# The Statistics Teacher Network

Number 44 — ASA/NCTM Joint Committee on the Curriculum in Statistics and Probability — Winter 1997


*Book Review*

## *Activity-Based Statistics*

### *Scheaffer, Gnanadesikan, Watkins, & Witmer (1996)*


***ISBN 0-387-94598-9 (Student) $20***

***ISBN 0-387-94597-0 (Instructor) $43***

***Springer-Verlag, 1-800-springer***

***Jones and Bartlett, 1-800-832-0034***


According to one paradigm—characterized variously as "established," "entrenched," "time-honored," "archaic," etc., depending on the characterizer's personal biases—students learn by absorbing information. Our job as teachers would then be to present knowledge at an optimal rate, given the abilities of our particular students. According to the reform paradigm, however, students learn by "constructing" their own knowledge. Our job as teachers would then be to create and maintain an environment in which students "discover" the concepts afresh through active participation—individually or in groups.

Some will argue that these reformers are throwing out the baby with the bath water, that the lecture format's virtues will be lost if we forsake the traditional approach. Others will argue that the lecture approach puts "coverage" of material ahead of student learning. How one sits on this debate determines to a great extent which texts one uses. So those sympathetic to the reform message will immediately embrace *Activity-Based Statistics*, as the authors are clearly committed to the reform approach and philosophy. Yet this resource is so well-conceived and structured that even many of those still committed to the lecture approach will consider using these activities in class.

The philosophy that pervades the book is that statistics "should be taught more as an experimental science and less as traditional mathematics." Skeptics of the new paradigm will be delighted that there is still plenty of structure in a statistics "lab," just as there is in other sciences. Students should learn that all statistics problems are motivated by a problem in a real-world context. So each activity begins with a description of the activity's "scenario." Immediately after this introduction is a concise question (in context) to focus students' attention, followed by a brief overview of the pedagogical objectives for the activity. The bulk of each activity is a meticulously outlined description of various tasks and questions that help students "discover" for themselves the activity's main concepts. Although this includes several "closed" questions with "correct" answers, students are forced to make links and generalizations and speculations on their own through lots of carefully chosen open-ended questions as well. Finally, each activity ends with a couple "wrap-up" questions and optional "extension" activities to check whether they have indeed mastered the concepts and can apply them in related situations.

The range of concepts illustrated by the activities spans nearly everything most people cover in an intro-level stat course—data production, exploratory data analysis, and inference—and then some: capture/recapture, bootstrap methods, randomization tests, and an entire unit on the nature of statistical modeling. Students are spared a gory (and unnecessary!) exposition on probability, complete with a for-

## Also In this Issue...

mal exposition of conditional probability and various probability "rules." Instead, they are treated to a series of activities that drive home the basic ideas about random behavior and sampling distributions with a minimum of mathematical and symbolic overhead. Similarly, the authors steer fairly clear of the formal aspects of hypothesis testing and devote more time to fundamental issues of estimation. And although students aren't exposed to a wide variety of design techniques, the included activities illustrate the fundamental design ideas far better than the scant sections on data production in most introductory texts.

Helpful notes for how to use these activities are compiled in the hardback companion *Instructor Resources*. In this guide, teachers will find not just lists of materials required for the activity and the prerequisite concepts and techniques, but also detailed procedures for how to initiate the activity session. The authors include sample results from their own experience—for the activities themselves, and for the wrap-up questions and extension exercises—to give instructors a feel for what to expect on the first attempt. There are also sample assessment questions for each activity, as well as a separate chapter on the kinds of assessment appropriate for the active learning approach in general.

Although the book's scope is comprehensive enough that a teacher more comfortable with the "constructivist" approach can easily build an entire course around this book as a primary text, the materials can alternatively be used as supplements to a more traditional course. I suspect that many reform-wary instructors will warm up to the new approach quickly, and will want to try more and more of the activities as they and the students become acclimated to the change in environment. But regardless of whether this book ends up converting the "old guard" into the avant-garde, many stat courses will soon be much more activity-enriched than ever before, thanks to the authors of *Activity-Based Statistics*.


