Teaching Statistics

George Cobb
MOUNT HOLYOKE COLLEGE

Introduction

This report on teaching statistics will present the Statistics Focus Group's recommendations under three headings, corresponding to *statistics* ("Recent Changes in the Field"), *mathematics* ("Some Differences Between Mathematics and Statistics"), and *teaching* ("What Research Tells Us"). A fourth section ("Examples") illustrates ways these recommendations can be put into practice, and a final section ("Making It Happen") offers two meta-recommendations about implementation.

> What do we want students to be able to *do*, themselves, in terms of performing statistical work after their course is completed? What kinds of statistical reasoning, or arguments, do we want them to be able to *understand*? What kinds of *experiences* should the students have had in the course?
>
> —*Jim Landwehr, AT&T Bell Labs*

Our group made a deliberate decision not to prescribe lists of topics for particular courses, but instead to seek a general intellectual framework within which we and others can fit a great variety of courses. (For one list of recommended topics, see Appendix D: Report on a Workshop on Statistical Education, by Robert Hogg, which presents a consensus view of thirty-nine statisticians on problems in statistical education and suggestions for reform.)

We intend our three recommendations to apply quite broadly.

> One introductory course may differ from another according to many factors: calculus prerequisite versus no calculus; engineering, technical audience versus arts, nontechnical audience; goal of understanding versus goal of doing; taught by mathematics or statistics department versus taught by user department; large research university versus small college; large clientele (100s–1000s) versus small clientele (less than 100); required course versus elective course; students bright, intellectually curious versus students dull, passive; PC's readily available versus computer facilities inadequate.
>
> —*Howard Taylor, University of Delaware*

> If I use applied regression as the vehicle to give students the experience they need and you use time series forecasting, that's fine. What matters most is the experience with practical reasoning about data.
>
> —*David Moore, Purdue*

Although our group spent most of its time thinking about several versions of the "standard" introductory course, several of us know from experience that the spirit of what we urge can infuse courses devoted to experimental design and analysis of variance, sample survey design, regression, time series, or multivariate analysis, to name just five. At Swarthmore, Gudmund Iversen teaches a pair of introductory courses ("Statistical Methods" and "Statistical Thinking") which are in some ways quite dissimilar, but the kind of general goals we shall describe are appropriate for both. At Mount Holyoke, the mathematics department offers two beginning statistics courses, one on experimental design and applied analysis of variance, with emphasis on experimental studies in biology, psychology, and medicine, and another on applied regression, with emphasis on observational studies in public policy, economics, etc. At St. Lawrence University, Robin Lock teaches an introductory course on applied time series. At University of Delaware, Howard Taylor teaches beginning statistics to engineers. At SUNY Stony Brook, Judy Tanur teaches statistics to sociologists, and does a lot with surveys. The goals we advocate are suitable for all these courses.

We have also sought to make our three recommendations relatively independent of level of preparation, in that we imagine them applying to courses designed for students at any band within an entire spectrum of technical facility, ranging anywhere from ultra-nimble to infra-numb. Level of preparation will determine how far or how deep a course can go, but need not determine the general direction; it may limit the degree of subtlety and sophistication with which students can come to understand random variability and it's underlying predictable patterns, but needn't prevent them from making a meaningful start.

Recent Changes in the Field of Statistics

> Statistics has moved somewhat away from mathematics back toward its roots in scientific inference and the analysis of data. ...The most important driving force in this shift of emphasis is the computer revolution.
>
> —*David Moore*

During the last two decades, statistics has been changing simultaneously on three levels, which correspond to technique, practice, and theory. On the technical level, cheap, powerful computing has made possible a number of important innovations: graphical methods for data display, iterative methods for data description, diagnostic tools for assessment of fit between data and model, and new methods of inference based on resampling techniques such as the bootstrap. On the level of practice, such things as pattern-searching, model-free description, and systematic assessment of fit have all become more prominent, at the expense of formal inference, most especially hypothesis testing. Statisticians now put more effort into the complex process of choosing suitable models, less effort into doing those things—simpler by comparison—which take the choice of model as given.

> How many of us still concentrate on hypothesis testing, and even within that narrow and questionable context permit students to finish an exercise in number crunching to a pre-specified *t*-test, with the sole application of the English language being the words "reject the null hypothesis?"
>
> —*Walt Pirie, VPI & SU*

> Mathematicians who teach the introductory course probably will be completely oblivious to the decreasing importance of hypothesis testing in the work of statisticians. To mathematicians, this may have the most profound implications for their introductory course because it calls into question the ultimate goals of the course.
>
> —*Ann Watkins, California State at Northridge*

On the level of theory, one can distinguish two kinds of changes that invigorate discussions about the reasoning of statistics. First, foundational discussions of long standing (should unknown parameters be regarded as fixed or as values of random variables? Should inferences be conditional or unconditional?) now much more often take place in the context of real applications. Second, statistical practice has partly outgrown its mathematical theories, which are consequently less relevant. Important new elements of data analysis (model-choosing, model-checking, and model-free description) don't fit the older theoretical frames, while the influential area of statistical process control offers new ways, not yet mathematically developed, to frame the enterprise of learning from data.

> [Alternatively,] a coherent theme built around modern information processing can encompass the general concepts of EDA (exploratory data analysis), graphics, and computing as they relate to the introductory course in statistics.
>
> —*Dick Gunst, Southern Methodist*

A deeper and more detailed treatment of the recent changes in statistics may be found in David C. Hoaglin and David S. Moore (Eds.), *Perspectives in Contemporary Statistics,* MAA, 1992. Suggestions for additional reading can be found at the end of this report.

What does all this mean for the teaching of statistics? We offer the following recommendation:

Recommendation I: EMPHASIZE STATISTICAL THINKING

Any introductory course should take as its main goal helping students to learn the basic elements of statistical thinking; many advanced courses would be improved by a more explicit emphasis on those same basic elements:

1. *The need for data.*

> To recognize in one's own citizenship, the need to base personal decisions and actions on evidence (data), and the dangers inherent in acting on assumptions not supported by evidence. This doesn't even necessarily invoke the concept of randomness, but is nevertheless inherent in statistics as a discipline.
>
> —*Walt Pirie*

2. *The importance of data production.*

> It is very difficult and time-consuming to formulate problems and to get data that are of good quality and really deal with the right questions. Data generally don't represent what people initially think. Moreover, most people don't seem to realize that this is the way things work out until they go through this experience themselves. This is the most important part of actually doing statistics, because if it is not done well all the subsequent analysis can't be worth much. ... Most people I deal with would be better off if they carried realistic notions about formulating problems and getting relevant and accurate data, rather than some vague notion of significance or confidence from the course they had.
>
> —*Jim Landwehr*

> I haven't had students plan studies and gather data in a couple of years, for various reasons. I wasn't very happy with the actual projects during the four semesters I did it. However, I was quite happy with the students' *experiences.* Almost every one of them, by the time they finished, was rather sheepish about what a poor study it turned out to be, because they could see all the ways it really should be improved.
>
> —*Mary Parker, Austin CC*

> Let's not forget that the most desirable process is usually (always?) starting with the model, in the sense of "design before data." In my experience the most frequent cause of poor (or even failed) experimentation is the lack of good prior design.
>
> —*Walt Pirie*

3. *The omnipresence of variability.*

Variability is ubiquitous. It is the essence of statistics as a discipline ... it is not best understood by lecture—it must be experienced.

—*Dick Gunst*

Some data have variability due to measurement error, where other data have variability due to the fact that the phenomenon isn't completely deterministic. As mathematicians, we notice that the same models can be used to analyze these two different kinds of variability, so we tend to identify them with each other. I think that students sense that these are different, and this is part of why they have a problem with the idea of a random variable. I had my students measure the length of the building by stepping it off, to illustrate that many measures are approximate even if we report them as exact. The scores for the whole class formed a normal distribution, with a few outliers, and we could see why we needed to discard them. In the same assignment, I asked each student to measure the heights of three females. These scores for the whole class also formed a normal distribution. We could see why it wasn't appropriate to discard outliers this time.

—*Mary Parker*

4. *The quantification and explanation of variability.*

   a. Randomness and distributions.
   b. Patterns and deviations (fit and residual).
   c. Mathematical models for patterns.
   d. Model-data dialogue (diagnostics).

—*David Moore*

The development of the ***history*** of statistical thinking should also be included in any general education statistics course. (And not in a separate unit, but infused throughout.) Part of the reason that many instructors resist changes in the introductory statistics course is that they don't understand how relatively young the subject is and that evolution can be expected.

