the sum of the deviations about the mean will be 0, except for possible rounding. Since the deviations will always add to 0, they are not useful for quantifying variation.

Instead of examining the difference between each data value and the mean, students next examined the distance of each data value from the mean. The distance that each data value is from the mean is simply the absolute value of its deviation from the mean:

$$\text{Distance from mean} = |\text{deviation from mean}|$$

The absolute deviations for distribution 2 are displayed in **figure 6.**

The total, or sum, of the distances for all data values from the mean indicates, overall, how different the data values are from the mean of 5. For these eight distributions, comparing the total of the distances is acceptable because each distribution has nine data values. The variations between groups of data of different sizes might be compared by using the average distance of the data values from the mean instead of the total of the distances from the mean. The average, or mean, of the distances is obtained by dividing the total of the distances by 9, which leads to the formula

mean absolute deviation (or MAD)

$$= \frac{\text{total distance (of all values from the mean value)}}{\text{number of values}}.$$

Alternatively,

$$\text{MAD} = \frac{\text{sum } [|\text{deviations from mean}|]}{\text{number of values}}.$$

```
  -3
  -3        -1
  -3   -2   -1                       3    4         6
--+----+----+----+----+----+----+----+----+----+--
  2    3    4    5    6    7    8    9    10   11
                  Distribution 2
```

Fig. 5 Deviations from the mean; mean equals 5 for distribution 2.

```
  +3
  +3        +1
  +3   +2   +1                       3    4         6
--+----+----+----+----+----+----+----+----+----+--
  2    3    4    5    6    7    8    9    10   11
                  Distribution 2
```

Fig. 6 Absolute deviations from the mean; mean equals 5 for distribution 2.

For distribution 2,

$$\text{total distance (of all values from 5)} = 26$$

and

$$\text{MAD} = \frac{26}{9} \approx 2.89.$$

The MADs for the eight distributions in **figure 4** are given in **table 1.**

TABLE 1
MADs for Eight Distributions in Figure 4

Distribution No.	MAD
1	0.00
2	2.89
3	2.44
4	1.78
5	2.44
6	2.67
7	2.22
8	4.00

Interpreting the MAD

AFTER STUDENTS FIND THE MAD, IT IS IMPORtant for them to interpret the MAD in context. What does the MAD tell us about the data? The MAD is a gauge of "on average" how different the data values are from the mean value. The MAD for the nine data values in distribution 2 is 2.89. This number indicates that the values differ on average by approximately 2.89 units from the mean of 5. In context, the mean size for these nine families is five people, and the MAD indicates that the actual nine family sizes differ from 5, on average, by 2.89 (almost 3) people. Note that both the mean and the MAD quantities are expressed in the same measurement units as the original data.

A small MAD indicates that the data values are similar to the mean. A large MAD indicates that the data values are quite different from the mean. Several correct student interpretations of a small MAD follow:

- It doesn't deviate very much. It's much closer to the mean.
- The answer is close to the mean.
- There was not much difference between the other numbers on the chart from 5.

A large MAD indicates the following:

- It deviates more. It's further from the mean.
- The answer is far from the mean.
- There was a lot of difference between the other numbers and the average.

The ordering for the distributions selected by the students (4, 7, 3, 5, 2, 6) is close to being the same as the ordering indicated by the MAD. One problem is that distributions 3 and 5, which are different distributions, have the same MAD, 22/9, so students cannot distinguish one from the other by knowing both the mean and the MAD. Also, distribution 2, MAD = 26/9 ≈ 2.89, was placed before distribution 6, MAD = 24/9 ≈ 2.67; therefore, according to the MAD, these two distributions are out of order. Thus, on the basis of the MADs, two correct orderings for these distributions are possible: 4, 7, 3, 5, 6, 2 or 4, 7, 5, 3, 6, 2. Additionally, except for distributions 3 and 5, which have the same MAD, students can distinguish the other six distributions from one another by knowing the MAD.