**Reviewed by Doug Andrews**
**Wittenberg University**
**Springfield, Ohio**


Remember to request your application forms for the 1997 Poster and Project Competition. Forms are available from Sue Kulesher at the ASA office. Call 703-684-1221 or E-mail sue@amstat.org. The deadline for entries is April 15, 1997.

*Statistics in the Classroom*

## A Quick Introduction to the Statistical Capabilities of the TI-83 Calculator

**Disclaimer:** *I have no affiliation with Texas Instruments. There are other calculators in the world that perform many of the same functions as the TI-83. The TI-83 seems to be easy to use and Texas Instruments calculators are used by many school districts and colleges around the country.*

**Introduction to the TI-83**

I have a confession to make. I bought a TI-82 calculator soon after they became available, but I really didn't learn how to use it. It didn't change the way I taught Introductory Statistics. The middle school and high school teachers I met at the time were excited about the mathematical and graphing capabilities, but the statistical capabilities of the TI-82 were essentially limited to descriptive statistics and graphs.

Then I heard about the TI-83. I heard that it would be more statistical—that it would build on the capabilities of the TI-82 and would include tables of statistical distributions, tests, confidence intervals, and some financial functions as well. I borrowed a TI-83 from Texas Instruments last spring, and I am now the proud owner of two TI-83s. The TI-83 has the potential to change the way we teach introductory statistics.

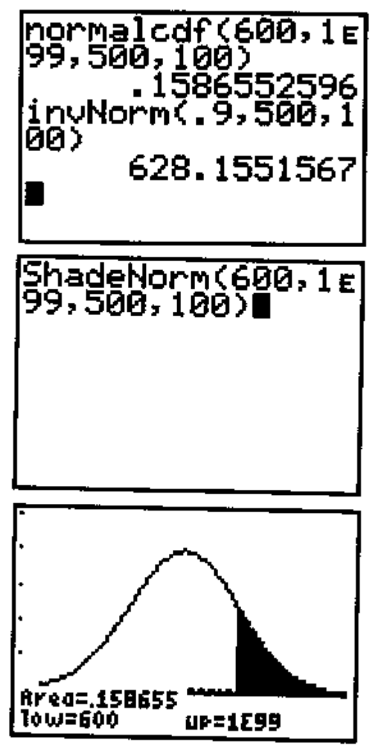


**Figure 1**

**Probability Distributions on the TI-83**

Here are some examples of the basic statistical features that make the TI-83 so exciting.

Suppose that the mean score on the

Mathematics portion of the SAT is 500 with a standard deviation of 100. The probability that a single randomly selected student will score above 500 is about 0.16. The normalcdf(lowerbound, upperbound,mu,sigma) function makes this calculation easy. The ShadeNorm(lowerbound,upperbound,mu,sigma) function will even draw a picture of the answer. See Figure 1 for the calculator commands and output for this example. The invNorm(area,mu,sigma) can help to answer questions like this: How high must a student score on the Math portion of the SAT to be in the top 10%? The calculator reports that the answer is about 628.

The TI-83 can compute probabilities for discrete distributions as well. Suppose that 20 students at a small high school will be taking the SAT this year. Assuming their SAT results are independent, the number of them that will score above 600 on the SAT Math test follows a Binomial distribution with parameters n=20 and p=0.16.

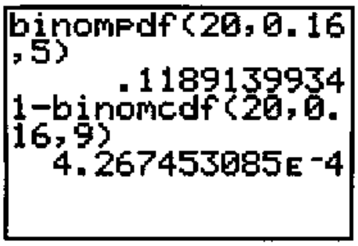


**Figure 2**

The TI-83 reports that the probability that exactly 5 of the 20 will score 600 or above on the SAT Math test is binompdf(20,0.16,5)=0.1189. The probability that 10 or more will score above 600 on the Math portion is 1 - binomcdf(20,0.16,9) = 0.0004. (See Figure 2).

The TI-83 also contains built-in functions for the Student's t, chi-square, F, Poisson, and geometric probability distributions. Soon there will be little need for tables in the backs of statistics textbooks! Hooray!