—*Ann Watkins*

Just as there was mathematics before Euclid and Archimedes, there was statistics before Karl Pearson and Ronald Fisher, but comparing the times when these four lived gives a sense of how new a subject statistics is. In the context of two millenia of changes within mathematics, essentially *all* changes in statistics are recent changes.

—*George Cobb, Mount Holyoke*

Some Pertinent Differences Between Statistics and Mathematics

Much thinking about statistics falls victim to an unfortunate habit of locating quantitative courses along an intellectual continuum which stretches from the mechanical to the mathematical, from recipes (bad) to theory (good). Two natural but destructive consequences of this one-dimensional view are that data tend to get associated with recipes and dismissed as devoid of intellectual challenge, and mathematical theory is taken to include all of statistical thinking, so that non-mathematical concepts of statistics are not taught. Real life is not so simple.

Statistics in practice involves a dialogue between models and data that is quite different from the deductive derivation of properties of models met in theory-based courses. Models are used to analyze data, but the data are allowed to criticize and even falsify a proposed model (diagnostics). Several different methods may be applied fruitfully to the same problem. The results of one study often suggest another study, not a final conclusion. Students need to meet and use this way of thinking.

—*David Moore*

The key is to teach statistics like statisticians instead of like mathematicians.

—*Walt Pirie*

Qualification: Math remains essential; you can never be too rich or too thin or know too much mathematics.

—*David Moore*

The question is, who is to be master—that's all.

—*Humpty Dumpty, Through the Looking Glass.*

Recommendation II: MORE DATA AND CONCEPTS: LESS THEORY, FEWER RECIPES

Almost any course in statistics can be improved by more emphasis on data and concepts, at the expense of less theory and fewer recipes. (To the maximum extent feasible, automate calculations and graphics.)

—*David Moore*

Statistical concepts are best learned in the context of real data sets. Fortunately, using the computer to automate calculations and graphics makes it possible to work with real data without becoming a slave to the mechanics.

The introductory course in statistics should focus on a few broad concepts and principles, not a series of techniques. Suggested concepts are: graphing data (as in Cleveland's book—this is not trivial), randomness (the idea of producing a predictable pattern through randomness is difficult for a student to grasp; it is not intuitive), inferential reasoning (ideas illustrated through bootstrap-like simulations are easiest to grasp and the formulas, for those who insist on using them, are approximations), experimental design (I've seen eyes light up for some reasonably intelligent people when they were able to set a stat-a-pult right on target after collecting information on a half rep of a two to the fourth factorial experiment; they did not think it could be done).

—*Dick Scheaffer, Florida*

I am less concerned with the mechanics than I am with students being able to display data quickly by hand or computer and then *interpret* the display. Examine many of the introductory texts today. See how many ask in chapter exercises for students to interpret the salient features of the graphs. The inability of students to interpret the graphics they produce makes the process merely an abstract exercise. So too with descriptive statistics. They are initially presented as a calculation exercise, without regard to their informational content.

—*Dick Gunst*

[...] was the first introductory textbook attempt that I know of to approach a data set with a set of candidate tools (parametric, non-parametric, and robust), and the attitude of trying to determine which would be best for this particular problem. When I embraced that approach, a lot of students were really enthusiastic about it. But the resistance by many of my colleagues was intense and intractable. Now the second edition is out, and sure enough, the non-parametric and robust material has been relegated to chapters at the end of the book like all other texts, where instructors can (and do) ignore it. Want to compare means of two groups? Push a button and out comes a t-test, the only thing we have to offer. No need to *think* about alternative approaches or make a judgment, just memorize the formula.

—*Walt Pirie*

Thinking about the role of probability can be a useful exercise in clarifying differences between statistical and mathematical thinking. Many statistical concepts don't rely on probability theory at all, and a course which puts statistical concepts ahead of mathematical theory will recognize that fact.

> "Yes, Virginia, there is statistics without probability." Many basic ideas can be discussed prior to any discussion of probability. Most of EDA (exploratory data analysis) and graphing falls into this category, as does the basic principle of model building and experimental design. EDA allows for exploration and summarization of data and the formulation of questions, some of which may have obvious answers just from the exploration. Modelling looks at relationships but, at the basic level, does not require a goodness-of-fit test. (Does the scatterplot go up or down? Are the points close to a line or is there lots of scatter? Does interest in sports seem to be associated with the sex of the respondent?) Experimental design involves planned investigations to answer specific questions, but the plan and the data that result can be looked at without probability notions. (It seems that I can get to school faster by route A than by route B. Stereo A is really the better buy when you consider its reliability.
>
> —*Dick Scheaffer*

Some important statistical concepts do, of course, depend on probability, but it is all too easy to ignore Humpty Dumpty and forget which is to be master.

> Probability should be introduced in a less threatening way than we have traditionally done, and only in the framework of enabling students to draw statistical conclusions that enhance their simple graphs, simple statistics. The topics chosen should be directed toward a continuity or constancy of purpose (to borrow from Deming). I would like students to be able to use probabilities to make statistical conclusions: to understand the difference between phenomena that are likely and those that are unlikely, and, above all, to understand the distinction between phenomena that are "real" (in a statistical sense) and those that are likely to have occurred by chance.
>
> —*Dick Gunst*

The distinction between mathematical theory and statistical concepts remains an important one even in thinking about the standard introduction to mathematical statistics.

> I don't think students who take the standard mathematical statistics course come away with even the faintest appreciation for what statistics is about. Unless students have had a previous course that does justice to data analysis, and so provides a meaningful context for the mathematical statistics, the course is mainly an opportunity to practice advanced calculus techniques. I think only three positions are tenable here:
>
> 1. The mathematical statistics course should *never* be taught to students who haven't first taken an applied course;
> 2. The mathematical statistics course must be radically revised, to integrate data analysis with the statistical theory; or
> 3. The mathematical theory of statistics should be introduced via an optional adjunct to the beginning applied course.
>
> —*George Cobb*

> Those students who have had my Statistical Thinking course early (freshmen, sophomores) are having a ball with mathematical statistics as juniors or seniors. Others struggle more because they get bogged down by probability theory and mathematical niceties like moment generating functions, and they have a harder time seeing what statistics is all about. This points to a need to hear statistics twice before it makes sense, and we cannot lose the connection to real data.
>
> —*Gudmund Iversen, Swarthmore*

My personal peeve about statistics courses in the typical undergraduate mathematics major, the upper-division offerings, is that they tend to 75% or more probability theory with possibly a second semester of mathematical statistics that most students do not take. Our mathematics majors often see little of the material taught in the introductory non-mathematical statistics courses for, say, psychology majors. This is especially true if a non-statistician teaches the course in the mathematics department. The 1981 CUPM report urged that the standard junior-level one-semester probability/statistics course taken by a mathematics major include a month of work with data. Most texts for this course still have nothing about working with data.

—Alan Tucker, SUNY Stony Brook

I think the probability–mathematical statistics sequence is important but should be preceded by a data analysis course.

—Jack Schuenemeyer, Delaware

In an ideal Grinnell I'd like my students to have an applied methods course as sophomores and then have a more traditional mathematical statistics sequence.

—Tom Moore, Grinnell

One of my colleagues was visiting from Norway, where they have a more extensive undergraduate statistics program than we do. Their experience was that many of the best students—those you particularly want to attract to statistics—were turned off by the early course without the mathematical underpinnings, and decided that they weren't interested in statistics. This surprised me at the time, but since then I have had a couple of quite bright mathematics majors take my elementary statistics course, with similar experiences. They seemed rather negative about statistics as we went through the elementary course, and neither was very interested in taking a mathematical statistics course because they felt they already knew all of that stuff. While I think I have convinced both of them to go on and take mathematical statistics, this experience indicates to me that we need a different beginning course than the typical elementary statistics course. And I suspect that the ideal beginning course for these mathematics people will not be the same as the ideal beginning course for the rest of our audience.

—Mary Parker

This sequence should not be the first course, or, at the least, should be accompanied by a lab that shows the other side (I almost said "the real nature") of the subject. For example, since the randomized comparative experiment is arguably the most important statistical idea of this century, one which has revolutionized the conduct of research in many fields of applied science, it's a sin to teach a first course that doesn't mention this idea and emphasize the contrast between observational and experimental studies. Because the two-sample t procedures and the mathematical model on which they are based ignore this distinction, the usual statistics theory course for mathematics majors also ignores it. That, as I said, is a sin. This example shows in brief what's wrong with the introduction we often give our majors.

