What is standard setting?

It is commonly accepted that standards are an important element in providing a sound education for America’s students. They strive to clarify learning expectations and to ensure that students receive the kinds of quality instruction that leads to achievement of the standards. Yet, if ever terminology caused confusion, “standard setting” is probably it. Some of the best information on this topic has been produced by the U.S. Department of Education (USDE) and the Council of Chief School Officers (CCSSO) through its state collaborative on Comprehensive Assessment Systems for IASA Title I (Handbook for the Development of Performance Standards: Meeting the Requirements of Title I, 1998). This document particularly targets the issues involved in setting standards and developing assessments for all students, including students with disabilities and students with limited English proficiency (LEP).

While the development of content standards, performance standards, and standard setting is an iterative process, the development of content standards is often a first step in the process. These standards set forth expectations for what students should know and be able to do. In most jurisdictions, these standards are delineated by content areas and grade levels (e.g., primary language arts; 5th grade mathematics; high school chemistry). There are a number of approaches to establishing such standards, but standard setting usually revolves around reaching consensus among content experts, curriculum experts, classroom teachers, policymakers, and representatives of the public about what is important for children to learn in schools.

Performance standards have most often been described as statements about how good is good enough with respect to any of the content standards. Often this has been interpreted to mean a “cut score,” e.g., the score a student must obtain to be “proficient,” or the score a student must obtain to be “advanced.”

What are performance standards?

Actually, performance standards are not cut scores at all, or at least not just cut scores. Performance standards are a system consisting of several highly interrelated components that include performance levels, performance descriptors, exemplars of student work, and cut scores.

The performance levels label each level of achievement. Performance descriptors describe each level of performance. Exemplars are samples of student work at each level of performance. Cut scores separate the different levels of performance.

Methods of developing each of these components may vary from state to state and assessment to assessment. The information provided by these performance standards helps to guide the design, interpretation and use of assessments and the results they produce.

While performance levels and performance descriptors are often determined before an assessment system is designed, such is not always the case. In fact, when various aspects of a performance standard are developed depends on why they are developed and the kind of system for which they are developed. Because the process is iterative, it may be desirable to draft descriptors and then revise them based on the results of field-testing. However, when performance levels and descriptors are the first steps in development, they can help clarify the design of the system and enhance, in advance, alignment with the content standards.
Exemplars of student work from the classroom or from “formal” assessments may be used in developing these performance levels and performance descriptors. Exemplars will also be useful in training those who actually score student work on the tests.

**PERFORMANCE STANDARDS**

- **Performance levels** ➔ **Labels each level of achievement**
- **Performance descriptors** ➔ **Describes each level of performance**
- **Exemplars** ➔ **Sample student work at each level of performance**
- **Cut Scores different** ➔ **Scores that separate the levels of performance**

Cut scores are usually set after the test has been administered for the first time. This procedure can involve some important technical issues. The guidance of test developers and assessment experts is important to ensuring technical quality in this process.

- Performance levels are labels for each level of achievement. The number of performance levels a state uses can vary, but, typically, states have from three to six levels. For example, the performance labels may include:
  - Advanced or exemplary for a performance that exceeds the required level of proficiency.
  - Proficient or mastery for a performance that is at the required level.
  - Partially proficient, needs improvement, or partial mastery for a performance that is below the required level.
  - Non-proficient, novice, or non-mastery for a performance that indicates little or no knowledge of the standard.

- Performance descriptors are narrative descriptions of performance at each level. For example, at the proficient level, the performance descriptor may be:
  - A description of proficient performance on the fifth-grade writing assessment.
PERFORMANCE LEVELS

- Advanced or Exemplary
- Proficient or Mastery
- Partially proficient or Needs Improvement
- Novice or Non-proficient

PERFORMANCE LEVEL RESULTS for READING

<table>
<thead>
<tr>
<th>Grade</th>
<th>Novice</th>
<th>Needs Improvement</th>
<th>Proficient</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>4th</td>
<td>11</td>
<td>31</td>
<td>63</td>
<td>31</td>
</tr>
<tr>
<td>8th</td>
<td>50</td>
<td>5</td>
<td>63</td>
<td>50</td>
</tr>
<tr>
<td>10th</td>
<td>10</td>
<td>5</td>
<td>60</td>
<td>5</td>
</tr>
</tbody>
</table>

- A description of proficient performance on the fifth-grade reading assessment.

- A description of proficient performance in writing overall.

- Exemplars are examples of student work from a representative sample of all students that illustrate the full range of performance at each level. For example, samples may include:
  - Needs improvement, proficient, and advanced examples of writing from the fifth-grade level, covering the different types of writing in the content standards.
  - Needs improvement, proficient, and advanced samples of mathematics problem solutions from the fifth-grade level, covering a variety of types of problems, processes, and mathematics areas.

- Cut scores are scores on assessments that separate the different levels of performance. For example, a cut score might be:
  - An obtained score of at least 60% correct to be classified as proficient.
  - A scaled score of 270 on a possible scale of 200 to 280 to be classified as proficient.

As with content standards, there are a number of approaches to setting performance standards. In general, setting performance standards requires experts in the field (e.g., educators, policymakers, content experts) to make judgments about student performance at each of the performance levels. Procedures that involve test developers and teachers in setting performance standards are most defensible because they help assure sound relationships between and among the content standards, what is taught, what is tested, and what students know and can do.

For more information about content and performance standards, see:


---

**PERFORMANCE STANDARDS**

**Cut Scores**

- **Percent correct:**
  - 0% 30% 60% 85% 100%
  - Novice Partially Proficient Proficient Advanced

- **Scaled scores on scale of 200 to 280:**
  - 200 230 250 270 280
  - Novice Partially Proficient Proficient Advanced

Points of view or opinions expressed in the paper are not necessarily those of the U.S. Department of Education, or Offices within it.