00</td>
<td>405 – 540</td>
</tr>
</tbody>
</table>

Final Effectiveness Score and Rating

The changes to the cut points for the Professional Practice score affected the cut points for the combined Final Effectiveness Rating. A change was also made in determining the minimum cut point for Highly Effective. Previously, the minimum cut point for Highly Effective was determined by adding the minimum score for Accomplished to the minimum score for More than Expected. It is now determined by adding the midpoint score for Accomplished to the minimum score for More than Expected.

The following table shows the changes in cut scores for the final effectiveness score and rating for Principals.
The figures below provide a visual comparison of the former and revised systems for determining a Highly Effective rating.

**Former System**

![Table showing the former system for determining a Highly Effective rating]

**Revised System**

![Table showing the revised system for determining a Highly Effective rating]

**Justification**

In the former system, the Highly Effective category had the largest possible range of scores. Additionally, school and district leaders found that it was more difficult to have effective feedback conversations around growth with teachers rated Highly Effective. Thus, the revised system sets a higher expectation for earning a Highly Effective rating in order to make the rating more accurate and to facilitate educator growth.