—David Moore

Recent Research On How Students Learn

Shorn of all subtlety and led naked out of the protective fold of educational research literature, there comes a sheepish little fact: lectures don't work nearly as well as many of us would like to think. This rather discouraging assertion is supported by two clusters of research results. The first cluster shows part of what makes learning hard and lecturing often ineffective; the second shows the kinds of things that do seem to work when lecturing doesn't.

A. *Basic concepts are hard, misconceptions persistent.* As teachers, we consistently overestimate the amount of conceptual learning that goes on in our courses, and consistently under-estimate the extent to which misconceptions persist after the course is over.

> Ideas of probability and statistics are very difficult for students to learn, and conflict with many of their own beliefs and intuitions about chance and data. Students connect new ideas to what they already believe, and correct or abandon erroneous beliefs reluctantly, only when their old ideas don't work or are inefficient. Learning is enhanced by having students become aware of and confront their misconceptions.
>
> —*Joan Garfield, University of Minnesota*

> I am still chagrined by an experience in class several weeks ago. Using IQ scores $N(100, 15)$ generated by the computer, my class "discovered" the central limit theorem. No problems. They were fairly experienced with sampling distributions, expected the normal shape, knew that the expected value of the sample average equals the population mean, and weren't surprised by the standard error. We then went on to examine a case where the population was decidedly not normal, using the ages of pennies that we brought to class. Suddenly they had no idea where the expected value of the sample average should lie for their samples of size four. We spent a lot of time establishing that. I am chagrined because in previous semesters, using synthetic data, I had essentially assumed that the location of $\mu(\overline{x})$ was obvious to students, spending time instead establishing the shape and standard error of the distribution of sample means.
>
> —*Ann Watkins*

B. *Learning is constructive.* To absorb the full impact of these three words, you have to push their implied metaphor to its limits: concepts are constructs, learning is building. (Moreover, any building that students do will start with whatever conceptual raw materials they bring with them to the course. There's no such thing as starting from scratch.) By taking these construction images to the edge of the literal, and applying common-sense principles of carpentry to the process of teaching and learning, you can arrive at much the same conclusions as those who do research on how students learn: If you want to teach your students to build, you won't spend much time just talking, and what talking you do will occur on-site, where students who are learning-by-doing will want and need your comments on their work.

> Effective learning requires feedback. Students learn better by active involvement; they learn to do well only what they practice doing; they learn better if they have experience applying ideas in new situations.
>
> —*Joan Garfield*

Taken together, the two sets of results lead to a third recommendation:

Recommendation III: FOSTER ACTIVE LEARNING

As a rule, teachers of statistics should rely much less on lecturing, much more on the following alternatives:

1. *Group Problem Solving and Discussion.*

> I do not lecture at all. Instead, students are required to read the textbook, guided by a student handbook I have written containing study questions, sample problems, etc. Each day we first discuss the study questions, often arguing about issues. ... After our large group discussions, students then work in permanent small groups on activities, usually analyzing a set of data and discussing questions about these data sets.
>
> —*Joan Garfield*

2. *Lab Exercises.*

Statistics should be taught as a laboratory science, along the lines of physics and chemistry rather than traditional mathematics. Students must get their hands dirty with data. The laboratory must be a requirement and must contain more than just a few computers. This approach involves real data but also involves manipulative devices that include spinners, cards, bead boxes, a quincunx, stat-a-pults or similar devices for experimental design, and so on. Many things seem to work in the design of experiments realm (George Box's paper helicopters, Lego cars, rubber band guns, popcorn (see Hogg and Ledolter), balloons, melting ice cubes and on and on ...), the important idea is that an experiment is to be designed to answer a specific question and at least two important factors can be controlled.

—*Dick Scheaffer*

3. *Demonstrations Based on Class-Generated Data.*

Whether by counting the number of red M&M's in bags, taking surveys, or conducting simple experiments, the scientific enterprise referred to as the field of statistics must be experienced. This is very difficult to do in large lecture sections, so opportunities for demonstrations in lieu of hands-on individual experiences are a necessary alternative. Ideally, several demonstrations or experiences would occur in a single course. They will, I contend, prove to be one of the features of the course that will be remembered long after the formal analyses are forgotten.

—*Dick Gunst*

4. *Written and Oral Presentations.*

Students come to us with primarily an intuitive understanding of the world. It is part of our job to ferret out those intuitive processes and correct the incorrect ones. As far as I know, this can only happen by having students discuss and write about their understandings and interpretations of problems.

—*Dick Scheaffer*

5. *Projects, Either Group or Individual.*

Students in a first course should learn by doing. They will buy into the course and subject if they can formulate and design projects, collect and analyze data. Easy to use statistical software with graphics should be emphasized. Calculations should be de-emphasized.

—*Jack Schuenemeyer*

I'm doing student projects for the first time in an introductory course, and I have never had such enthusiasm. The assignment is to do a survey about some issue of interest at California State University at Northridge.

—*Ann Watkins*

Examples

Ah, generalities, which like fish glitter but stink. Here we escape glittering generalities.

—*David Moore*

Our proposed escape route is marked by two clusters of examples. The first cluster relates principally to our first two recommendations, on what to teach, and consists mainly of examples of entire courses. Taken together, these examples illustrate how remarkably different courses may be, both as to technical level and as to statistical topics, while nevertheless serving the general goals we have spelled out. The second cluster of examples relates principally to our third recommendation, on how to teach, and consists mainly of examples of parts of courses. These examples offer teachers of statistics a variety of alternatives to lecturing, all of them compatible with the examples of course content from the first cluster.

Teach Statistical Thinking

The following examples illustrate how the maxims "more data, fewer recipes" and "more concepts, less theory" improve students' statistical thinking. We begin with the junior high and high school level (Quantitative Literacy Project), then present three undergraduate general education courses (Chance, Quantitative Reasoning, and Statistical Thinking), then three courses that are technically somewhat more demanding (Statistical Methods, Time Series, and Multivariate Statistics), and end with two quite different variants of the standard mathematical statistics course.

These courses rely on computers in various ways. For example, both Mount Holyoke's quantitative reasoning course (No. 3) and Oberlin's mathematical statistical course (No. 7) use computers to analyze moderate-to-large archival data sets. (Most of the other courses also use computers to analyze data sets.) A very different approach to the mathematical statistics course (No. 8) relies on computers for simulation-based empirical investigation of the properties of estimators.

1. *The Quantitative Literacy Project.* (Dick Scheaffer, University of Florida)

The National Council of Teachers of Mathematics (NCTM) recently released their *Curriculum and Evaluation Standards,* which have a carefully delineated strand in statistics throughout the curriculum and an emphasis on modelling from data in other areas such as algebra and functions. That emphasis on data analysis should be woven through the mathematics curriculum and connected to other components of the curriculum is seen in the following:

> This standard should not be viewed as advocating, or even prescribing, a statistics course; rather, it describes topics that should be integrated with other mathematics topics and disciplines.

The NSF-funded Quantitative Literacy Project (QLP), a joint project of the American Statistical Association (ASA) and the National Council of Teachers of Mathematics (NCTM), served as the basis from which the strand in statistics was developed for the *Standards.* The QLP provides curriculum materials in certain areas of data exploration, probability, and inference, in a style that makes the material accessible to teachers and students, and provides a model framework for in-service programs to enhance the skills of teachers in the area of statistics and probability.

The curriculum units developed by the QLP explore elementary topics in data analysis, probability, simulation, and survey sampling, with new units being planned for exploring measurements and planning experiments. The approach is to use real data of interest to the students and simulations of real events to show how to use statistical ideas to extract useful information from numbers. Many of the statistical tools are graphical and reflect the latest thinking among practicing statisticians. Teachers using these materials are provided with opportunities to make heavy use of hands-on activities, group discussion, student projects, and report writing. (From "The ASA-NCTM Quantitative Literacy Project: An Overview and Possible Extensions." See also the report of the Focus Group on Quantitative Literacy in this volume.)

