



Incorporating ACT Scores into Your Statewide Assessment

Introduction

One of the most challenging issues a state must resolve in designing a statewide standards and college readiness assessment is that of how student scores should be reported. The ACT is an effective and reliable measure of student readiness for college and work, but in some cases states may wish to augment the ACT with tests of their own design. In these cases where multiple tests from different sources are included in the student assessment, how do you create a common score scale that combines information about student achievement across tests?

This question is frequently asked of ACT by states that use, or are considering using, the ACT as part of their state assessment. For example, one state wanted to augment the ACT Science Test with an additional customized science assessment. At first, the problem of how to integrate scores on state-developed tests and scores on ACT tests into meaningful combined scores may seem extraordinarily complicated, but with ACT's assistance a state may find the task not so daunting after all, as the following discussion shows.

ACT's simple, straightforward approach is based on the use of *scale* scores, in which raw scores from different forms of an assessment are converted to a common scale for reporting, and scale scores from assessments with differing requirements or differing numbers of items are combined onto a single, common scale. This method ensures both that equivalent scores on the different forms of the assessments will refer to an equivalent level of achievement and that equivalent aggregate scores made up of the individual scale scores from each assessment will refer to an equivalent level of achievement.

Raw Scores: The Starting Point

Scoring any test starts with documenting the student's performance on that test. In a multiple-choice test, this can be as simple as counting the number of questions the student got correct. In an essay test, scoring can be more complicated, whether based on an analytical approach (such as counting the number of grammatical errors a student makes) or a holistic approach (in which a piece of writing is assigned a score from, say, 0 to 4, with each score point based on a generic description of an essay's characteristics and overall quality). On either type of test, the number resulting from such documentation—55 out of 70 questions correct; 3 points out of a possible 4—is known as the *raw score*.

It's clear from our two examples that raw scores are not necessarily comparable. This is true even for two tests of the same type. A raw score of 65 on a 70-question multiple-choice test is not the same as a raw score of 65 on a 150-item multiple-choice test, nor does a raw score of 65 on a very easy 70-question test represent the same performance level as a score of 65 on a very difficult 70-

question test. The first step in developing comparable scores is to develop a common scale along which they can all be reported.

Scale Scores: Comparing Apples to Apples

The reporting score scale for an individual test or the composite of several tests can be as basic as rank ordering (for example, a student who answered 9 out of 10 questions correctly on a single form would be ranked higher than a student who answered 8 out of 10 questions correctly), as simple as a scale from 0 to 100 (in which the students in our example would have scored a 90 and an 80, respectively), or as abstract as a scale from, say, 30 to 230. Whatever the chosen scale, the important point is that it must enable scores to be arranged on a continuum that permits useful comparisons among them. Where, as we have seen, raw scores can only be compared within a single form, scale scores are always comparable across forms.

As an example, let us consider a science assessment with a composite score comprising the ACT Science Test and a state supplemental science essay test. Let's say the ACT Science Test is reported on a scale from 1 to 36 and that the supplemental test is scored holistically using a 5-point scale from 0 to 4. The state has mandated that the comprehensive science composite score is to be reported on a scale from 50 to 250. In order to come up with the composite score, we would need first to establish the weight of each test component in determining the final score. For example, the state may wish to have the ACT Science Test count 60 percent toward the total score, and the science essay test to account for the other 40 percent. Then, using a *conversion table*, we would plot all 36 possible ACT Science Test scores against all 5 possible essay test scores, with the ACT Science Test score weighted 60 to 40 against the essay test score, to get all possible combined scores. These scores would be expressed as values along the continuum from 50 to 250, giving us the final science composite score according to the state-mandated reporting scale. Because this composite score is a scale score, the values from 50 to 250 can then be used to make useful comparisons among differing degrees of student performance in science.

Conclusion

Before crafting a statewide college readiness assessment, a state must answer a number of key questions:

- What scores are to be reported? (English, Language Arts, Math, Science, Social Studies, Total, etc.)
- What should the score scale look like? (e.g, 0 to 100)
- Who are the target test takers on whom the score scale will be based? (e.g., high school juniors, high school seniors)
- If a score is based on multiple measures (for example, an English Language Arts score may be based on a multiple-choice usage test and an essay test), how much emphasis should each measure be given in the overall score?

ACT stands ready to assist you in the task of combining state-developed assessments with ACT assessments to produce a final test whose results are an accurate reflection of your state's standards and the college readiness of your high school students. ACT can help you with any or all of the following tasks:

- Determining which state assessments will be augmented with ACT tests, which can stand alone, and which content domains can be assessed with ACT tests exclusively
- Establishing the reporting scale on which state results will be reported
- Deciding how much weight to assign to each component of a multiple-measure assessment
- Choosing the population on whose performance the scale is to be based
- Collecting data and assembling conversion tables
- Conducting quality control of score results prior to reporting
- Producing and distributing score reports
- Documenting procedures for monitoring reliability of the score scale
- Developing technical manuals

With ACT's assistance, your state can determine a score scale for a state standards and college readiness assessment containing multiple measures that usefully and efficiently combines information about student achievement across tests. ACT's straightforward scale-score method ensures that a total score made up of the individual scores from each assessment will yield a meaningful "snapshot" of student performance.

For more information, please contact us.