Introduction to Using Released Items with Student Work

"Student Work" Added to Released Items

It is important for both students and teachers to have clear examples of the quality of work required to earn maximum credit on each of the constructed response items. Teachers can use these examples to lead classroom discussions of what the characteristics are that make the higher scored samples worthy of the higher scores. Ultimately, of course, the goal is that students will use this awareness to self-monitor their responses and hopefully to improve them.

Other Information Is Provided for Each Item

Each item includes the standard tested, the number of points each item is worth, the *mean* score it received when each item was field tested, the rubric/scoring guide, and sample student responses

Educators are familiar with standards, benchmarks, point values and rubrics/scoring guides but may not be familiar with the mean score statistic. In short, the mean score indicates what fraction of the possible points was earned by students statewide on each item (expressed in a decimal with two places).

This statistic is helpful in two ways. First, knowing the state mean score provides an index of the relative difficulty of items. For example, a 2 point item that has a mean of 1.75 is a much easier item than one with a mean of .53 (which means that statewide, students on average only earned a little over ¼ of the points possible for the item compared to the 1.75 item where students earned 7/8 of the points. Students and teachers will profit by discussing why one item would be so much harder than another and what more is demanded of the student by the harder item (and then practicing producing more well developed responses on subsequent constructed response practice activities.)

Second, knowing what the average percentage of points students across the state earned on an item, will provide some context and possibly comfort as students work to improve their own performance. On many 4 point items, for example, a mean of anything over 2 is considered a relatively high score by the psychometricians who provide the data at SBA Data Review Workshops. Students should not be discouraged if they don't 'earn' full point value for a response but rather recognize that many students are in the same situation; then they can focus on ways to increase the fullness of their responses on subsequent practice activities and on the SBA.