2. *Chance.* (J. Laurie Snell, Dartmouth)

This course is being developed at six colleges: Dartmouth, Grinnell, King, Middlebury, Princeton, and Spelman. The course studies chance issues currently in the news such as statistical issues in AIDS, gender issues in SAT examinations, the use of DNA fingerprinting in the courts, scoring streaks and records in sports, reliability of current political polls, interpreting data graphics, and the role of cholesterol in the prevention of heart disease. Such topics form the focus of the course. Concepts in probability and statistics are developed only to the extent necessary to understand the issues. Students read and compare treatments in the newspapers, popular science journals such as *Chance Magazine*, *Science*, and *Nature*, and original research articles. The goal is to make students more literate in probability and statistics and to permit them to make more intelligent choices when faced with chance issues. Experimental versions of the course are being given in the 1991 Fall term at Grinnell College by Tom Moore, at Middlebury College by William Peterson, and at Princeton University by Peter Doyle and Laurie Snell. These courses incorporate an emphasis on good writing, group learning, the use of the computer, and student projects.

3. *A Quantitative Reasoning Course.* (George Cobb, Mount Holyoke College)

In 1982, at Mount Holyoke College, a group of faculty began to plan what was eventually to find its way into the course catalog as Interdepartmental 100—Case Studies in Quantitative Reasoning. After gaining provisional approval by the faculty in 1986, the course was taught for the first time in 1987. As of fall 1990, the QR course, as it has come to be called, had been taught for eight consecutive semesters all told. Fifteen faculty from six departments had by then taught 22 sections of the course to a total of more than 250 students. The course, which has no prerequisites, is designed to appeal to students with a broad range of academic interests and widely differing mathematical backgrounds. It is not a simple presentation of technical methods followed by practice problems. Instead, case studies from a variety of disciplines form the subject matter of the course. Different quantitative methods are introduced and used in the attempt to develop understanding of these examples. The emphasis is not on rote computation, but on reasoning; not on formulas, but on ways to construct and evaluate arguments. The goals are to help students strengthen their analytical skills and acquire more confident understanding of the meaning of numbers, graphs, and other quantitative summaries they will encounter in many subsequent courses, no matter what their majors.

For example, witchcraft in seventeenth-century New England forms the central problem for investigation in the first third of the course, which concentrates on exploratory data analysis, with heavy emphasis on graphs, two-way tables, and other informal statistical tools for finding and presenting patterns in numerical data. In this case the data pertain to the 141 people accused of witchcraft in 1692 in Essex county, and to residents of Salem Village whose names appear in various tax records, petitions to the General Court, and minutes of the Village meetings. Discussions and assignments stress the process of working from numerical patterns to plausible explanations, and vice-versa, from possible explanations of the historical phenomena to relevant numerical evidence. The major project is to write a paper formulating, investigating, and discussing some hypothesis about the relationship between wealth and power as reflected in the historical records, along with some other hypothesis of the student's own choosing. ("The Quantitative Reasoning Course at Mount Holyoke College," in Samuel Goldberg (Ed.), "The New Liberal Arts Program: A 1990 Report," Alfred P. Sloan Foundation.)

4. *Statistical Thinking and Statistical Methods.* (Gudmund Iversen, Swarthmore College)

At Swarthmore we have carried this idea [Jim Landwehr's distinction between doing and thinking] so far that we have two introductory courses. In Statistical Methods I expect the students to be able to do their own analyses after they have finished the course. Most students end up doing the statistics in papers, lab reports, and senior theses. But this means going through regression, analysis of variance, and contingency tables. This year I am using the Moore/McCabe book (*Introduction to the Practice of Statistics,* Freeman, 1989). In Statistical Thinking I do not expect students to be able to do anything, but the goal is to have them understand uses of statistics they see all around them in scientific journals, books, newspapers, television, news magazines, etc. I am using the Moore paperback (*Statistics: Concepts and Controversies, 3rd edition,* Freeman, 1991) with additional material for this course. (A more detailed description appears as Appendix B.3.)

5. *Time Series Analysis As a First Course in Statistics.* (Robin Lock, St. Lawrence University)

A wide variety of fundamental statistical ideas can be conveyed through the study of time series. For example, the general notions of an underlying model for some real world phenomenon, estimation of its parameters from data, and diagnostic checking of the model assumptions are central themes in statistics. The models encountered in forecasting are fairly straightforward, yet can be used to effectively illustrate important principles such as parsimony, variability in parameter estimates, and criteria for choosing between competing models. The analysis of residuals to check model assumptions, suggest alternate models, or gauge the accuracy of the fit is a featured part of time series methodology which is often neglected in traditional introductions to statistics. Similarly, statistical graphics are used at many points throughout a time series analysis.

The point here is not to demonstrate that the field of time series analysis uses important statistical techniques, but rather that many of the fundamental concepts in applied statistics can be effectively introduced to students within the context of time series analysis. (Adapted from "Forecasting/Time Series Analysis: An Introduction to Applied Statistics for Mathematics Students," SLAW Technical Report No. 90-001. This and other technical reports of Statistics and the Liberal Arts Workshop are available from Don Bentley, Mathematics, Pomona College.)

6. *Multivariate Descriptive Statistics.* (Frank Wolf, Carleton)

A course in multivariate descriptive statistics that presupposes preparation in linear algebra can be used as a course to introduce students to statistical ideas. Such a course requires students to make heavy use of mathematical modelling and of concepts, techniques, and results from linear algebra and, hence, contributes to the students' understanding and appreciation of applied mathematics. The course is data-driven, and students learn to deal sensibly with real data. The course should be acceptable to a typical liberal arts mathematics department for credit in mathematics. It can be taught so as to assume no earlier training in statistics and yet be open and very useful to those students who have already taken one or more courses in that area. Such a course has been taught many times at Carleton. (From "Multivariate Descriptive Statistics: An Alternative Introduction to Statistics," SLAW Technical Report No. 90-004.)

7. *Data Analysis in the Mathematical Statistics Course.* (Jeff Witmer, Oberlin)

I believe it is imperative that students learn something of how statistical theory is applied in practice, but it is particularly difficult to cover much material on applied statistics while at the same time covering the mathematical statistics topics. I address this problem by offering an additional, one-credit course at the upper level.

In the spring of 1988, I started the course by discussing some of the statistical packages that are available on computers. I used MINITAB. I then presented some ideas from exploratory data analysis ... the use of normal probability plots and transformations of data ... compound smoothers ... statistical quality control ... I devoted roughly half the course to the general topic of regression, relying heavily on MINITAB.

In the spring of 1989, rather than present lectures, I involved the entire class in a data analysis project. We analyzed the results of a survey I helped conduct of libraries at liberal arts colleges. The students helped me explore roughly 200 variables measured on 97 colleges. I gave each of the students access to the computer file that contained the data and told them, "Explore, generate graphs, fit models, and let me know what you learn." We spent class periods discussing the data set and the statistical methods we were using to analyze it.

I am planning to teach the course in 1990 in a similar fashion. The reactions of students to this course have been positive. I believe, and they seem to agree, that seeing statistical methods applied to real data motivates students to want to learn more about the subject. (From "Data Analysis: An Adjunct to Mathematical Statistics at Oberlin College," SLAW Technical Report No. 90-003.)

8. *Computer-Enhanced Mathematical Statistics.* (Marsha Davis, Mount Holyoke)

Under a grant from FIPSE (Fund for Improvement of Post-Secondary Education), Mathematical Statistics at Mount Holyoke College has been redesigned, and now meets in a computer classroom. With the aid of technology, the new course incorporates a constructivist approach to mathematics instruction. Laboratory projects have been designed to support theoretical material and to guide students in discovering concepts for themselves. In one project, for example, students use computer simulation to examine the sampling distributions of sample means as an introduction to the normal distribution and the central limit theorem. Another project, "The Taxi Problem," places students in a hypothetical situation where they must estimate the parameter N of a discrete uniform distribution on the integers $1-N$. Students suggest plausible estimators, generate ideas of reasonable criteria for selecting a "best" estimator, modify estimators based on theoretical considerations, and present their choice of estimator based on results from a simulation study. Through this process students gain an understanding of how a research statistician works, as well as a chance to experience the interplay of working with theory and testing ideas with simulation.

Foster Active Learning

The following examples correspond to the various alternatives to lecturing mentioned in our third recommendation: group problem solving and discussion, lab exercises, demonstrations, written and oral presentations, and projects. As many of the examples illustrate, there are effective alternatives to lecturing that do not use computers. (See also Robin Lock and Tom Moore, "Low-Tech Ideas for Teaching Statistics," SLAW Technical Report No. 91-008.)

