mathematics and statistics into their proper roles—and can improve the teaching and learning of both.

Examples of Assessment Items

A critical look at a few proposed and used test questions may help clarify the difference between mathematical reasoning and statistical reasoning. All these examples are listed in the sources as questions on data analysis and are chosen in part because they are fairly typical of questions that find their way onto assessments in statistics. The comment sections provide a rationale on whether or not answering the question actually requires statistical thinking. If you are feeling adventurous, you might like to pause long enough after each test question to gauge your own reaction before reading the comments.

Example 1

The students in Mr. Kirby's class voted for their favorite book of the past three months.

The three books that they read were these:

Babe, the Gallant Pig

Sarah, Plain and Tall

Stone Fox

Here are some clues about how the vote came out.

a. 34 students voted.
b. The winner got the most votes but got fewer than half the votes.
c. There was a two-way tie for second place.

In the space below, make a chart or graph that shows voting results that fit all three clues.

(*Source:* Gawronski and Collins 2005, pp. 158–61.)

Comment

This is not a data analysis question, even though it does ask something about data. It is a question about mathematical reasoning (properties of numbers), not about statistical reasoning (using numbers in context—data—to answer a practical question of interest). The clues are given to challenge number sense and computation, not to direct students toward a deeper statistical reasoning about the data, how they were obtained, how they might be interpreted, and to what use they might be put.

Many, if not most, assessment items categorized as data analysis or statistics questions are really about computation and computational algorithms. The following three questions show a progression of this type of question.

Example 2

The average weight of 50 prize-winning tomatoes is 2.36 pounds. What is the combined weight, in pounds, of these 50 tomatoes?

A. 0.0472
B. 11.8
C. 52.36
D. 59
E. 118

(*Source:* NAEP Sample Questions; nces.ed.gov/nationsreportcard/)

Example 3

Joe had three test scores of 78, 76, and 74, while Mary had scores of 72, 82 and 74. How did Joe's average (mean) score compare with Mary's average (mean) score?

A. Joe's was one point higher.
B. Joe's was one point lower.
C. Both averages were the same.
D. Joe's was 2 points higher.
E. Joe's was 2 points lower.

(*Source:* TIMSS eighth-grade released items; timss.bc.edu/timss2003i/released.html)

Example 4

TIME CARD Name: J. Jasmine				Number of Hours	Average Hourly Wage	Total Daily Earnings
Mon.	10:00 a.m.	–	3:00 p.m.	5	5.50	27.50
Tues.	9:00 a.m.	–	4:00 p.m.	7	5.50	38.50
Wed.	3:00 p.m.	–	7:00 p.m.	4	5.75	23.00
Thurs.	2:00 p.m.	–	8:00 p.m.	6		
Fri.	5:00 p.m.	–	10:00 p.m.	5	6.00	30.00

a. According to the information above, what is the average hourly wage for Thursday's earnings if the total earnings for the five days was $153.50 ?

Answer:________________

b. The hourly wage rate changes at some hour during the day. At what time does the hourly wage rate change?

Answer:________________

Comment

Item 2 requires knowledge only about how a mean relates to a total; the context is incidental to the problem. (Does the student care that these are prize-winning tomatoes?) If a student can calculate one mean then that student can calculate two means, so item 3 adds nothing to the computational skill issue and certainly adds nothing to the statistical reasoning issue, even though the item may look more sophisticated than item 2. Again, the context is superficial and not germane. Item 4 is actually quite a good question on reading a table and making appropriate calculations, but, once again, this is a computation problem and not a problem about data analysis in a statistical-reasoning sense. When these types of items are used on an assessment, they should not take the place of questions that do, in fact, require statistical reasoning. But such questions are not easy to write, even among well-intentioned item developers, as the next example demonstrates.

Example 5

The table below shows the daily attendance at two movie theaters for 5 days and the mean (average) and the median attendance.

a. Which statistic, the mean or the median, would you use to describe the typical daily attendance for the 5 days at Theater A? Justify your answer.
b. Which statistic, the mean or the median, would you use to describe the typical daily attendance for the 5 days at Theater B? Justify your answer.

	Theater A	Theater B
Day 1	100	72
Day 2	87	97
Day 3	90	70
Day 4	10	71
Day 5	91	100
Mean	75.6	82
Median	90	72

(*Source:* NAEP Sample Questions, nces.ed.gov/nationsreportcard/)

Comment

A question that looks like it has the makings of a good statistics question is sabotaged by the required answer; the answer to part (a) is required to be "median" and the answer to part (b) is required to be "mean." This is a result of what could be sound statistical reasoning turned into an overly simplistic algorithm: when an outlier is present, use the median as the measure of center; when there are no outliers, use the mean. Appropriate statistical measures cannot be assigned to the data until the context is established with sufficient information to understand how and why the data were collected and what question needs to be answered with these data. If the question is to compare the two theaters, then the median should not be used for one and the mean for the other. If these are the same days (paired data), then the daily differences in attendance might be the appropriate measures to analyze. Shouldn't the 10 be investigated? Perhaps it was a mistake. In short, this turns out to be neither a good statistics question nor a good mathematics question, but it could become a good statistics question with a little more context and a little more leeway on the correct answer.

Example 6

The following table gives the times each girl has recorded for seven runnings of the 100-meter run this year. Only one girl may compete in the upcoming tournament. Which girl would you select for the tournament and why?

