Balanced Assessment for the Mathematics Curriculum: classroom packages

Balanced Assessment Project—Pearson Learning (Dale Seymour Publications)

Summary

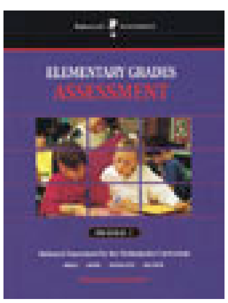
A typical Balanced Assessment (BA) Package offers ten to twenty tasks, ranging in length from 5 to 45 minutes. Some of the tasks consist of a single problem, while others consist of a sequence of problems. Taken together, the tasks provide students with an opportunity to display their knowledge and skills across the broad spectrum of content and processes described in the *Principles and Standards for School Mathematics (PSSM)*. It takes time to get this kind of rich information—but the problems are mathematically rich and well worth the time spent on them. Materials are provided to help teachers use the tasks in their own classroom and how to think about student work.

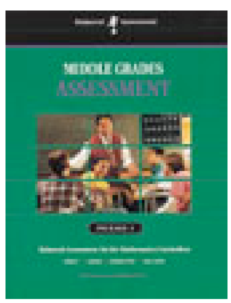
Purpose

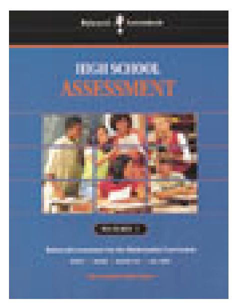
To provide assessment tasks, for Grades 3-12, aligned with *Principles and Standards for School Mathematics (PSSM)*, so as to recognize the broad range of performance these standards require, with some associated classroom materials.

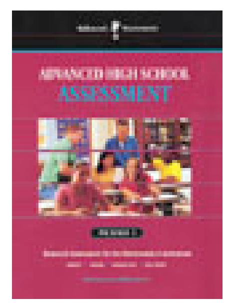
The Tool

There are two assessment packages at each grade band. Pearson Learning (Dale Seymour Publications) publish the assessment packages for upper elementary, middle grades, high school, and advanced high school students.









Using the tool

The tasks in the packages can be used for multiple purposes:

- **Implementing performance assessment under controlled situations**. Each package contains a balanced set of tasks appropriate for such on-demand, high-stakes assessment. Schools may choose to use any of the tasks as a basis of finding out how students across a particular grade level are making sense of their mathematical experiences. The tasks could be used for alignment of curriculum and instructional programs.
- **Classroom-based performance assessment**. Teachers may use the tasks as embedded assessments within their curriculum. The tasks may provide new insights into students' thinking and reasoning at the end of a chapter or unit of instruction. Additionally, a task could be used to see how students are transferring their knowledge from known content and curriculum-based problems to unknown situations and problems that are not familiar to them.

- **As a lever for reform**. The collection of tasks may be used as a lever to help teachers and students move toward standards-based curricula. In this case, the tasks may be used as enrichment for classroom activities. Upon using the tasks with students, teachers frequently seek additional resources that will help them prepare students to be better able to respond to constructed response tasks.
- **For professional development:** The high-quality tasks in each package provide rich opportunities to engage teachers in discussion about curriculum, teaching, learning, and assessment. Professional development through assessment provides opportunities for teachers to broaden their knowledge of mathematics while learning how to use the tasks for student assessments. Conversations that focus on student understandings and next instructional steps prove to be quite valuable to a teacher's professional growth.

Implementation Issues

The BA tasks are more complex, engaging, and differ significantly from the routine exercises that can be found in many mathematics textbooks. Teachers may find it challenging to find a "good fit" with their existing curriculum materials, since the tasks often cross content strands. The packages include tasks where students are required to solve multi-step problems, model a situation, evaluate, generalize, justify, explain reasoning, and communicate their findings to a specified audience. The process of writing and explaining thinking may not be a routine practice in the mathematics classroom. Teachers will need guidance to interpret and score student work and use the results to inform instruction.

The tasks in a package vary in time expectations and may require 5 to 45 minutes of class time for student completion and additional time for whole group discussion. Though some teachers may be reluctant to use a whole class period to solve and discuss a longer task, they express a need to provide opportunities for their students to solve these types of problems. Teachers gain valuable insights into their students' understandings and struggles when examining their written work on performance tasks and classroom discussions around ways of reasoning about the mathematics in the tasks.

What's included with each task?

Overview

- Task Description—the situation that students will be asked to investigate or solve.
- Assumed Mathematical Background—the kinds of previous experiences students need to have had to engage the task productively.
- Core Elements of Performance—the mathematical ideas and processes that are central to the task.
- Circumstances—the estimated time for students to work on the task; the special materials that the task require; whether students work individually, in pairs, or in small groups; and any other such information.

Task Prompt. These papers are black line masters that teachers may duplicate for the students in their classroom.