On the other hand, certain kinds of lab exercises (e.g., Nos. 3, 7, 8 above) and class demonstrations (No. 3b below) are impossible without computers, as are most projects (Nos. 5a and 5b below). In particular, demonstrations that rely on even the simplest dynamic graphics (No. 3b below) can be extremely effective, and are impossible without computers.

1. *Group Problem Solving and Discussion.* (Joan Garfield, University of Minnesota)

In my classes, I do not lecture at all, which takes a while for students to adjust to. Instead, students are required to read the textbook before coming to class, guided by a study guide/student handbook I have written containing study questions, sample problems, etc. When students come to class each day we first discuss the study questions, often arguing about issues such as which is the best measure of center to use, which type of plot gives the most information, etc. They rapidly learn that there is often not one right answer nor one way to solve a problem. At first they think this means that anything is OK, but then learn what is important is being able to justify a claim, defend a point of view, and judge the appropriateness of a solution process. After our large group discussions, students work in permanent small groups on activities, usually analyzing a set of data and discussing questions about these data sets. I find the Quantitative Literacy Project materials very useful for these small group activities. Each write-up is turned in that day and the group receives a group score. I also give weekly quizzes on the homework problems. This method works extremely well. Students tend to enjoy the course and express amazement that they are actually doing statistics (and it's fun!). Many bring in data sets of interest to them to analyze for their "real life" problem assignments, and learn how to deal with messy data that are not easily plotted, missing values, and related issues. (A more detailed description appears as Appendix B.2.)

2. *Lab Exercises.* (Dick Gunst, Southern Methodist University)

I have a "nuts and bolts" experiment that I've used many times to illustrate (a) the need for data, (b) application of a simple two-factor factorial experiment, and (c) the ability of simple graphics (point plots) to convey important information. The experiment involves students selecting nuts and bolts from a tall, small mouth jar containing a variety of sizes of nuts and bolts. The objective is to select and screw four nuts onto four bolts in as short a time as possible. Students operate in groups, taking turns selecting and fastening the nuts and bolts and timing those who do. After the initial times are obtained and plotted, discussion ensues over why the times are unacceptable and what could cause a reduction of the times. A simple experiment is then conducted using four combinations of jar sizes and nut sizes and the results are again plotted. The sorting of jars and nuts results in lower times *for all four groups*, an unexpected result. (A more detailed description appears as Appendix B.4.)

3a. *Demonstrations Based on Self-Generated Time Series.* (Howard Taylor, Delaware)

(There are no prerequisites for this one.) The students are asked to write down a random series of $n = 100$ numbers chosen from the digits 1, 2, ... , 20. The form provided has columns with 1. ____, 2. ____, etc., to 100. ____, a psychological nudge to write the numbers down sequentially. (But some of the brighter students will skip around, entering a bunch of 1's, then a bunch of 2's, and so on.) For comparison, one or more students (or the instructor) uses a table of random numbers, or a programmable hand calculator, to form their lists.

The analysis:

(i) The students are asked to do a frequency tabulation and draw a histogram. The result is that most students have frequencies that look fairly realistic.

(ii) Next, a time series plot is drawn. These typically look less realistic when compared to the control plot. Some students go up-down too much; others are too smooth.

(iii) This suggests a plot of $X(t)$ vs $X(t+1)$ which typically looks less random, showing either a positive or negative correlation.

(iv) Students are asked to count how many times $X(t) = X(t+1)$, and how often this should happen in a random time series.

(v) Students are asked to calculate the number of runs above and below the median R. The critical points for a test of randomness are $R \leq 40$ or $R \leq 62$. These critical points are written on a slip of paper and given to a student prior to the calculations of the R's. Then the time series are divided into two groups according to the R values. This virtually always very nicely divides them into human generated versus computer or table generated, effectively demonstrating how hard it is for humans to be truly random.

(Two more examples of demonstrations appear in Appendix B.1.)

3b. *Simple Computer-Based Demonstrations.* (Ann Watkins, California State University at Northridge)

(i) Standard deviations: One computer demonstration I use in class requires a large screen projector but only the simplest software. Students can construct a population of, say, 100 numbers between 1 and 100. The program computes the mean and standard deviation. The game we play in class is to find the population with the largest standard deviation (and then with a given standard deviation). This sounds incredibly simple-minded, but students find it challenging as they are just coming to grips with the idea of standard deviation.

(ii) Regression: You can build effective labs or class demonstrations using programs that allow students to change the points of a scatterplot and watch the least squares line (and coefficients) change, or leave the points fixed but change the fitted line, and watch the residual sum of squares change.

4. *Written and Oral Presentations.* (Gudmund Iversen, Swarthmore College)

There is an increased emphasis on writing in today's undergraduate curriculum, and papers can play an important role in an introductory statistics course. With the existence of good interactive statistical software it is possible to move the classic introductory statistics course away from the study of formulas to the study of statistical thinking and the role of statistics in society. In such a new course students get an increased understanding of statistical ideas by writing papers across a wide range of topics; actual topics have ranged from a comparison of statistics and religion to a study of the relationship between the time of first class in the morning and the distance from the bed to the alarm clock. (See "Writing Papers in a Statistics Course," Proceedings of the Section on Statistical Education, American Statistical Association, 1991. See also Norean Radke-Sharpe, "Writing As a Component of Statistics Education," *American Statistician,* Vol. 45, No. 4, November 1991.)

5a. *Projects.* (Tom Moore, Grinnell, and Katherine Halvorsen, Smith)

Student projects can teach concepts not usually encountered in introductory or second-level statistics courses. Questions about study design, study protocols, questionnaire construction, informed consent, confidentiality, data management, data cleaning, and handling missing data may arise when students deal with collecting and analyzing their own data. ...This paper describes the student projects we have used in introductory courses and in second-level statistics classes. It addresses the issues of motivating, monitoring, and evaluating student projects, and discusses some unique problems student projects present for instructors using them. ...In our experience students usually conclude that the project was one of the most useful parts of the course. Some comment that the project made them apply everything they learned as soon as they learned it. For some it is the first time they have stood in front of an audience to present their work. On the whole we would encourage other faculty to use this kind of project in their classes. (See "Motivating, Monitoring, and Evaluating Student Projects," Proceedings of the Section on Statistical Education of the American Statistical Association, 1991.)

5b. *Projects.* (Don Bentley, Pomona)

There are at least three types of exercises involving data for students to work with at the introductory level: the standard fictional data set, real data extracted from the literature, and statistical consulting. Students who have a potential interest in statistics as a career should be given the opportunity to become meaningfully involved in the analysis of data from original scientific investigations ...to encounter the excitement of being the first to know the results of the statistical analyses. ...

A list of projects with which students have been involved include soaking and cleaning solutions for hard contact lenses, sterilizing solutions for soft contact lenses, proving efficacy of intraocular lenses, evaluation of a retroprofusion process used with angioplastic surgery, and investigation of the use of ambulatory tocodynamometry in high risk pregnancies. These industry-generated projects provide a wealth of opportunities for students to become involved in the analysis of meaningful data. The necessity for accuracy is clearly defined. This is an element frequently lacking in the classroom experience. The importance of the role of statistics in the research is also made very clear to the student, both by the financial implications of the project to the corporation, and the implications for health care in general. (The consequences of both the Type I and Type II errors gain real meaning.) (See "Recruiting and Training Undergraduates Through Statistical Consulting," Proceedings of the Section on Statistical Education, American Statistical Association, 1991.)

Making It Happen

The way faculty change and learn is probably not so very different from the way students change and learn. In particular, reports alone probably have about as much effect on the way most faculty teach statistics as lectures alone have on the way most students understand statistics. Our Focus Group may be full of sound and fury, but without effective follow-through, we'll fail to signify even at the ten percent level.

Change must overcome four inertias: one logistical, one intellectual, one interpersonal, and one institutional.

1. Logistical Inertia:

Good data sets are hard to find. Anyone reading this report could easily invent enough examples to fill a lecture on differentiating polynomials, and it would take at most five minutes or so. But how long would it take you to come up with just one real data set, say from cognitive psychology, to illustrate the effect of outliers in one-way analysis of variance?

Automating statistical busy-work has a high start-up cost. Unless you're already set up with a data analysis package for your classes, the ordeal of first choosing a good one, then getting it installed, then learning how to use it yourself, and finally learning how to teach your students to use it—all this may not be at the top of your list of ways to spend what would otherwise be your next vacation.