**Storing Data on the TI-83**

Like the TI-82, the TI-83 accepts data into lists. The list structure is more flexible on the TI-83 for several reasons: the lists can be longer; one can enter and store more than six lists at a time; and the lists can be named.

If the data are entered into a program, then running the program can set up data in the list editor, complete with column names, and create plots and analyses as well. Thanks to Al Coons for sharing this suggestion for storing datasets in programs for easy sharing and retrieval, and for more efficient storage. Figure 3 contains a data storage program for the milk production data, which will be discussed in the next section.

**Statistical Inference with the TI-83**

Consider a hypothetical study of the effect of a new hormone on milk production in cows. Suppose that twenty calves were randomly divided into two groups of ten. The two groups were placed on identical diets, but the treatment group received the hormone. The ten cows in the control group did not receive the hormone.

After the calves have grown into milk producing cows, their production is measured for a one day period. Here are the data collected on that one day, measured in pounds of milk produced.

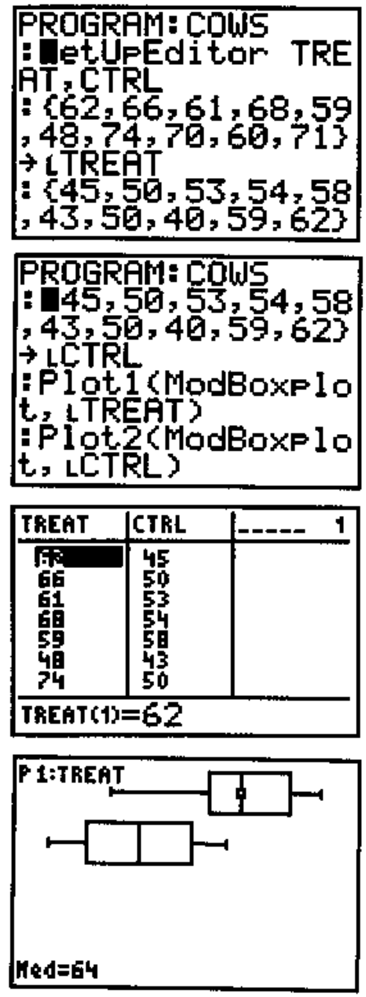


**Figure 3**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Treatment Group | 62 | 66 | 61 | 68 | 59 | 48 | 74 | 70 | 60 | 71 |
| Control Group | 45 | 50 | 53 | 54 | 58 | 43 | 50 | 40 | 59 | 62 |

The TI-83 can create side by side boxplots of the data, but so could the TI-82. The TI-83 can also perform statistical tests. In this case, a two-sample T-test seems to be appropriate.

First, the calculator is asked to perform a two-Sample F-test for the equality of the variances using the command 2-SampFTest. After determining that the variances seem to be

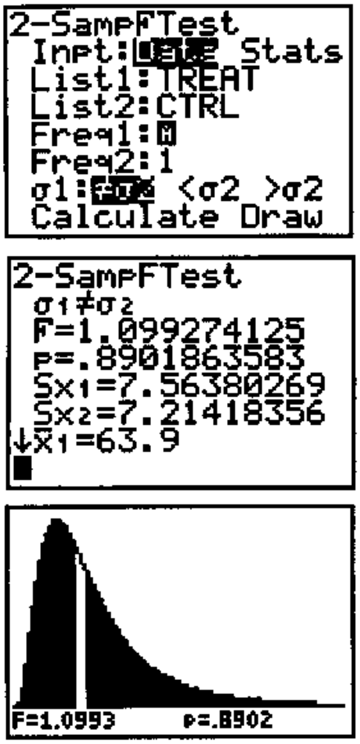


**Figure 4**

equal, (see Figure 4), a Pooled 2-SampTTest is performed using the 2-Sample T-Test function. The t-test tests the equality of the average milk production levels for cows in the treatment group compared to the control group. The test reveals that there is a statistically significant difference in the mean milk production levels of the two groups. The p-value for the test is 0.0014. Figure 5 contains the results of the two-sample t-test.