RACE #	1	2	3	4	5	6	7
Suzie	15.2	14.8	15.0	14.7	14.3	14.5	14.5
Tanisha	15.8	15.7	15.4	15.0	14.8	14.6	14.5
Dara	15.6	15.5	14.8	15.1	14.5	14.7	14.5

(*Source:* Burrill and Collins [2005, pp. 201–04])

Comment

This open-ended question explores ways to use real (or realistic, at least) data to answer an important practical question. It requires statistical thinking, since there are a number of plausible answers, no one of which is obviously "best" under all considerations. The statistical ideas clearly stand out, as the mathematics to be used here is elementary.

Example 7

Animal-waste lagoons and spray fields near aquatic environments may significantly degrade water quality and endanger health. The National Atmospheric Deposition Program has monitored the atmospheric ammonia at swine farms since 1978. The data on the swine population size (in thousands) and atmospheric ammonia (in parts per million) for one decade are given below.

a. Construct a scatterplot of these data.
b. The correlation coefficient for these data is .85. Interpret this value.
c. On the basis of the scatterplot in part (a) and the value of the correlation coefficient in part (b), does it appear that the amount of atmospheric ammonia is linearly related to the swine population size? Explain.
d. What percent of the variability in atmospheric ammonia can be explained

Year	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Swine Population	0.38	0.50	0.60	0.75	0.95	1.20	1.40	1.65	1.80	1.85
Atmospheric Ammonia	0.13	0.21	0.29	0.22	0.19	0.26	0.36	0.37	0.33	0.38

(*Source:* AP Statistics Exam, 2002, available at apcentral.collegeboard.com/)

Comment

The clear context, real data, and questions that require the answer to involve the context make this a sound question in statistical reasoning. Again, the mathematics necessary to answer the questions is elementary, but the statistical ideas are fairly deep. A complete answer to such an item might require that the student understand the difference between association (seeing a statistical relationship) and causality (which cannot be determined solely on the basis of observational data of this type).

Concluding Remarks

The use of "marriage" as a key word in the title of this article was not by accident. A strong marriage brings two people together not for the purpose of becoming one but so that the strengths of each can be validated, supported, and enhanced. Unconditional support, each for the other, fosters confidence for growth and creativity that neither person might achieve on his or her own. As a result, the whole is bigger than the sum if its parts. The same can be true of statistics and mathematics education, although the marriage analogy can be pushed too far. Each can support the other and both can become stronger as a result, but they should not be merged into one. With such mutual support, creative growth can emerge on both sides. And that is to the benefit of all.

The CD accompanying this yearbook includes an article from the *Mathematics Teacher* on reaching beyond computation when designing statistical tasks (Curcio and Artzt 1996).

REFERENCES

American Diploma Project. *Ready or Not: Creating a High School Diploma That Counts.* Washington, D.C.: Achieve, Inc., 2004. Available from www.achieve.org/achieve.nsf/ADP-Benchmarks-Samples?OpenForm, 2005.

Britz, Galen, Donald Emerling, Lynne Hare, Roger Hoerl, and Janice Shade. *Statistical Thinking.* Special Publication. Milwaukee, Wis.: American Society for Quality Control,

Burrill, John, and Anne M. Collins, eds. *Mathematics Assessment Sampler, Grades 6–8: Items Aligned with NCTM's "Principles and Standards."* Reston, Va.: National Council of Teachers of Mathematics, 2005.

Curcio, Frances R., and Alice F. Artzt. "Assessing Students' Ability to Analyze Data: Reading beyond Computation." *Mathematics Teacher* 89 (November 1996): 668–73.

Franklin, Christine, Gary Kader, Denise S. Mewborn, Jerry Moreno, Roxy Peck, Mike Perry, and Richard Scheaffer. "A Curriculum Framework for Pre-K–12 Statistics Education." Proposal presented to the American Statistical Association Board of Directors, March 2005. Available from it.stlawu.edu/~rlock/gaise, 2005.

Gawronski, Jane, and Anne M. Collins, eds. *Mathematics Assessment Sampler, Grades 3–5: Items Aligned with NCTM's "Principles and Standards."* Reston, Va.: National Council of Teachers of Mathematics, 2005.

National Commission on Excellence in Education. *A Nation at Risk.* Washington, D.C.: U. S. Government Printing Office, 1983. Available from www.ed.gov/pubs/NatAtRisk/index.html, 2005.

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

Salsburg, David. *The Lady Tasting Tea.* New York: W. H. Freeman & Co., 2001.

Steen, Lynn. *Achieving Quantitative Literacy.* MAA Notes #62. Washington, D.C.: Mathematics Association of America (MAA), 2004.

———, ed. *Mathematics and Democracy: The Case for Quantitative Literacy.* National Council on Education and the Disciplines. Princeton, N.J.: Woodrow Wilson Foundation, 2001.

Tukey, John. "The Future of Data Analysis." *Annals of Mathematical Statistics* 33 (1962): 1–67.

ADDITIONAL READING

Conference Board of the Mathematical Sciences. *The Mathematical Education of Teachers.* Providence, R.I.: American Mathematical Society, 2002

Freedman, David, Robert Pisani, and Roger Purves. *Statistics.* New York: W. W. Norton & Co., 1998.

Moore, David. *Statistics: Concepts and Controversies.* 5th ed. New York: W. H. Freeman & Co., 2001.

Moore, Thomas, ed. *Teaching Statistics: Resources for Undergraduate Instructors.* MAA Notes #52. Washington, D.C.: Mathematics Association of America (MAA), 2000.

Utts, Jessica. *Seeing through Statistics.* 2nd ed. Belmont, Calif.: Duxbury Press, 1999.