2. Intellectual Inertia:

Learning to handle the ambiguities of statistics takes time, practice, and hard thought. Even with software installed and data sets in hand, doing a proper analysis and interpretation is a kind of challenge that many who teach statistics are not prepared to meet, mainly because, through no fault of their own, they've rarely if ever seen it done, and their training and experience have not prepared them either to do it or to value the doing of it. Remember that in Plato's curriculum, students were to devote their entire first decade of study to mathematics, because the other subjects, being less clear-cut, were understood to be *harder.* (Mathematics is often mistaken for being harder because the absence of ambiguity makes the subject much less forgiving of low quality effort.)

3. Interpersonal Inertia:

Giving up the familiar role of "I talk, you listen" doesn't always come easily. Teachers who are skillful and comfortable delivering a lecture on volumes of revolution may, through lack of role models and lack of practice, find themselves neither skillful nor comfortable with the group dynamics of labs or discussions, where the direction of conversation is not set in advance, and there will often be more than one defensible position on what the data have to say.

4. Institutional Inertia:

Most deans and department heads don't care very much whether statistics is taught well. One consequence is that untenured mathematics faculty are all too often being quite realistic to see a fork in their career path—to the left, doing what it takes to teach statistics well, or to the right, doing what it takes to get tenure. A second consequence, apparently, is that many mathematics departments have not been willing to do what it takes to recruit and retain a Ph.D. statistician. According to Moore and Roberts (*American Statistician,* 1989, 43:2) only 26 of 80 mathematics departments at responding liberal arts colleges could claim a Ph.D. statistician. Half had no one with even a masters-level degree in statistics.

In the hope of overcoming these inertias, we offer two clusters of meta-recommendations:

Meta-Recommendation I: WORK AT THE GRASS ROOTS

The statistics and mathematics professions should do more to support those who want to teach statistics well, through (1) new sections in journals, (2) a newsletter, (3) improved access to data, (4) more e-mail conferencing, (5) more workshops for teachers, and (6) better preparation for teaching in graduate schools.

It seems to me that ASA could do considerably more about undergraduate teaching than it currently does. For example, the other ASA to which I belong (American Sociological Association) maintains a Teaching Resource Center that collects and circulates such teaching materials as syllabi and reading lists and holds teaching workshops. The Association also sponsors a journal entitled *Teaching Sociology* that is a veritable treasure chest of ideas for managing classes and of research on the efficacy of teaching strategies.

—Judy Tanur, SUNY Stony Brook

1. *Establish New Sections in Existing Journals.*

It would be easy to assume that most of the things we recommend should be done by the ASA, but some of this must be done in MAA journals. Quite a few statistics teachers in mathematics departments, members of MAA but not ASA, need to be encouraged to participate in this conversation.

—Mary Parker

There should be sections devoted to statistical education in some of the existing journals. A section on statistical education devoted to undergraduate teaching and learning should not be used as a vehicle to publish low quality research papers. I'm opposed to starting a new journal. Libraries are cutting circulation and I can't find time to read existing journals.

—Jack Schuenemeyer

The "Teacher's Corner" in the *American Statistician*, although interesting, does not address our concerns.

—Mary Parker

I'd like things set up so most submissions would be short, each one describing a project or lab or class demo. There should be a well-planned classification scheme, so that a highly structured index would be easy to construct, to make it possible for someone wanting a demo on the central limit theorem for a large lecture class to locate quickly a write-up of Howard Taylor's example; ditto for someone wanting examples of two-hour computer labs on regression diagnostics for a small class of students with no linear algebra background, or ...

—George Cobb

2. *Establish a Newsletter.*

I like the idea of a newsletter to be sent out three or four times a year. The "Undergraduate Mathematics Education" newsletter (*UME Trends*) began as a funded project (NSF, I think) and now is supported by paid subscribers like myself. It is an excellent source of information about current projects, programs, research, books, software, conferences, workshops, teaching ideas, etc., all related to teaching undergraduate mathematics courses.

—Joan Garfield

3. *Expand Efforts to Make Data Sets Readily Available.*

Statisticians should contribute more often to the bank of data sets available by e-mail via the StatLib file server at Carnegie Mellon University. (See Appendix C for a brief description and instructions on how to get data sets.) Data sets should be classified and indexed by area of application, structure (e.g., experimental two-way factorial in complete blocks), and interesting statistical features (e.g., data need to be transformed to logs), so that teachers could use the index to retrieve examples to fit particular pedagogical needs. A newsletter or section of a journal could list such summary information for data sets that had been contributed since the last issue. Periodically, anthologies of these data sets should be published.

4. *Expand Efforts to Involve Statistics Teachers in E-mail Conferencing.*

Tim Arnold of North Carolina State has just recently set up a procedure for joining such a group, one established by the Statistical Education section of the ASA. (See Appendix C for a description and instructions on how to join.)

> In the ten years or so I have been teaching statistics, I have not seen many opportunities to participate in the kind of discussion we are having in this group. I think that statistics teachers would profit by, and enjoy, participating in such discussions. ...We need to be providing the kind of forum and inspiration to people that will encourage each of them to think of ideas of their own. For me, reading what other people do is marvelously inspiring. I think it would be to others.
>
> —*Mary Parker*

5. *Run Workshops at National and Regional Meetings.*

6. *Establish Programs Within University Departments to Help Prepare Graduate Students to Teach.*

> Our department also has a formal program for teaching our graduate students to teach. (Of course, what they teach is introductory sociology, not statistics, but I would think some of the ideas are transferable.) In the fall semester the teaching practicum meets once a week to study the art and craft of teaching, observe teachers who have a reputation for expertise and discuss their techniques with them, etc. Each member of the practicum prepares a syllabus for an introductory course, and gives several practice classes to the other members of the practicum. In the spring semester, each graduate student teaches his/her own section of the introductory sociology course under the supervision of the practicum instructor (often me). Again, it seems to work. The graduate students are considerably less nervous and considerably more effective with their classes than they would otherwise have been, and they have won far more than their share of university-wide awards for excellence in teaching. But again, this system is very demanding of resources.
>
> —*Judy Tanur*

Meta-Recommendation II: WORK AT THE TOP

Statisticians and mathematicians should work together with academic administrators—first to improve the way colleges evaluate work by statisticians and others who teach statistics, and second, to insure that improved evaluation reshapes institutional criteria both for faculty recruitment and for subsequent personnel decisions.

> Without additional motivation and encouragement, I don't think our material will be read by those who need it most. That kind of motivation usually has to come from one's senior colleagues within a department. What if young faculty see that the reward system is weighted mostly toward refereed publications? We need to enlist the help of others who will influence them. The obvious group is Department Heads. That could be done through the auspices of the MAA and ASA, both of whom do have periodic Department Heads meetings.
>
> —*Walt Pirie*

> We need to think hard about how to reach not just teachers of statistics, but also their chairmen who assign them the statistics courses. We need to impress presidents and provosts and deans that

things are bad, that they need to do something about the teaching of statistics in their schools, and that we are available to help.

—*Gudmund Iversen*

The comment which follows, by a newly elected Fellow of the ASA, came early in our group's work, but serves well as a final summary:

> At Delaware we teach an introductory, non-calculus based, traditional, two-semester course to 1000 students per year. Problems: too much lecturing, too much emphasis on hypothesis testing, too much material, too little involvement by students, and too little reward for those who spent time trying to improve the situation.
>
> —*Jack Schuenemeyer*

Focus Group Participants

bushaw@wsuvml.bitnet	DONALD W. BUSHAW, *Washington State University.*
gcobb@mhc.bitnet	GEORGE W. COBB, Moderator, *Mount Holyoke College.*
gunst@smuvml.bitnet	RICHARD F. GUNST, *Southern Methodist University.*
iversen@swarthmr.edu	GUDMUND IVERSEN, *Swarthmore College.*
jml@research.att.com	JAMES M. LANDWEHR, *AT&T Bell Laboratories.*
dsm@chi.stat.purdue.edu	DAVID S. MOORE, *Purdue University.*
mooret@grinl.edu	THOMAS L. MOORE, *Grinnell College.*
parker@math.utexas.edu	MARY R. PARKER, *Austin Community College.*
wltpirie@vtvml.edu	WALTER R. PIRIE, *Virginia Polytechnic Inst. & State Univ.*
scheaffe@orca.stat.ufl.edu	RICHARD L. SCHEAFFER, *University of Florida.*
jacks@brahms.udel.edu	JOHN H. SCHUENEMEYER, *University of Delaware.*
jtanur@sbccvm.sunysb.edu	JUDITH M. TANUR, *SUNY Stony Brook.*
pamh@brahms.udel.edu	HOWARD M. TAYLOR, *University of Delaware.*
awatkins@vax.csun.edu	ANN E. WATKINS, *California State University at Northridge.*
pqa6031@ca.acs.umn.edu	JOAN GARFIELD, *University of Minnesota.*

Suggested Reading

EXPLORATORY DATA ANALYSIS

1. Hoaglin, David C.; Mosteller, Frederick; Tukey, John W. (Eds.). *Exploring Data Tables, Trends, and Shapes.* New York: John Wiley, 1985.
2. Hoaglin, David C.; Mosteller, Frederick; Tukey, John W. (Eds.). *Understanding Robust and Exploratory Data Analysis.* New York: John Wiley, 1983.
3. Mosteller, Frederick and Tukey, John W. *Data Analysis and Regression.* Reading, MA: Addison-Wesley, 1977.
4. Tukey, John W. *Exploratory Data Analysis.* Reading, MA: Addison-Wesley, 1977.