After a significant difference is found, the size of the difference can be estimated with a confidence interval. The TI-83 function 2-SampTInt produces a confidence interval for the difference between the means. As shown in Figure 5, a 95% confidence interval for the difference in the group means is (5.56, 19.44) pounds of milk per day.

The TI-83 has other statistical capabilities, corresponding to the topics contained in a typical full year introductory statistics course. The built-in functions include one variable regressions and one way analysis of variance, and stop just short of multiple regression analysis.

**The TI-Graph Link and Calculator-Based Lab**

For about $50.00 one can purchase a TI-Graph Link cable and software that connects TI graphing calculators to a personal computer. The Graph Link allows printing of calculator screens, storage and retrieval of data and programs on the computer, and program creation on the computer as well.

The Graph Link software can generate program files that can be stored on the World Wide Web. Educators around the world can then either print and enter the programs into calculators, or, with a Graph Link of their own, link programs from the Web directly into the calculator. A number of programs for the TI-83 are already available at the following Texas Instruments site: ftp://ftp.ti.com/pub/graph-ti/calc-apps/83/

I don't own a Calculator-Based Laboratory (CBL) system, although I have seen a demonstration of one. The CBL allows for automated data collection and storage, and can provide fascinating interactive examples in the classroom.

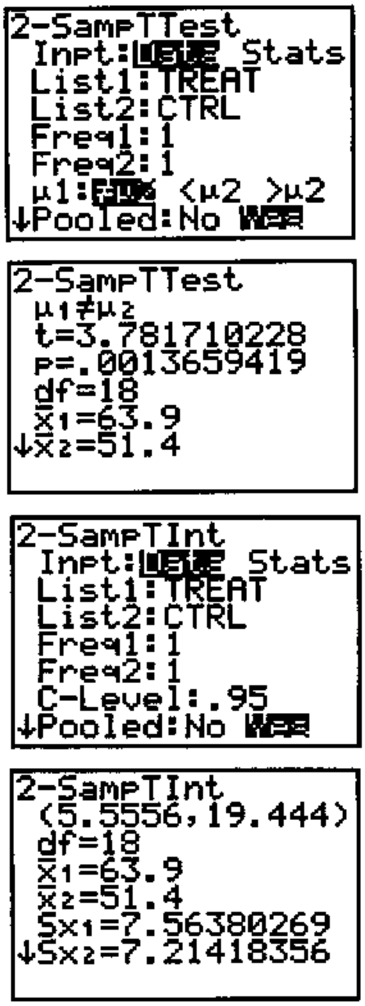


**Figure 5**

**Implications**

With the statistical tools necessary for an entire first course, and with the storage and retrieval capabilities using the Graph Link, the TI-83 provides a relatively inexpensive personal computing device for students and teachers. It is my belief that tables of distributions that have lived in the backs of statistical textbooks for more than half a century are doomed to extinction. Calculators like the TI-83 will reduce the computational burden for learners and allow teachers to focus on statistical concepts such as data collection, display, analysis, and interpretation.

**Tom Short**
**Villanova University**
**Villanova, Pennsylvania**

*Book Review*

## *Graphing Statistics & Data: Creating Better Charts*

**by Anders Wallgren, et al.**
**Sage Publications, 1996**
**ISBN 0-7619-0599-5**
**94 pages, paperback, $17.95**

The subtitle of the book is appropriate. This book is aimed at anyone who wishes to learn how to create better charts—from professional statisticians to students and teachers. No formal training in statistics is required, but some elementary understanding of the basic principles of statistics would be helpful (I repeat, helpful, not required). This is not a textbook—as no exercises or classroom activities are included. However, it does provide guidelines and principles for creating effective and easy-to-read charts, and more than 250 examples are included that illustrate different types of charts and different ways of charting the same data, which illustrate both what the authors consider suitable and unsuitable charts.

The book focuses on charts that might appear in a report, book, or newspaper, or that might be used in a presentation. About one-half of the book deals with different types of charts: bar charts, frequency plots, time series plots, scatter plots, boxplots, flow charts, and statistical maps. The other half deals with topics such as choosing chart types, the building blocks of charts (text in charts, patterns, shading, and color), chart philosophy, chart layout, and charts in practice. The book also covers many of the "refinements" of graphics software that often allow us almost unlimited opportunities for producing unsuitable charts. A number of the smaller data sets are listed so the readers can use their favorite software to try to reproduce the charts.

