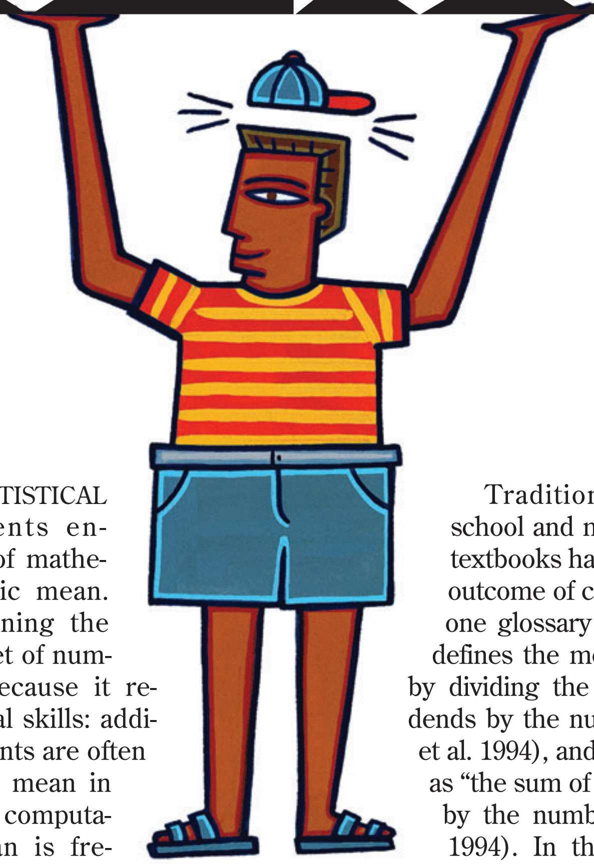
ELIZABETH GEORGE BREMIGAN

# Developing a Meaningful Understanding of the MEAN



ONE OF THE FIRST STATISTICAL measures that students encounter in their study of mathematics is the arithmetic mean. The procedure for determining the arithmetic mean of a given set of numbers is relatively simple, because it requires only two computational skills: addition and division. Thus, students are often introduced to the arithmetic mean in grades 4 or 5. At this level, computation of the arithmetic mean is frequently presented as an application of division rather than as a statistical concept. Initially, the arithmetic mean is often called the *average*; in this article, the arithmetic mean will be referred to simply as the *mean*.

BETH BREMIGAN, *egbremigan@bsu.edu, teaches at Ball State University, Muncie, IN 47306. She enjoys designing group activities and projects for students and for preservice and in-service mathematics teachers.*

Traditionally, many elementary school and middle school mathematics textbooks have defined the mean as the outcome of computations. For example, one glossary in a sixth-grade textbook defines the mean as a number obtained by dividing the sum of two or more addends by the number of addends (Bolster et al. 1994), and another explains the term as "the sum of a collection of data divided by the number of data" (Hoffer et al. 1994). In the classroom setting, students and teachers might easily assume that one's ability to successfully compute the mean of a given data set is synonymous with achieving an understanding of the concept of the mean. However, the possibility exists for students to successfully compute the mean without developing an understanding of what the mean actually represents, how it is related to the numbers in the data set, or how it is related to other measures of the center or spread of a data set (Zawojewski and Shaughnessy 2000).

For many years, I taught students the procedure for finding the mean without paying much attention to their conceptual understanding of this important statistical idea. My instructional decisions emphasized the development and application of a numerical procedure. The assessments I used measured students' abilities to successfully execute this procedure. The focus of the problem situations I presented and the questions I posed to students were limited. Although I knew that understanding the mean involved more than simply performing a procedure, it was not clear to me what a conceptual understanding of the mean was. I found myself asking the question, "What does it mean to understand the mean?"

An answer to my question came when the results of some mathematics education research caught my attention. In particular, Strauss and Bichler (1988) analyzed the concept of the arithmetic mean into its properties and investigated the development of children's understanding of these properties. The results of this research helped me to define conceptual understanding of the mean for students. In turn, my definition significantly changed how I taught this concept to middle school students and is clearly reflected in my current teaching of preservice elementary and middle school teachers.

Strauss and Bichler (1988) identified seven properties of the arithmetic mean that students must comprehend (see **fig. 1**). These properties serve as a useful guide for my instructional decisions as I seek to help students develop a rich conceptual understanding of the mean. My awareness of these properties has influenced my selection and construction of tasks that have the potential to lead students to encounter and pay specific attention to various aspects of the mean. The types of questions that I ask students during class discussions and the ways I respond to students' comments and questions have been affected by the properties that Strauss and Bichler defined. The results of their research have provided me with a framework to evaluate and improve my teaching of this mathematical topic in both significant and practical ways.