Sample Solution. Each task is accompanied by at least one solution; where there are multiple approaches to the problem, more than one may appear.

Using this Task. Suggestions for launching the task and helping students understand the context are given in this section. Some tasks have pre-activities; some have students conduct initial exploration in pairs or as a whole class then write individual conclusions.

Characterizing Student Performance. This section contains descriptions of characteristic student responses the task is likely to elicit. These descriptions, based on the Core Elements of Performance, indicate various levels of successful engagement with the task. They are accompanied by annotated artists' renderings of typical student work. The examples were chosen to show something of a range and variety of responses to the task, and various aspects of mathematical performance it calls for.

The commentary is intended to exemplify these key aspects of performance at various levels across several domains. The discussions of student work are deliberately qualitative and holistic, avoiding too much detail. They are designed to focus on the mathematical ideas that "count," summarized in the Core Elements of Performance for each task. They offer a guide to help teachers and students look in some depth at a student's work in the course of instruction, considering how it might be improved.

More information may be found at the Pearson Learning website:

Take a tour of the packages at *http://pearsonlearning.com/dalesey/ba_default.cfm*.

Sample assessments from the Elementary Grades, Middle Grades, High School, and Advanced High School Assessment packages may be obtained at:
http://pearsonlearning.com/dalesey/samplepages/dsp_smplr_assessment.cfm

A sample middle grades task and some of its supporting materials follows.

Task 8

Overview

Choose appropriate statistical tools to analyze data.
Draw conclusions based on statistical measures.

Best Guess?

Short Task

Task Description

Students are given data sets on three people's estimates of when 30 seconds has passed. Students use the data sets to decide who is best at estimating a 30-second interval.

Assumed Mathematical Background

It is assumed that students have had experience with using statistical tools (including measures of center and spread) to analyze data.

Core Elements of Performance

- choose appropriate statistical tools to analyze variation and center
- apply statistical tools accurately
- use results from data analysis to draw and support conclusions for a situation

Circumstances

Grouping:	Students complete an individual
Materials:	No special materials are needed f
Estimated time:	15 minutes

Best Guess?

Name Date

Best Guess?

This problem gives you the chance to

- *choose appropriate statistical tools*
- *apply statistical tools accurately*
- *analyze data and use results to draw and support conclusions*

On your own

Hugh, Rita, and Sandy were trying to see who could make the closest estimate of when a 30-second interval had passed. They took turns guessing while someone else kept time with a stopwatch.

When it was Hugh's turn to guess, the following times were recorded from the stopwatch:

31 25 32 27 28

When it was Rita's turn to guess, the following times were recorded from the stopwatch:

37 19 40 36 22

When it was Sandy's turn to guess, the following times were recorded from the stopwatch:

32 38 24 32 32

1. Who do you think is best at estimating 30 seconds? ____________

2. Give reasons for your choice.

A Sample Solution

Task 8

Hugh is the best guesser because the range of his guesses is smallest (the difference between his lowest guess and his highest guess is only 7 seconds), and he is never more than 5 seconds off on any turn. Hugh's median guess is 28, only 2 seconds off. Sandy's median guess is 32 which is also only 2 seconds off, but she has two guesses that are 6 and 8 seconds off and on opposite sides of the 30-second mark, making her range of guesses larger than Hugh's. In this situation, the mean is not a reasonable measure to evaluate the people's guessing ability because Rita's guesses are between 6 and 11 seconds off and yet balance out with the mean to be very close (30.8) to the desired goal of 30 seconds.

More on the Mathematics

Students may choose to use a number of different statistical measures in determining their answer. For this task, spread of the data around 30 is the most powerful measure for determining who is the best guesser. A strong argument would consider measures of center and the spread of the data and support a choice based on these measures and their relationship. Using these tools, Hugh and/or Sandy could be supported as being the better guesser. Any combination of statistical measures could not support Rita as a reasonable choice.

The spread of the data shows that Hugh does the best with his guessing. Rita's guesses are very spread out (range of 19 to 40, spread of 21). Sandy's guesses are more spread out (range of 24 to 38, spread of 14) than Hugh's, whose data is the most compact (range of 25 to 32, spread of 7).

Sandy has a mode guess of 32, neither of the other two has a mode gues
mode itself is not a very useful statistic for this situation. What makes S
mode worth noting is that she is very close to the desired goal of 30 sec
To argue that she is the better guesser *only* because she guessed the san
number three times is a weak response.

Looking at median, Hugh and Sandy would be tied as the better guesse
median guess of 28 and 32, respectively, (both 2 from 30) compared to I
median of 36.