Chapters 1–3 deal with general concepts and principles. Chapter 1, "The Power of charts," is a short introductory chapter covering a little of the why, when, and how to use charts (in particular, charts versus tables). The authors state that "As chart constructors we must find a balance between details and the whole." Chapter 2, "Choosing chart types," briefly describes the basic chart types and considers how the structure of the data and the main purpose of the chart can assist us in choosing the best type of chart for a particular situation. Chapter 3, "The building blocks of charts," discusses building charts from the basic elements of lines, areas, and text. It clearly defines the building blocks of charts-plot area and chart area, framing and grid lines, titles and other text, curves, areas, line and area patterns, color, legends, labels, axes, axis scales, and even axis ticks–and suggests how to format each of them so that the finished chart is effective and easy to read.

Chapters 4–10 deal with different types of charts. Chapter 4, "Bar charts—our basic chart," covers vertical and horizontal bar charts, grouped bar charts, and stacked bar charts. It discusses the choice between grouped and stacked bar charts and topics such as the gaps between bars, gridlines, the order of the bars, overlapping bars, and negative values in bar charts. A brief description of a dot chart is also included. (There is more on dot charts in later chapters.) Chapter 5, "Showing frequencies," focuses mostly on showing distributions of continuous variables. Types of charts include histograms, population pyramids, frequency polygons, stem-and-leaf charts, and ogives. The chapter also covers pie charts for illustrating percentage distributions of qualitative variables and as an alternative to bar charts. Topics include class size, comparing several sets of continuous data, pie charts for several groups, and exploding a sector of a pie chart.

Chapter 6, "Showing development over time," is a little longer than the previous chapters. It covers basic types of time series charts plus index charts, semi-logarithmic charts, and accumulated line charts (area charts). There is also discussion of topics such as the appropriate proportions of a time series chart (the time axis versus the y-axis), should zero be included on the y-axis, showing a time series with periodical data, and showing several series in the same chart. Chapter 7, "Showing relationships," illustrates how scatterplots can be used to show the relationship between two quantitative variables and discusses such topics as aggregated data, the relative proportions of the axes, the risk of optical illusions, and scatterplots with several x-variables. Then bar charts and dot charts are discussed again as tools for showing relationships when the x-variables are qualitative. Finally, the use of barometer chars, a kind of bar- and dot-chart hybrid for showing y-differences for x-variables with two classes, is described.

Chapter 8, "Showing variation," covers the use of boxplots and Lorenz charts as alternatives to histograms for showing variation and as tools to use to compare variation among different sub-groups. The Lorenz chart uses the ranked observations to compare two cumulative distributions. Chapter 9, "Showing flows," covers flow charts, although not just the typical kind of sequential process flow charts with which we are most familiar. It discusses input-output flow diagrams that can be used to show such relationships as sources of income versus expenditures or energy type versus the usage of each type, or flows between geographic areas.

Chapter 10, "Showing geographical variation," covers various types of maps in some detail. It includes the following: basic statistical maps for showing location-specific information; choropleth maps or shaded maps for showing ratios such as proportions or intensities within some administrative subdivisions; isopleth maps where, instead of administrative boundaries, the characteristics of the statistical variable determine the boundaries directly; density maps; and cartograms, a cross between charts and maps where the map provides a background map for the identification of geographical areas and small charts or symbols are drawn in the various areas.

Chapter 11, "Some chart philosophy," discusses some of the fundamental ideas behind the "art" of designing good charts. It covers in detail, and with many examples, how to convey the intended message of the chart in an effective way. Subheadings and topics include the following: subject matter and target group—using information about the subject matter (in particular the comparisons to be made) and the extent of the target group's knowledge of the subject to produce an effective chart; the design and function of charts—letting the data dominate, data density, the degree of clarity, and making true comparisons; and about perception-optical illusions and how easy is it for the eye to make different types of comparisons. This last section discusses some of William Cleveland's results on making different graphical comparisons.

