

Using NBME Subject (Shelf) Test Scores as Part of a Final Course or Clerkship Score

The National Board of Medical Examiners (NBME) reports student scores from its subject (shelf) exams in what the NBME calls a Subject Test Score (STS). While reported as a two-digit score, the STS is not a percent score. Consequently, the STS should not be treated as such in calculating a student's final percent score in a course or clerkship.

The Subject Test Score is a standard score. What happens when **raw scores** (the actual number of items answered correctly on an exam – e.g., 51 would be a student's raw score if she answered 51 of 55 test items correctly) or **percent scores** (e.g., 51 of 55 would be 93%) are **standardized**? The **mean** and **standard deviation** are calculated from ALL exam scores. We now can tell a student how much above or below the mean she is for the group taking the exam. Thus, if T1 Susan Smith were 1.17 standard deviations above the mean on her Physiology exam, we would say her standard score is 1.17.

If an exam does not yield scores that are normally distributed (typically the case with local medical school exams), the value of knowing a student's standard score is limited. We know that Ms. Smith scored higher than most of her classmates, but we cannot interpret her score in normal distribution terms.

However, standard scores for an NBME shelf exam are calculated from a large population certain to yield a normal distribution of scores. If Susan Smith were a T3 in Surgery, her standard score of 1.17 on the Surgery shelf exam might tell us that she is at the 88th percentile, or that she scored better than 88 percent of the national group on which the standard scores was based.

While psychometricians prefer to report standard scores in standard deviation units, this is not necessary. At its origin USMLE Step standard scores had a mean of 200 and a standard deviation of 20. A student with a score of 220 was one standard deviation above the mean or at the 84th percentile.

The NBME subject (shelf) exams are reported as standard scores using a mean of 70 and a standard deviation of 8. This unfortunate choice of a two-digit scale can be easily confused with percents. The take-home message is: **DO NOT TREAT A SUBJECT TEST SCORE FROM A SHELF EXAM AS A PERCENT**. This is an avoidable measurement error, and the resulting course or clerkship final "percent" score will be both invalid and a disservice to students.

What NOT TO DO:

- Medical schools sometime use two methods to “convert” shelf exam standard scores to percents. One approach is to change the mean STS of 70 to a typical local test mean percent score, say 80%, and use the STS standard deviation of 8 to adjust scores above and below the mean. A related approach is simply to multiply the STS score by some factor (e.g., 1.15 or 1.25). **Both methods are psychometrically unsound practices that should be avoided.**

What TO DO:

- The NBME will provide students’ actual percent scores from a shelf exam (if you are insistent). You will find that the distribution of shelf exam percent scores is lower than your local exam scores. It is defensible to adjust shelf exam percent scores to conform to local standards.
- Some medical schools use the percentiles paralleling STS standard scores as cutoffs for the Honors grade. For example, to obtain an Honors in Surgery, you must (1) successfully complete all required procedures, (2) be rated superior on all faculty/resident clinical evaluations, AND (3) be at or above the 90th percentile on the NBME shelf exam.

This approach has the advantage of controlling the number of Honors grades in a clerkship. Further, the NBME will send quarterly norms for your shelf exam, allowing you to use a changing standard score cutoff to determine Honors as the year progresses. These quarterly norms reduce the advantage of taking a clerkship later in the year when the student will have benefited from her experiences in prior clerkships.