
Responding to Questions about Assessment Results



Office of Educational Accountability

May 2007

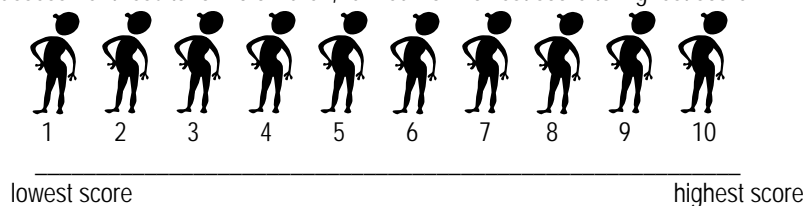
How many items do students need to get right?

We often receive questions about how many items need to be answered correctly to be proficient on the statewide exams. This can be a very challenging question to answer. Many people understand assessment based on their own experiences with classroom assessments. Often in those situations, test questions are scored either as correct or incorrect and overall competency on the classroom assessment is then reported as a percentage (number correct divided by the total number of items). Grades are assigned based on a teacher's assertion of what level of performance constitutes an A, B, C, D, or F.

The *Wisconsin Knowledge and Concepts Examinations (WKCE)* are scored by a different process. The WKCE takes advantage of a statistical procedure known as "item response theory" (IRT) to evaluate scores. Each test question – or item – is treated independently. Unlike classroom tests, all items are not created equal, but are viewed as samples of the knowledge domain with varying difficulty. The IRT calculations do not weight questions per se, however, an incorrect answer on a difficult item does not affect scores as much as an incorrect answer on an easier item. Similarly answering a more challenging item correctly will have a greater affect on the overall score than answering a more challenging item incorrectly. For large-scale standardized tests, IRT scoring is a more accurate representation of students' knowledge of the content area.

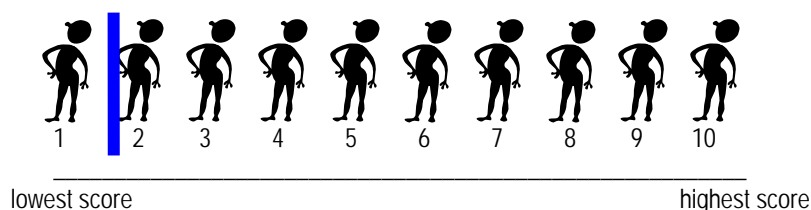
How can a student at the 20th percentile be proficient?

Assessment results are reported as criterion-referenced or norm-referenced scores. While the assessment questions may be similar, there is a significant difference in interpretation of the results. For example, a common norm-referenced report of results involves percentiles. Consider the following illustration of assessment results for 10 children, ranked from lowest score to highest score.



In the diagram above, using a norm-referenced interpretation, Student 5 scored at the 50th percentile. This means that this child performed equal to or better than half of the other students. Likewise, Student 2 scored at the 20th percentile and equal to or better than one-fifth of his/her peers. The results in this scenario answer the question, "How well did the student perform in comparison to his/her peers?" Reporting results in terms of percentiles, however, does not provide information on whether or not the students were proficient. It is possible that all students performed poorly and that none are proficient or that all are proficient. Rank ordering the student scores does not necessarily reflect an amount of knowledge and skills measured by the test.

Consider the same picture. Again, student scores are rank-ordered. This time a line has been drawn to signify the score needed to be proficient. In the case of the classroom assessments, recall that the score for proficient (or the criteria for a passing grade) is often determined by individual teacher judgment. In the case of the WKCE, a statewide panel of Wisconsin educators determined the passing score for "proficient" through a consensus process. Proficiency means that the student demonstrates the knowledge and skills necessary to be successful on grade-level work related to Wisconsin's *Model Academic Standards* and the *Assessment Frameworks*.



In this situation, one can still make comparisons between students. Student 5 continues to fall at the 50th percentile and Student 2 continues to perform at the 20th percentile. We know, however, from the determination of what constitutes proficiency that Students 2 through 10 are proficient. Thus, a student at the 20th percentile falls within the proficient range. While student 2 is proficient, students 3 through 10 are performing at a higher level relative to Student 2.

If the previous scenario was a classroom assessment, Students 2 through 10 would all receive passing grades. Under a traditional grading system, it is very possible that 90% of a class receive a passing grade. As demonstrated in the diagram, this means that students at the 20th through the 100th percentile rank passed and were considered proficient. All have reached the teacher-established benchmark for proficiency despite their rank order.

Since the WKCE is a criterion-referenced test, the goal is for all students to meet the established criterion for competency. In the accountability context of *No Child Left Behind*, the goal established by federal law is for all students to be proficient by the year 2014. In this situation, as student performance increases the number of students moving across the bar for proficient increases until all students are proficient regardless of percentile rank. The diagram below demonstrates this scenario. Note that in each graphic, Student 2 is at the 20th percentile and Student 5 is at the 50th percentile.