Chapter 12, "Charts and layout," briefly discusses the design and relative positioning of the chart relative to the main body in which it will appear. Chapter 13, "Charts in practice," applies the principles given in the earlier parts of the book to several realistic examples. Topics revisited include patterns and order in bar charts, a bar chart with an extra dimension, the zero on the y-axis, time series (index charts and periodical data), and dot charts for situations with many comparisons. A final section discusses some amusing news graphics taken from Tufte's book *The Visual Display of Quantitative Information.*

Chapter 14, "Check-list," provides a simple check list for before you start and when you think you have finished. The last chapter, Chapter 15, "Choice of technology for producing charts," briefly discusses the technology used in the production of the book itself.

Note that this book was translated from Swedish. Some examples were revised using data from the U.S.A., but most of the examples from the Swedish edition were kept. However, these do not detract from the readability of the book. In summary, this book provides sound advice on the basics of drawing charts, and is written in an easily digestible style that should make it beneficial to a wide audience. It does not attempt to be as entertaining as Tufte's books, or as innovative as Cleveland's; but, it is enjoyable reading and contains much practical, common-sense, advise on chart building. I highly recommend it. (Note: there is also a hardback version, ISBN 0-7619-0598-7, $45.)

**Reviewed by John Schollenberger**
**Cleveland Chapter, ASA**

*Statistics in the Classroom*

## A Statistician Goes to First Grade

Recently I "volunteered" to teach some special lessons to my daughter's first grade class on data analysis: what is data, how can students collect it, and how can they analyze it using simple bar charts. It was a lot of fun and very rewarding because the children loved it and it gave me some very positive feedback. Also, it was a good way for me to partake in the efforts of ASA's Quantitative Literacy program. It had the added benefit that it gave me a real good foot in the door of my child's classroom. I thoroughly enjoyed the experience and would like to share how I went about it and what we did.

My daughter, Alexandra, is in the first grade and her teacher has parents volunteer to help in the classroom. The first time that I was in the classroom I spent the hour or so cutting apples and making apple sauce; the second time I spent the hour weaving strips of construction paper. Then I suggested that I might be able to do a math activity since I am a statistician. The teacher agreed and we made a six-week commitment (to be canceled if the first session bombed) and set up a regular schedule of about 45 minutes per week.

I started with the Used Numbers series from Dale Seymour Publications which is particularly good since each book is directed towards two grade levels (e.g., the first book is for first and second grades). It gives the reader many different approaches towards the same activity, broken up into suggested 45 minutes worth of activity. (It's amazing how little you get accomplished with a group of first graders in 45 minutes!) Also, it gives sample dialogues between teacher and students, which will give you some idea of how to motivate the students. One thing I must say is that teaching graphing to first graders is nothing like lecturing at a statistics conference!

In our first couple of sessions we talked about how to count the number of children in the classroom, emphasizing methods for counting accurately, how to sort out discrepancies in counts and how to check if the answer is correct. It was heart warming when the students moaned when the lesson was over because it was "too short."

Our second activity was to collect the ages of the children in the classroom and graph them in bar charts. We first produced one large chart with the students sticking on a dot to indicate where their own ages belonged. Then each of the students made his or her own individual chart of the ages of the children in the class. On another day the children drew self portraits on Post-It notes and put the notes in the appropriate places to build a large bar chart. The discussion included hypothesizing how the chart might look one month from the date or a year from that date.

One gratifying thing that I noticed was that the teacher was picking up on what I was doing and started including the collecting and charting of data in her lessons and homework assignments. For example, she showed me a project that she had the class produce for presentation to the Board of Education: it was a large bar chart showing the year each student's ancestors arrived in America!

Then we had the students collect their own data. They were given data collection sheets and asked to collect the ages of their siblings and parents. To graph these data we used five different colors of Post-It notes (one for them, their brothers, sisters, mothers, and fathers) and produced a 30-foot-long bar graph of the ages of their family members. In order to put this large graph together, we needed to hang the paper in the school hallway and work on it and discuss it there.

