Discussion of the Items

ITEM 1 IN **FIGURE 1** ASKS STUDENTS TO FIND the various measures of central tendency. This item is intended to assess whether students know and can distinguish among the procedures for finding each. Brown and Silver (1989) reported that although the high school students did better than the middle school students in 1985–1986, fewer than half, generally only about 40 percent in each grade, responded correctly with 15 for the mode and 16 for the median and mean, when they responded at all (see **table 1**). Another item administered at the same time used the term *average*, asking for the average age of six children ages 13, 10, 8, 5, 3, and 3. Almost all the students responded to the item, perhaps indicating a greater familiarity with the term *average*. Interestingly, the percent of correct answers from responding seventh graders was not much higher for the *average* item (46%) than for the *mean* item (40%); however, the high school students were more successful on the item that used the term *average*. You may be interested in determining whether your students respond differently when the word *average* is used instead of *mean*.

These students did not understand the advantages of each statistic

Item 2 in **figure 1** asks students to identify the median when the data are represented in a scatterplot. The item assesses the combined knowledge of interpreting a graph and using the procedure for finding the median. Zawojewski and Heckman (1997) reported that only a little more than a fifth of eighth-grade students and fewer than a third of twelfth-grade students in the NAEP samples for those grade levels responded correctly, as shown in **table 2.** One common wrong answer was 55 (choice E), which is disturbing because this response may indicate that some students may have added the labels on the *y*-axis (30 + 40 + 50 + 60 + 70 + 80 = 330), then divided by 6 to get 55. If so, these students are not only confused about the median and the mean but also unable to use and interpret information given in graphical form. If students in your class respond with choice E, ask students in a follow-up interview question or writing prompt why they chose this answer.

Item 3 in **figure 1** is different from the first two because it requires students to make a choice between mean and median rather than find the measures of central tendency. This type of item assesses students' conceptual understanding of mean and median, which is different than just knowing the procedures for finding them. The written responses indicated that a number of students had not made their choices on the basis of the mathematical characteristics of the two measures of central tendency. Instead, when faced with a choice of mean or median, some students selected the mean, apparently without regard for the shape of the distribution. Some claimed that the mean is the better choice because it is the typical value, or the average, as illustrated by one student's comment, "The mean. To make a generalization of 'typical' attendance, averages are used, not middle points." These types of responses seem to imply that students may think that the median is not representative of a typical value, whereas the mean is. Others claimed that the mean was better because it was superior to the median in some way, as illustrated by the student who wrote, "The mean. An average gives a more accurate # because it involves all the #s."

The idea that the mean is more precise, or more accurate, than the median may actually reveal some understanding that the procedure for the mean incorporates all the values, whereas the median reports just one value. These and other responses suggest that these students did not have an understanding of the trade-offs and relative advantages of each statistic. The prevalence of explanations indicating an "absolute" belief that the mean is better than the median, no matter what, may in part explain why only 4 percent of the grade-12 students in the 1996 NAEP responded with correct answers for item 3—that is, that the median is appropriate for Theater A and the mean, for Theater B—with a complete explanation for at least one measure (see **fig. 2**).

In examining students' difficulty with summary statistics, we must also think about how the problems, or tasks, and their scoring rubrics are designed. For example, item 3 in **figure 1** asked students to respond without knowing why they needed to decide between the two measures of central ten-

TABLE 1
Percent Correct and Response Rate for Students in Grades 7 and 11 on Item 1 (Fourth NAEP, 1985–1986)

	Percent Correct [Response Rate]	
Item	Grade 7	Grade 11
a. What is the mode?	26 [.65]	40 [.41]
b. What is the median?	38 [.65]	47 [.41]
c. What is the mean?	40 [.66]	41 [.72]

(Brown and Silver 1989, 29)

dency. To get the highest score (called "extended" by NAEP) on this item, a response had to include a statement that attendance on Day 4 for Theater A (10 attendees) was much lower than for the other days and how this outlier can affect the mean. Because the NAEP scoring rubric accepted as "extended" only those responses that addressed the distribution of the data, the implication is that whenever a set of data contains an outlier, the median is the best statistic to use regardless of purpose. Imagine a situation in which the attendance figures of the two theaters are to be compared directly; in this instance, it could be argued that the identical statistic should be reported for both theaters. In fact, reporting both the mean and the median for each theater would be very effective for making direct comparisons. A statistician might assume that the theater attendances were to be compared using a specific statistical test, such as a *t*-test; if so, only the mean, not the median, would be required for Theaters A and B. Because the students taking the test were neither given nor asked to make up a reason for choosing the mean or the median, many may have simply chosen the statistic with which they were more familiar.

