



1. How are Interim and Modular scaled scores calculated?

There are two kinds of scores that can be seen in the AIR Ways, item scores and scale scores. The item score indicates how many points a student earned for his or her response. For example, a student may earn 1 point for a multiple-choice item or up to 4 points for an extended response item. This score can be found when looking at student responses, but currently there is not a summary report for these raw score totals in AIR Ways.

Scale scores are reported at the overall and subcategory levels in AIR Ways. The scale score factors in the difficulty of each item along with the student's response. Reporting in this manner is necessary for an adaptive test since students are tested on the range of depth of content, but at different difficulty levels. This difference in difficulty level is why one student may have a higher total item score, or raw score, but a lower scaled score than another student on the same test. Each of the correctly answered questions of greater difficulty contribute more to that student's scaled score. The proficiency determination is determined from the scaled score.

2. What do the Performance levels represent in AIR Ways?

The Performance levels aid in interpreting how well a student did on a Modular or Interim assessment and provide an indication of probable performance on the Summative assessment later in the year, if no interventions or additional learning takes place in between. If a student achieves a "Low" performance level, then that student is not likely to be proficient on the Summative assessment, assuming no further learning occurs before the spring summative testing window. A student achieving an "At or Approaching" performance level indicates the student is more likely to be proficient on the Summative assessment, with further instruction. If a student achieves at the "On or Above" level, the student is most likely to be proficient or advanced on the Summative assessment.

3. What is the range of scale scores for each Performance Level?

WY-TOPP proficiency cut scores will be set during standard setting in the summer of 2018. Meanwhile, WDE and AIR decided to use a temporary proficiency cut so that stakeholders can infer performance of each student for this first year of Interims and Modulares. When using a temporary proficiency cut like this, it is better to use a confidence interval to classify students' performance level. Usually, standard error of measurement (SEM) from each student is used to create the confidence interval around the temporary proficiency cut score. This methodology is used widely in the large-scale assessment field and was also approved by Wyoming Technical Advisory Committee (TAC) members.

4. Why are some students' performance levels not categorized similarly to their peers?

What is important to understand with the methodology of using confidence intervals based on SEM to classify students' performance level is that SEMs can be very high at the upper and lower end of the scale, depending on the items that were administered to a student and their responses to those items. Therefore in a few instances, confidence intervals can be very large because of the student's high SEM. In these cases the student is assigned the middle performance level, "At or Approaching."

While this is a mathematically correct approach, there will be a few students who have a low or high scaled score and still classified as “At or Approaching” due to their large SEM. This is also why there are not defined ranges of scaled scores for each performance level for these tests at this time. Please note that this is temporary. After standard setting in summer of 2018, all 2018-19 Interims will be reported using the same scale and performance cut scores as the Summative assessments.

Please note, because Modular assessments measure a small portion of the standard within the test and do not align with the Summative blueprints, the same cut scores will not apply. Rather, Modulares will be reported with the cut scores for Proficient used in conjunction with the student score and SEM to estimate their performance below, at/near or above standard.

5. How will the range of scale scores for each Performance Level change after Standard Setting in the summer of 2018?

During the standard setting event, Wyoming educators and the WDE will set what will be the proficiency levels for WY-TOPP for the 2018 Summative tests and all future Interims and Modulares. After these proficiency cut scores are determined, all future WY-TOPP tests will be reported on the 4 performance levels of Below Basic, Basic, Proficient, and Advanced.