GRAPHICS

1. Becker, Richard A.; Cleveland, William S.; Wilks, Alan R. "Dynamic graphics for data analysis," *Statistical Science*, 1986, V. 1, pp. 355–395.

2. Chambers, J.M.; Cleveland, W.S.; Kleiner, B.; Tukey, P.A. *Graphical Methods for Data Analysis.* Belmont, CA: Wadsworth, 1983.
3. Cleveland, William S. *The Elements of Graphing Data.* Belmont, CA: Wadsworth, 1985.
4. Cleveland, William S. and McGill, M.E. (Eds.). *Dynamic Graphics for Statistics.* Belmont, CA: Wadsworth, 1988.
5. Kolata, Gina. "Computer graphics comes to statistics." *Science*, 217 (1982), p. 919.
6. Tufte, Edward R. *The Visual Display of Quantitative Information.* Chester, CT: Graphics Press, 1983.
7. Tukey, John W. "Data-based graphics: Visual display in the decades to come." *Statistical Science*, 5:3 (1990), pp. 327–339.
8. Wainer, Howard and Thissen, David. "Plotting in the modern world: Statistical packages and good graphics," *Chance*, 1:1, 1988.
9. Weihs, Claus and Schmidli, Heinz. "On-line multivariate exploratory graphical analysis." *Statistical Science*, 5:2 (1990), pp. 175–226.

MISCELLANEOUS

1. Barnett, Vic. *Comparative Statistical Inference, Second Edition.* New York: John Wiley, 1982.
2. Belsley, D. A., Kuh, E., and Welsch, R. E. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York: John Wiley, 1980.
3. Cook, R. Dennis and Weisberg, Sanford. *Residuals and Influence in Regression.* London: Chapman and Hall, 1982.
4. del Pino, Guido. "The unifying role of iterative generalized least squares in statistical algorithms." *Statistical Science,* 4:4 (1990), pp. 394–408.
5. Efron, Bradley and Tibshirani, R. "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy." *Statistical Science*, 1 (1986), pp. 54–77.
6. Hoaglin, David C. and Moore, David S. (Eds.). *Perspectives in Contemporary Statistics.* Washington, DC: Mathematical Association of America, 1992.
7. Walton, Mary. *The Deming Management Method.* New York: G.P. Putnam, 1987.

SOME WAYS STATISTICS DIFFERS FROM MATHEMATICS

1. Kempthorne, Oscar. "The teaching of statistics: Content versus form." *The American Statistician*, 34 (1980), pp. 17–21.
2. Landwehr, J.M. "Discussion of the role of statistics at four-year undergraduate institutions." *Proceedings of the Section on Statistical Education of the American Statistical Association*, 1990.
3. Moore, David S. "Should mathematicians teach statistics?" (with discussion). *The College Mathematics Journal*, 19 (1988), pp. 3–25.
4. Moore, David S. "Teaching statistics as a respectable subject." SLAW Technical Report No. 91-002, Department of Mathematics, Pomona College.
5. Moore, Thomas L. and Witmer, Jeffrey. "Statistics within departments of mathematics at liberal arts colleges." *American Mathematical Monthly*, 98 (1991), pp. 431–436.

Appendix A: Helping Students Learn

by Joan Garfield, UNIVERSITY OF MINNESOTA

A summary of research related to learning statistics, drawn from the areas of learning and cognition, mathematics education, and statistical learning:

A. Basic Concepts are Hard, Misconceptions Persistent

1. *Teachers should not underestimate the difficulty students have in understanding basic concepts of probability and statistics.* Ideas of probability and statistics are very difficult for students to learn and conflict with many of their own beliefs and intuitions about data and chance (Shaughnessy, in press; Garfield and Ahlgren, 1988). Students may be able to answer some test items correctly or perform calculations correctly while still misunderstanding the basic ideas and concepts (Konold, 1990; Garfield and delMas, in press). Students' misconceptions may be strong and resilient (Konold, in press; Garfield and delMas, in press; Shaughnessy, in press). "It is very difficult to replace a misconception with a normative conception, a primary intuition with a secondary intuition, a judgmental heuristic with a mathematical model. Beliefs and conceptions are slow to change" (Shaughnessy, in press).

2. *Learning is enhanced by having students actually become aware of and confront their own misconceptions.* Activities that help students evaluate the difference between their own beliefs and intuitions about chance events and actual empirical results help students learn even better (delMas and Bart, 1989; Shaughnessy, 1977). If students are first asked to make guesses or predictions about data and random events, they are more likely to be engaged and motivated to learn the results. When experimental evidence explicitly confronts their predictions, they should be helped to evaluate this difference. In fact, unless students are forced to record and then compare their predictions with actual results, they tend to see in their data confirming evidence for misconceptions of probability. Research in physics instruction also points to this method of testing beliefs against empirical evidence (e.g., Clement, 1987).

B. Learning is Constructive

1. *Learning is a constructive activity.* Students learn by constructing knowledge, not by passive absorption of information (Resnick, 1987; von Glasersfeld, 1987). Students approach a learning activity with prior knowledge, assimilate new information, and construct their own meaning (NCTM, 1989). They connect the new information to what they already believe. Students accept new ideas only when their old ideas don't work or are inefficient. Ideas are not isolated in memory but are organized and associated with language and experiences. Students have to construct their own meaning for what they are learning, regardless of how clearly a teacher or a book tells them something (AAAS, 1989). Students retain best the mathematics that they learn by processes of internal construction and experience (NRC, 1989).

2. *Students learn better by active involvement in learning activities.* Students learn better if they are engaged and motivated to struggle with their own learning. Teachers should involve students in their own learning by using activities that require students to express their ideas both orally and in writing. Students learn better if they work cooperatively in

small groups to solve problems and learn to argue convincingly for their approach among conflicting ideas and methods (NRC, 1989).

3. *Students learn to do well only what they practice doing.* Practice may be through hands-on activities, cooperative small groups, or computer work. Students also learn better if they have experience applying ideas in new situations. If they practice only calculating answers to well-defined problems, then that is all they are likely to learn. Students cannot learn to think critically, to analyze information, to communicate ideas, to make arguments unless they are permitted and encouraged to do those things over and over in many contexts. Merely repeating and reviewing tasks is unlikely to lead to improved skills or deeper understanding (AAAS, 1989).

4. *Effective learning requires feedback.* Learning is enhanced if students have opportunities to express ideas and get feedback. Feedback ought to be analytical, and come at a time when students are interested in it. There must be time for students to reflect on the feedback they receive, make adjustments, and try again (AAAS, 1989).

5. *Computers should be used to enhance learning, not just to crunch numbers.* Computer-based instruction appears to help students learn basic statistics concepts by providing visual representations as well as the capability for easy interactive data exploration (Rubin, Rosebery and Bruce, 1988; Weissglass and Cummings, in press).

6. *Statistics instruction should shift from lecture-based classes emphasizing learning mechanical skills, to student-centered courses emphasizing application, understanding, and communication.* Based on research in learning and cognition, and mathematics education, the National Council of Teachers of Mathematics (1991) is urging the following shifts in mathematical instruction, which also apply to statistics instruction:

- Toward classrooms as mathematical communities.
- Toward logic and mathematical evidence as verification, away from the teacher as the sole authority for right answers.
- Toward mathematical reasoning, away from merely memorizing procedures.
- Toward conjecturing, inventing, and problem solving, away from an emphasis on mechanistic answer-finding.
- Toward connecting mathematics, its ideas, and its application away from teaching mathematics as a body of isolated concepts and procedures.