Computing the MAD

USING LINE PLOTS WITH STICKY NOTES HELPS STUdents develop an understanding of how the MAD gauges variation in data from the mean. However, determining the MAD from this type of representation is not practical with most data sets. The following example presents a strategy for actually computing the MAD for data sets of reasonable sizes.

Consider the following nine family sizes: 3, 2, 4, 2, 9, 8, 2, 11, 4, which is distribution 2. The mean family size is 5. The deviations and absolute deviations from the mean for these data are displayed in **table 2.** The sums of the three columns in **table 2** are

$$\text{sum}\,[\text{data values}] = 45,$$

$$\text{sum}\,[\text{deviations}] = 0,$$

and

$$\text{sum}\,[\text{absolute deviations}] = 26.$$

Consequently,

$$\text{MAD} = \frac{26}{9} \approx 2.89.$$

TABLE 2
Computation Method for Finding MAD

DATA VALUE	DEVIATION FROM MEAN	ABSOLUTE DEVIATION
3	–2	2
2	–3	3
4	–1	1
2	–3	3
9	4	4
8	3	3
2	–3	3
11	6	6
4	–1	1
SUM 45	0	26

The Mean as a Measure of Center

THE MEAN AND THE MEDIAN ARE OFTEN REFERRED TO as "measures of center," the notion of center being that of some "halfway" point. The median is the value in the middle of the ordered data, for an odd number of measurements, or the average of the two middle values in the ordered data, for an even number of measurements. Approximately half the data values are below the median, and approximately half the data values are above the median. How is the mean a center? In distribution 4, both the median and the mean are 5. Four data values are below the 5, and four data values are above the 5. Additionally, in distribution 4, the total of the distances from the mean for all data values is 16. The total distance for the four values below the mean is 8, half of 16, and the total distance for the four values above the mean is 8, half of 16.

In distribution 2, six data values are below the mean and three data values are above the mean. So the median cannot be the same as the mean for this distribution. However, in distribution 2, the total of the distances from the mean for all data values is 26. The total of the distances for the six data values below the mean is 13, half of 26, and the total of the distances for the six data values above the mean is 13, half of 26. A check of the other distributions in **figure 4** supports the general rule,

total distance from mean for values below mean

= half of total distance from mean for all values

= total distance from mean for values above mean.

This result illustrates how the mean is a halfway point and is often called the *balance-point* interpretation of the mean. An excellent discussion of various interpretations of the mean can be found in Uccellini (1996).

Extension

THE ACTIVITY THAT I HAVE DESCRIBED OFFERS A learning experience for developing the statistical concept of variation in data from the mean. The objective is to develop student understanding of the notion of variation in data and the need to quantify the degree of variation in data from the mean. I have proposed the mean absolute deviation (MAD) as a gauge of the degree of variation from the mean and as a measure for distinguishing from one another those distributions with the same mean.

The MAD is not a commonly used quantity in statistics. Statisticians prefer to gauge the degree of variation from the mean with the variance and the standard deviation instead of the MAD. Zawojewski (1991) proposes that middle grade students "explore" and "compare and contrast" distributions by using the standard deviation. However, developing an understanding of the standard deviation may be too advanced for most middle-grades students. Where appropriate, this activity can easily be extended to develop students' understanding of the standard deviation as a measure of the degree of variation in data from the mean. This extension requires that students have knowledge of the Pythagorean theorem. The MAD is a much more natural measure of variation than the standard deviation; it is accessible to all students and should certainly be introduced before the standard deviation.

References

Friel, Susan N., Janice R. Mokros, and Susan Jo Russell. *Used Numbers: Middles, Means, and In-Betweens.* Palo Alto, Calif.: Dale Seymour Publications, 1992.

Uccellini, John C. "Teaching the Mean Meaningfully." *Mathematics Teaching in the Middle School* 2 (November–December 1996): 112–15.

Zawojewski, Judith S. *Dealing with Data and Chance. Curriculum and Evaluation Standards for School Mathematics* Addenda Series, Grades 5–8. Reston, Va.: National Council of Teachers of Mathematics, 1991.

Kader, Gary (March, 1999). Means and MADS. *Mathematics Teaching in the Middle School, 4* (6), 398–403.