Regarding mean, Rita has a mean guess of 30.8, Sandy has a mean gues
31.6, and Hugh has a mean guess of 28.6. The back-and-forth guessing
makes this a weak measure in helping to decide who is the better guess

Using this Task

Task 8

Review the aims of the assessment in the box at the top of the activity page. Read the problem aloud to your students. Because this task is short and straightforward, it would be appropriate simply to hand out the task and direct students to read and answer the questions individually.

Issues for Classroom Use

At first glance, this problem appears small and contained when actually it provides an opportunity to see which statistical tools students bring to bear on a problem as well as how they make sense of these measures. The fact that some students use only the mean to justify their choice, failing to think about the range and what it says about the data set, shows a lack of understanding of statistics on the part of the student. Students must also consider the spreads of the data for each guesser in order to address the task completely.

Characterizing Performance

Task 8

This section offers a characterization of student responses and provides indications of the ways in which the students were successful or unsuccessful in engaging with the task. The descriptions are keyed to the *Core Elements of Performance*. Our global descriptions of student work range from "The student needs significant instruction" to "The student's work meets the essential demands of the task." Samples of student work that exemplify these descriptions of performance are included below, accompanied by commentary on central aspects of each student's response. These sample responses are *representative*; they may not mirror the global description of performance in all respects, being weaker in some and stronger in others.

The characterization of student responses for this task is based on these *Core Elements of Performance*:

1. Choose appropriate statistical tools to analyze variation and center.
2. Apply statistical tools accurately.
3. Use results from data analysis to draw and support conclusions for a situation.

Descriptions of Student Work

The student needs significant instruction.

Student selects and attempts to apply inappropriate or vague/unclear statistical tool(s) (for example, student might find the sum for each set of data).

The student needs some instruction.

Student selects (or invents) and applies statistical tools that provide information about either the variation/spread or the center OR tools and calculations are absent or unclear. There is, however, evidence in the reasoning that the student was considering a method that would take into account both measure of variation and center.

The student's work needs to be revised.

Student selects and applies statistical tools that provide information about both the spread and the center (for example, student attempts to count the number of times the guess is within 2 seconds of 30; student attempts to use both the median and the range of guesses). While the student recognizes the importance of both types of information, there is no evidence that the student realizes that for this context the variation is the more important of the two aspects. There may be come calculation errors or problems in articulating reasoning.

The student's work meets the essential demands of the task.

Student correctly calculates or states measures that address both the spread and the center for two or three sets of data, recognizes the greater importance of spread for this context, and in the justification for the best guesser compares the statistics found between at least two of the guessers.

Best Guess? ▪ Student Work

Student B

This problem gives you the chance to

- *choose appropriate statistical tools*
- *apply statistical tools accurately*
- *analyze data and use results to draw and support conclusions*

On your own

Hugh, Rita, and Sandy were trying to see who could make the closest estimate of when a 30-second interval had passed. They took turns guessing while someone else kept time with a stopwatch.

When it was Hugh's turn to guess, the following times were recorded from the stopwatch:

31 25 32 27 28 143 total sec

When it was Rita's turn to guess, the following times were recorded from the stopwatch:

37 19 40 36 22 154 total sec

When it was Sandy's turn to guess, the following times were recorded from the stopwatch:

32 38 24 32 32 158 total sec

1. Who do you think is best at estimating 30 seconds? Hugh
2. Give reasons for your choice. Hugh ~~be~~ is because he had the most closest guess to 30 seconds and because he had the lowest number of seconds, when you add them up.

Best Guess? 153

Student B

The student needs significant instruction.

Student B found the sums of the data points, which alone is not helpful for deciding the best guesser. Further, while the reasons given refer to Hugh's having the "most closest guesses to 30 seconds," there is no method or tool used to provide the data analysis needed to support this reason.

Best Guess? ■ Student Work

Student C

This problem gives you the chance to

- *choose appropriate statistical tools*
- *apply statistical tools accurately*
- *analyze data and use results to draw and support conclusions*

On your own

Hugh, Rita, and Sandy were trying to see who could make the closest estimate of when a 30-second interval had passed. They took turns guessing while someone else kept time with a stopwatch.

When it was Hugh's turn to guess, the following times were recorded from the stopwatch:

31 25 32 27 28 28.6

When it was Rita's turn to guess, the following times were recorded from the stopwatch:

37 19 40 36 22 30.8

When it was Sandy's turn to guess, the following times were recorded from the stopwatch:

32 38 24 32 32 31.6

1. Who do you think is best at estimating 30 seconds? Hugh and Rita
2. Give reasons for your choice. Hugh got close to 30 every time

Hugh averaged 28.6
Rita Averaged 30.8 Add the 5#'s togather and ÷ by 5

154 *Middle Grades Package 1*

Student C

The student needs some instruction.

Student C finds the mean for each data set. While the student refers to the notion of variation/spread ("Hugh got close to 30 every time"), there is no defined statistical method or tool that is applied to the other data sets. The student needs some instruction and not just a revision of his work.