*Principles and Standards for School Mathematics* (NCTM 2000, p. 248) offers additional insight into the teaching and learning of this important statistical concept. The Data Analysis and Probability Standard advocates that teachers at all grade levels make instructional decisions that enable students to "select and use appropriate statistical methods to analyze data." In the middle school grades, students should learn to "find, use, and interpret measures of center and spread." This Standard draws our attention to additional aspects of the conceptual understanding of the mean, specifically, the idea that students must understand the mean's relationship to other measures of center (e.g., median and mode) and to measures of spread (e.g., range and standard deviation).

A. The mean is located between the extreme values.
B. The sum of the deviations from the mean is zero.
C. The mean is influenced by values other than the mean.
D. The mean does not necessarily equal one of the values that was summed.
E. The mean can be a fraction that has no counterpart in physical reality.
F. When one calculates the mean, a value of zero, if it appears, must be taken into account.
G. The mean value is representative of the values that were averaged.

**Fig. 1 Properties of the mean (Strauss and Bichler 1988)**

During a fund-raising candy sale, the twenty-nine students in one seventh-grade class are divided into five groups. A prize is offered to the group that is the most successful. The amounts of money collected by the five groups of students are given below. Your goal is to determine which group is the most successful and should win the prize. Describe several different ways that you could reach a decision about which group is the most successful.

| Group 1 | Group 2 | Group 3 | Group 4 | Group 5 |
|---|---|---|---|---|
| 50 | 50 | 100 | 70 | 70 |
| 40 | 50 | 45 | 60 | 60 |
| 35 | 35 | 20 | 50 | 50 |
| 30 | 25 | 20 | 20 | 20 |
| 30 | 25 | 15 | 10 | 10 |
| 25 | 25 | 10 | 0 | |

**Fig. 2 The Fund-Raising Contest**

As teachers, we are faced with the challenge of finding or creating problem situations and leading classroom discussions in which our students can explore the properties of the mean and the relationship of the mean to other measures of center and to measures of spread of the numbers in a data set. The following sections present two problem situations that I designed and used to meet these goals. References to Strauss and Bichler's properties of the mean are included in the discussions of the tasks.

## The Fund-Raising Contest

THE FIRST PROBLEM, THE FUND-RAISING CONtest, shown in **figure 2,** was inspired by a debate that occurred in the middle school where I was teaching. Students were struggling with the question of how to fairly determine which group had the best performance during a fund-raising contest when the groups were of different sizes. In designing this open-ended task for

| Money Raised by Students in Group 1 | Difference between Money Raised by Students in Group 1 and the Mean |
|---|---|
| 50 | 50 – 35 = 15 |
| 40 | 40 – 35 = 5 |
| 35 | 35 – 35 = 0 |
| 30 | 30 – 35 = –5 |
| 30 | 30 – 35 = –5 |
| 25 | 25 – 35 = –10 |
| Mean = 35 | Sum of deviations from the mean = 0 |

| Money Raised by Students in Group 3 | Difference between Money Raised by Students in Group 3 and the Mean |
|---|---|
| 100 | 100 – 35 = 65 |
| 45 | 45 – 35 = 10 |
| 20 | 20 – 35 = –15 |
| 20 | 20 – 35 = –15 |
| 15 | 15 – 35 = –20 |
| 10 | 10 – 35 = –25 |
| Mean = 35 | Sum of deviations from the mean = 0 |

**Fig. 3 Deviations from the mean for groups 1 and 3**

my seventh-grade class, I used the same context but adjusted the numbers to serve my instructional purposes.

I have found that this open-ended task generally works well when presented to small groups of students first, then discussed in a whole-class setting. Before performing any computations, many students suggest that the fund-raising group that collected the most money simply should be declared the winner. Students are surprised to find that each of the five groups collected the same amount of money. This finding leads students to consider other ways to compare the groups. Frequently, students then decide that finding the mean of each of the five groups could be useful and important. I usually present this task after students have learned how to identify the mean, median, mode, and range of a given data set. Solving the fund-raising task gives students an opportunity to apply their mathematical knowledge and skills in a real-world context.

