



New Meridian

Frequently Asked Questions about Performance Levels

December 2025

Each Testlet Report provides a snapshot of how a student performed on a short, focused assessment covering a specific set of skills. Because each testlet includes only a small number of items, the results are designed to support instructional understanding, not to make a final determination of overall grade-level achievement.

The performance levels on the Testlet Report are based on an estimate of the student's overall ability, calculated from their responses on that testlet and placed on the same statewide scale used for their summative score on the through-year report. That estimated scale score is then compared to the statewide proficiency cut score to determine whether the student's performance on that testlet is consistent with, approaching, or below grade-level expectations.

There is natural variation inherent in every assessment, and each testlet includes a relatively small number of items. To account for this, we create a range around the summative proficiency cut score to create the performance levels. A student's designated level depends on whether their estimated summative scale is below, within, or above that range. This ensures that students whose performance is close to the cut score are not over-classified into a higher level. In other words, the level boundaries are intentionally conservative and stable, and they do not change from one testlet to another, even though the content and difficulty of testlets may vary.

The performance levels are:

- Level 3: A student score in this level shows performance consistent with grade-level expectations for proficiency.
- Level 2: A student score in this level shows performance approaching proficiency, but not yet fully meeting grade-level expectations.
- Level 1: A student score in this level shows performance below grade-level expectations for proficiency.

Because each testlet measures a different set of standards and skills, and because the difficulty of the items varies across testlets, a student may show different performance levels on different testlets, even if the number of correct responses is the same. On a more difficult testlet, a level 3 performance may require fewer correct answers, while on an easier testlet it may require more. This difference reflects the difficulty of the content being assessed, not necessarily a change in the student's underlying ability.

Testlet Reports in the 24–25 school year used Performance Bands, which compared a student's score to the scores of other Montana students who took a similar testlet. In that model, students were grouped into thirds:

- Band 3: Highest-performing third of students statewide who took that testlet
- Band 2: Middle-performing third of students statewide who took that testlet
- Band 1: Lowest-performing third of students statewide who took that testlet

This approach was used because the statewide proficiency cut scores had not yet been finalized. The state's standard-setting process, which establishes the summative proficiency expectations, was completed in summer 2025. Until those proficiency standards were set, the Testlet Reports communicated performance by comparing students to peers, not to the statewide proficiency benchmark.

Now that the statewide proficiency standard is established, Testlet Reports use Performance Levels, which compare a student's performance directly to the grade-level expectation, rather than to other students. The new levels answer a different and more instructionally useful question: Does the evidence from this testlet suggest that the student is currently performing at, near, or below the level expected for grade-level proficiency?

Common Questions

Why does the distribution of my students look different this year compared to last year?

Last year's Testlet Reports used Performance Bands, which compared each student's performance to that of other students taking a similar testlet. This created a distribution that was always split into thirds statewide.

This year, Testlet Reports use Performance Levels, which compare student performance to the statewide proficiency expectation, not to other students. Because the levels are tied to grade-level standards rather than peer comparison, the distribution of students may look different from last year. This does not indicate that instruction, student performance, or school outcomes changed. It reflects a different, more useful method of interpreting the assessment results.

Why do I have so many students in Level 2?

Level 2 indicates that a student's performance on this testlet is close to the proficiency expectation but not yet demonstrating the level of understanding needed to be classified as proficient based with certainty on this limited evidence. Level 2 represents a range of performance around the proficiency cut score, not performance just below it. Students in Level 2 may ultimately be slightly below, at, or even slightly above the summative proficiency threshold based on the evidence from one testlet. Because each testlet contains a small number of items, we use a conservative range around the proficiency cut score to avoid over-classifying students as Level 3 when there is not yet strong enough evidence. This means it is not unusual to see many students in Level 2.

Why are some student performances categorized in Level 2 if they only missed one question?

Performance levels are not based solely on percent correct or the number of questions answered correctly. They are also based on the difficulty of the questions being asked and the number of questions. Because testlets vary in difficulty, a student may reach

Level 3 with fewer correct answers on a difficult test, while on an easier testlet, it may take more correct answers to indicate a performance consistent with proficiency requirements.