The NAEP results raise some questions that you may want to ask yourself as you review your students' performance:

- Do your students understand the procedures for finding the mean and median? For example, if they are able to find the mean or median of the years in item 1 rather than the inches of snow, you will be able to tell that your students know the procedure but do not understand when or where to apply it.
- Do your students understand the terminology of *mean* and *median*? For example, if you ask similar questions using *average* or *middle data point* instead of *mean* or *median*, you will be able to tell whether the students are connecting the words to known procedures.
- Do your students make mathematical connections between statistics and other branches of mathematics? For example, if they are able to interpret points on a scatterplot and apply their understanding of median, tasks such as item 2 will help you determine whether they can use their combined knowledge on a single task.
- Do your students understand the relationship between the distribution of the data set and the selection of mean and median? For example, when students explain their choices of mean or median in item 3, you can determine whether they are using information about outliers in their decision making.

TABLE 2

Percent of Students in Grades 8 and 12 Responding to Choices on Item 2 (Sixth NAEP, 1992)

	Percent Responding	
Choice	Grade 8	Grade 12
A	5	2
B	11	6
C	32	27
D (correct)	23	31
E	26	32

(Zawojewski and Heckman 1997, 215)

a. "The median. Day 4's attendance of 10 obviously lowered the mean a good bit so the median would be more typical."
b. "I would use the mean, since the median gives an artificially low number—it does not reflect at all the two days of high attendance."

Fig. 2 Sample correct response for item 3

Implications for Teaching

THE NAEP DATA WE HAVE SHARED HERE CAN help illuminate aspects of students' understanding of measures of central tendency that need attention, such as confusion about the procedures for finding mean and median, as well as difficulty selecting appropriate statistics. Furthermore, other summary statistics are equally important for developing students' conception of a distribution, such as measures of spread and variation (i.e., range, standard deviation, confidence interval, and so on). Unfortunately, past NAEP assessments had few items assessing students' understanding of variation and spread, and none of these items has been released yet.

You can, however, incorporate some of your own questions into items that are similar to these NAEP items to assess your students' understanding of spread, as well as of center. For example, you might want to implement the middle school activity on standard deviation described by Wilmot (1991) in *Dealing with Data and Chance*. Only through additional data gathering can you, as the teacher, understand student difficulties and use that insight to guide your instruction. As the classroom teacher, you are in a good position to probe students' knowledge by asking, orally and in writing, such follow-up questions as those suggested previously. You can also modify tasks to elicit more explanation from students or to provide a familiar context that may elicit better responses.

In addition, you can enhance your curriculum by selecting or creating additional worthwhile tasks to contribute to both teaching and assessment opportunities for data analysis. One of the reasons that students do not find the concepts of mean and median easy may be that they have not had sufficient opportunities to make connections between *centers* and *spreads*; that is, they have not made the link between the measures of central tendency and the distribution of the data set. As you take time for action, you can create teaching and assessment opportunities by selecting items from large-scale assessments, such as the NAEP, as well as tasks from supplementary curriculum materials, to guide your own data-driven instructional decisions.

References

Brown, Catherine A., and Edward A. Silver. "Data Organization and Interpretation." In *Results from the Fourth Mathematics Assessment of the National Assessment of Educational Progress*, edited by Mary M. Lindquist, 28–34. Reston, Va.: National Council of Teachers of Mathematics, 1989.

Wilmot, Barbara. "Exploring Standard Deviation." In *Dealing with Data and Chance, Curriculum and Evaluation Standards for School Mathematics* Addenda Series, Grades 5–8, 29–32, 68–69. Reston, Va.: National Council of Teachers of Mathematics, 1991.

Zawojewski, Judith S., and David Heckman. "What Do Students Know about Data Analysis, Statistics, and Probability?" In *Results from the Sixth Mathematics Assessment of the National Assessment of Educational Progress*, edited by Patricia Ann Kenney and Edward A. Silver, 195–224. Reston, Va.: National Council of Teachers of Mathematics, 1997.

Zawjewski, Judith, and J. Michael Shaughnessy. "Data and Chance." In *Results from the Seventh Mathematics Assessment of the National Assessment of Educational Progress*, edited by Edward A. Silver and Patricia Ann Kenney. Reston, Va.: National Council of Teachers of Mathematics, in press.

"Take Time for Action" encourages active involvement in research by teachers as part of their classroom practice. Readers interested in submitting manuscripts pertaining to this theme should send them to "Take Time for Action," MTMS, *NCTM, 1906 Association Drive, Reston, VA 20191-9988.*

Preparation of this article was supported in part by a grant to the National Council of Teachers of Mathematics from the National Science Foundation, grant no. RED-9453189, with additional support from the National Center for Education Statistics. Any opinions expressed herein are those of the authors and do not necessarily reflect the views of the National Council of Teachers of Mathematics, the National Science Foundation, or the National Center for Education Statistics.