Students' attention begins to focus on differences among, and within, the five groups. One difference among groups that quickly becomes apparent is that group 5 has one fewer person than the other four groups. Although the total amount of money collected for each group is the same, the mean for group 5 is greater than the mean of the other four groups because the number of students in group 5 is smaller. In examining the data sets and means of groups 4 and 5, students directly encounter property F as they compare the impact of a student's contribution of $0 in one group with the impact of being one member short in another group.

Differences in the spread, or distribution of the data in each group, also emerge as an interesting aspect of this problem situation. Determining the range of amounts collected by each fund-raising group is a common decision that small groups make in the classroom. Students attempt to find a way to measure the contributions of individual members of each fund-raising group in terms of how "equal" they are. Some students argue that a smaller range means that each member of the group has contributed "more equally" to the fund-raising effort. This observation can lead easily into a discussion of the location of the mean between the extreme values of the data set, as noted in property A, and recognition that the mean does not necessarily equal one of the values that was summed, as noted in property D.

Another way students attempt to measure the equality of contribution is to compare the money raised by individual group members with the group mean. Students are often surprised to discover that in all five groups, the sum of the differences of the mean and the numbers below the mean is the same as the sum of the differences of the mean and the numbers above the mean. If the differences between the mean and the numbers below the mean are assigned negative values and the differences between the mean and the numbers above the mean are assigned positive values, then we can say that the sum of the differences, or deviations from the mean, is 0. **Figure 3** shows that the sum of the deviations from the mean is 0 for both groups 1 and 3, even though the pattern of individual differences from the mean is quite different within, and between, the two groups. Guiding students to explore and discuss this aspect of the problem situation not only leads them to encounter property B but also lays the foundation for their future study of the standard deviation of a data set.

## What Happens If . . .?

ANOTHER TASK I CREATED THAT WAS DIRECTLY inspired by the properties identified by Strauss and Bichler (1988) is shown in **figure 4.** Note that each of the numbers presented in the given data set is a positive integer and that the data set is presented without a context. I deliberately constructed this data set so that the mean would not be an integer to give students a situation in which the mean can be a fraction, which is noted in property E.

One of the challenges presented to students in this task is to construct a realistic context that this data set could represent. In my experience, many

Suppose that you are given the following data set:

1, 1, 2, 2, 2, 2, 3, 3, 4, 5

(a) Determine the mean and the median of the given set of numbers.
(b) Identify a real-life situation that this set of numbers could realistically represent.
(c) What happens to the mean if a new number, 2, is added to the given data set? Explain why this result occurs.
(d) What happens to the mean if a new number, 8, is added to the given data set? Explain why this result occurs.
(e) What happens to the mean if a new number, 0, is added to the given data set? Explain why this result occurs.
(f) What happens to the mean if two new numbers, 2 and 3, are added to the given data set? Explain why this result occurs.
(g) Find two numbers that can be added to the given data set and not change the mean. Explain how you chose these two numbers.
(h) Find three numbers that can be added to the given data set and not change the mean. Explain how you chose these three numbers.
(i) What happens to the mean if a new number, 30, is added to the given data set? How well does the mean represent the new data set? Can you find another statistical measure that better represents the data set?
(j) Find two numbers that can be added to the given data set that change the mean but not the median. Explain how you chose these two numbers.
(k) Find two numbers that can be added to the given data set that change the median but not the mean. Explain how you chose these two numbers.

**Fig. 4 The What Happens If . . .? task**

students construct a context for this data set in which only whole numbers "make sense," such as the number of televisions in a household or the number of pets owned by students in the class. In these contexts, the numbers in the data set represent discrete quantities, but the mean, as property E states, "can be a fraction that has no counterpart in physical reality." For example, if the data set represents number of pets owned by students in the class, students have the opportunity to consider how 2.5 can be the mean number of pets while recognizing that having 2.5 pets is physically impossible. Other students may construct contexts in which the numbers in the data set represent continuous quantities that can be divided into fractional parts, such as money or time. In these contexts, the mean does have a counterpart in physical reality. The discussion that emerges surrounding property E of the mean is surprisingly rich and was never a part of my classroom before I used these properties to guide my instructional decisions.

Focusing students' attention on property C, that the mean is influenced by values other than the mean, was my primary motivation for constructing parts c–f of this task. Students should develop understanding of how the mean of the ten numbers given in the data set is affected when various new numbers are added to the data set. The questions posed in this task can be approached from several directions. Students may actually compute the new means, then be encouraged to explore why the resulting changes in the means occurred. These questions also can be answered without direct computation of the new mean by using the power of property B. Adding a new number to the data set creates a new deviation from the mean that must be taken into account. The direction of change in the mean can be determined by identifying if the new deviation from the mean is positive, negative, or 0.