This drew the attention of several other teachers and classes who were intrigued by what we were doing. Many mathematical concepts could be discussed by comparing the positions of the colored notes. In fact, we spent nearly a whole class with my posing questions that the students could answer by looking at the bar graph, and then having the students raise their own questions.

Just before Christmas our "treat" lesson involved counting the number of M&Ms in the individual packs and displaying the counts in graphs along with individual graphs of the distribution of the colors of the M&Ms in the packs. This was one of their favorite lessons.

If you haven't guessed by now, things were going so well that the teacher asked me if I'd be willing to extend my commitment past the six weeks originally agreed upon.

The final project was a large one that lasted over a month and produced some interesting results and charts that showed many different concepts about the distribution of data. The data involved estimating and counting handfuls of Fruit Loops (a popular cereal) taken by the first graders. Each child took a handful and placed them in a paper cup. Then they guessed how many there were. They took their cups into a third grade classroom and had the third graders guess, and then did the same with sixth graders, and finally counted how many were in their cup. The data were collected on data sheets I prepared, and we started graphing the results. We produced bar graphs for each of the three groups of guesses and also for the counts, which were color coded by sex. The guesses piled up around the low teens for the first graders, in the high teens for third graders, and in the twenties for the sixth graders. The counts were a uniform distribution between 16 and 50. (The color scheme of the counts showed that the boys, in general, took larger handfuls, and this led to an boys had larger hands or the boys were hungrier.) Since we couldn't measure the hunger levels, we chose to measure hands; sure enough, the boys (and the two girls who took more than the other) had the larger hands in the classroom. Lastly, we tallied three comparisons of who guessed closer to the count, comparing first graders versus third graders, first versus sixth, and third versus sixth, and it was clear that the sixth graders were the best guessers—not too surprising, but the first graders found it fascinating.

We summarized the findings in a large poster that was displayed for the whole student body to see. We considered submitting this to the

Poster Competition sponsored by ASA's Statistical Graphics Section but were not able to since the rules specify that the poster needs to be produced by a maximum of four students, and the whole class participated in producing this one.

I feel that teaching data analysis in my daughter's classroom was very beneficial both to me and to the students and teachers. It influenced the teaching of one classroom teacher and, in fact, the other teachers in the school were impressed by what we had done and started doing similar activities. It was fun and interesting for me and gave me a meaningful way to be involved in my daughter's schooling.

I must relay one additional vignette. While I was putting my daughter to bed, she asked me why Miss Robin told the class that I was doing math. I answered "because I am" and asked her why she was asking. She said, "you can't be doing math because what you are doing is FUN!" Certainly, "math is fun" is a concept we want all children to have, even from the earliest ages.

**Lorraine Denby**
**AT&T Bell Laboratories**

**Editor's Note:** *Lorraine wrote this article for* Amstat News *in May 1993. She permitted me to include it here to encourage statisticians and teachers to replicate her experiment.*


## Keep Us Informed...

*The Statistics Teacher Network* is a newsletter published three times a year by the American Statistical Association—National Council of Teachers of Mathematics Joint Committee on the Curriculum in Statistics and Probability.

We need your letters, announcements, articles, and information about what is happening in statistics education! Please send hard copy, and, if possible, a disk written in standard ASCII text to the editor:

**Jerry Moreno, Dept. of Mathematics, John Carroll University, University Heights, OH 44118**

or **moreno@jcvaxa.jcu.edu**

or **fax: (216) 397-3033**

**Production Editor:** *Kristen Fernekes, American Statistical Association*

**To be added to the mailing list or make an address change,** please send your name and address to: *Statistics Teacher Network,* c/o American Statistical Association, 1429 Duke St., Alexandria, VA 22314-3415; (703) 684-1221; fax: (703) 684-2037; E-mail: veronica@amstat.org

Printed in the U.S.A



**Statistics Teacher Network**
**c/o American Statistical Association**
**1429 Duke Street**
**Alexandria, VA 22314-3415**
**USA**

Non-Profit
Organization
U.S. Postage
PAID
Permit No. 361
Alexandria, VA

*Let us know your Zip +4!*