For instance, suppose that 8 is the new number added to the data set (see part d of **fig. 4**). The mean of the original ten numbers was 2.5. The difference between the new number being added to the set and the mean is found to be a positive number, $8 - 2.5 = 5.5$; thus, the sum of the deviations from the original mean in the new set is no longer 0. We can conclude, therefore, that the mean has changed when 8 is added to the set and that the mean must have increased to accommodate the new positive deviation. Students can easily confirm this conclusion by adding the eleven numbers in the new data set, dividing by 11, and determining that the mean of the new data set is 3. They can then verify that the sum of the deviations for the new set is 0 for the new mean of 3.

Similar reasoning can be used in parts (e) and (f) of this task. When 0 is added to the original data set, the difference between the new number, 0, and the mean is negative, $0 - 2.5 = -2.5$, and the mean decreases. Property B also explains why the mean does not change when the two new numbers, 2 and 3, are added to the original data set with mean 2.5. The mean does not change because the effects of the new positive deviation, $3 - 2.5 = 0.5$, and the new negative deviation, $2 - 2.5 = -0.5$, have a sum of 0, which means that they offset each other.

Parts (g) and (h) of the problem situation are open-ended and give students the chance to find numbers that can be added to the data set but do not change the mean. This task offers another opportunity for students to encounter property B because the sum of the deviations from the mean must be 0 if numbers are added to the set but the mean is unchanged. Thus, solutions to part (g) include pairs of numbers that are equally distant from the mean, with one number larger than the mean and the other number smaller than the mean. Note that students can find both integer and fraction solutions to part (g).

Finding three numbers that can be added to the data set and not change the mean is more challenging. As students experiment with various combinations of numbers, they will discover that one solution to part (h) is to add one number that is the mean itself, 2.5 in this problem, and two other numbers that are equally distant from the mean, such as 1 and 4. Another solution to part (h) consists of using two numbers above the mean and one number below the mean or vice versa. Using property B once again, we can determine that although multiple solutions to this problem are possible, finding a solution in which all three new numbers are integers is impossible.

In part (i), students are asked to consider the effect on the mean of adding an extreme value, or an outlier, to the data set. The mean of this new data set changes significantly and helps to draw students' attention to the conditions under which the mean is a "good" representative of the data set. Understanding that the mean value is representative of the data set is stated as property G, but discussion about property G must also include situations such as this one, in which the mean may not be the best representative measure. One way to prompt such a discussion is to compare the mean to other measures of center. For the given data set, the addition of an outlier does not affect the median; thus, the median may be considered a more representative measure for this data set. In solving parts (j) and (k) of this task, students are challenged to consider not only the relationship of the mean and the median to the numbers in the data set but also the relationship of these two statistical measures of center to each other.

## In Your Classroom

THE RESULTS OF STRAUSS AND Bichler's research (1988) serve as a useful framework to guide instruction in the concept of the arithmetic mean. I encourage you to examine the problems in the curriculum materials you use and the discussions held in your classroom with respect to both the seven properties defined by Strauss and Bichler and the Data Analysis and Probability Standard (NCTM 2000). In addition, you may encourage your students to use various technological tools as they explore and discuss these properties of the mean. For instance, performing computations and examining the mean and the sum of deviations from the mean can easily be done using a basic spreadsheet. This tool can be used for the tasks presented in this article and is particularly helpful when working with large data sets. These seven properties of the mean have been a valuable resource in helping me (and, in turn, helping my students) to understand what it means to not just find, but to meaningfully understand the mean.

## References


Bolster, L. Carey, et al. *Exploring Mathematics.* Glenview, Ill.: Scott Foresman and Co., 1994.

Hoffer, Alan R., Martin L. Johnson, Steven J. Leinwand, Richard D. Lodholz, Gary L. Musser, and Tina Thoburn. *Mathematics in Action*. New York: Macmillan/McGraw Hill School Publishing Co., 1994.

National Council of Teachers of Mathematics (NCTM). *Principles and Standards for School Mathematic*s. Reston, Va.: NCTM, 2000.

Strauss, Sidney, and Efraim Bichler. "The Development of Children's Concepts of the Arithmetic Average." *Journal for Research in Mathematics Education* 19 (January 1988): 64–80.

Zawojewski, 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, pp. 235–68. Reston, Va.: National Council of Teachers of Mathematics, 2000. □