References

1. American Association for the Advancement of Science. *Science for all Americans.* Washington, DC: American Association for the Advancement of Science, 1989.
2. Clement, J. "Overcoming Students' Misconceptions in Physics: The Role of Anchoring Intuitions and Analogical Validity." Proceedings of the Second International Seminar, Misconceptions and Educational Strategies in Science and Mathematics. Ithaca, NY: Cornell University, 1987.
3. Garfield, J.B. and Ahlgren, A. "Difficulties in Learning Basic Concepts in Probability and Statistics." *Journal for Research in Mathematics Education*, 19 (1988), 44–63.
4. Garfield, J.B. and delMas, R. "Students' Conceptions of Probability." In *Proceedings of the Third International Conference on Teaching Statistics.* Dunedin, New Zealand. (In press.)

5. Konold, C. "The Origin of Inconsistencies in Probabilistic Reasoning of Novices." In *Proceedings of the Third International Conference on Teaching Statistics,* Dunedin, New Zealand. (In press.)
6. National Council of Teachers of Mathematics. *Curriculum and Evaluation Standards for School Mathematics.* Reston, VA: National Council of Teachers of Mathematics, 1989.
7. National Council of Teachers of Mathematics. *Professional Standards for Teaching Mathematics.* Reston, VA: National Council of Teachers of Mathematics, 1991.
8. National Research Council. *Everybody Counts: A Report to the Nation on the Future of Mathematics Education.* Washington, DC: National Academy Press, 1989.
9. Resnick, L. "Education and Learning To Think." Washington, DC: National Research Council, 1987.
10. Rubin, A.; Rosebery, A.; and Bruce, B. "ELASTIC and Reasoning Under Uncertainty." Research Report No. 6851. Boston, MA: BBN Systems and Technologies Corporation, 1988.
11. Shaughnessy, J.M. "Misconceptions of Probability: An Experiment With a Small-group Activity-based Model Building Approach to Introductory Probability at the College Level." *Educational Studies in Mathematics,* 8 (1977), 285–316.
12. Shaughnessy, J.M. "Research in Probability and Statistics." In D.A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning.* New York: Macmillan, 1977.
13. Von Glasersfeld, E. "Learning As a Constructive Activity." In C. Janvier (Ed.), *Problems of Representation in the Teaching and Learning of Mathematics.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1987, pp. 3–17.
14. Weissglass, J. and Cummings, D. "Dynamic Visual Experiments with Random Phenomena." In Walter Zimmerman and Steve Cunningham (Eds), *Visualization in Teaching and Learning Mathematics.* MAA Notes No. 19. Washington, DC: Mathematical Association of America, 1991.

Appendix B: Examples

1. Classroom Experiments (Howard Taylor)

Central Limit Theorem. (Prerequisite: Probability distribution, means and variances; a class of 40 to 100.) I start with a jar containing tags numbered 0, 1, ..., 9 and select a tag at random, writing its number on the board. Then I select ten tags, with replacement, and compute their sample mean. A student repeats the experiment, providing another single digit and another sample mean. Then in the interest of speed, I pass out slips of paper, each slip containing a single digit plus ten more digits and their sample mean. (I create these using APL.) The students call out their single digit in order and we create a frequency tabulation. Usually you have to smooth it into five groups to get a reasonable histogram. You can compare the observed frequencies with the uniform probabilities. Of course the sample experience does not exactly conform to the population of the sample means. Prior to the tabulation, I predict the frequencies in the various groups from normal table calculations that I make ahead of time. From the histogram, we observe that the sample means are centered about the same number as the individual observations, there is less variability in the sample means, and the bell-shaped curve has started to appear. Again, we compare the observed frequencies to the predicted frequencies. (The students are often amazed that I can make these predictions.)

Acceptance Sampling. (Prerequisite: Elementary binomial probabilities.) We compute the probability of accepting the lot when a sample size of 20 (and 50) items is selected, and we accept the lot if there are c or less defectives in the sample, as a function of the fraction p of defectives in the very large lot. We draw the operating characteristic curves for $c = 0, 1, 2$ and $0 < p < .20$. We then choose $n = 50$ and $c = 1$ and pass out simulated data, 5×10 arrays of 0's and 1's representing each student's sample. The first data is from a low acceptable value of p. We count how many student's accepted their lots, and how many rejected theirs, and compare the fraction accepted with the operating characteristic curve. Then we repeat with simulated samples from a higher value of p.

I find acceptance sampling a natural way to introduce students to hypothesis testing. Obviously, I treat Type I and Type II errors right from the start. One of the things pointed out at the end is that we could use a similar simulation to evaluate a very complicated acceptance rule whose operating characteristics we could not compute analytically.

2. Teaching Without Lecturing (Joan Garfield)

Case Study: An Alternative Approach to Teaching Introductory Statistics. This course was developed to introduce undergraduate students to the basic statistical ideas and methods. The only prerequisite for the course is a high school level course in algebra. The course is based on the following components:

1. *Involvement with real data.* Students are involved in the analysis of real data in order to solve problems of interest to them. This involves analyzing existing data sets gathered from a variety of resources as well as collecting data of their own.
2. *Emphasis on exploring data.* The focus of the course is on exploring data and this theme is continued throughout the course. The intent is to help students develop the attitude that statistical work is like detective work: looking for clues in the data, speculating about the real situation that a sample of data represent, and learning how different ways of representing data give different information.
3. *Use of software designed to build conceptual understanding.* Software is used not just as tools to crunch numbers and to display data more quickly and easily than by hand, but as a means to build concepts, confront misconceptions, and explore data by creating multiple representations.
4. *Emphasis on oral and written communication.* Students are given opportunities to communicate about statistics both orally (as they solve problems and analyze data in groups) and in written form (writing interpretations of data and results of student projects).
5. *Confronting misconceptions.* Students are confronted with their faulty reasoning, incorrect intuitions, and misconceptions through carefully designed experiences on the computer, group problem solving, and class discussion.

Course Details. The text used is *Statistics and Data Analysis* by Andrew Siegel. A study guide, written by the instructor, offers additional information, study questions to guide students' reading of the text, and sample problems with full solutions. Students are instructed to read the assigned chapters, jot down answers to the study questions, and work out the sample problems before coming to class. The class meets three times a week for ten weeks. Two sessions are two hours long and the third meets for fifty minutes (where students turn in homework and take a quiz).

Each two-hour session begins with a discussion of the assigned material and questions students have about the reading and homework assignment. This takes about 30 minutes. The next 45 minutes students work in assigned groups of 3 to 4 students per group, analyzing one or more small data sets provided to them which use the methods from their assigned readings. These data sets are accompanied by questions which help students focus on what to look for in their analysis, and require them to summarize in writing what they did and what they learned.

The next 45 minutes are spent in a computer lab using software to illustrate techniques, explore data, build conceptual understanding, or confront misconceptions. The software programs currently used are not yet commercially available, but are available from the authors. They are the Chance Plus Software (Clifford Konold, SRRI, Hasbrouck Lab, University of Massachusetts, Amherst, MA 01003), and ELASTIC Software (Andee Rubin, TERC, 2067 Massachusetts Ave., Cambridge, MA 02140).

Evaluation. Students receive grades in the course based on their scores on each of the following: weekly homework assignments; weekly quizzes; daily group activities; mid-term and final exams; and two student projects (where students individually collect and analyze a set of data and write up a full report).

I collect informal evaluation information about the course in several ways:

1. I usually have students complete a mid-quarter survey which asks about their reactions to the course and its nontraditional format (i.e., no lectures, reading the book before coming to class, and working in groups on problems during class time). Their responses are 90% enthusiastic, with comments such as: "I never knew statistics could be fun," "I can't believe I actually enjoy this course," etc.
2. I use a standard student evaluation form at the end of the quarter which indicates generally favorable student reactions. I know this is not just a teacher effect because I have trained six different graduate students to use my materials and methods, and students have expressed high levels of satisfaction in those sections of the course as well, even with two foreign students who often get lower ratings in traditional lecture classes.
3. I use what I call "Real Life Problems." These are structured written reports of student projects where students have gathered and analyzed a set of data. Although this is used as a learning activity, it also provides a useful assessment of students' ability to use the statistical language appropriately, choose appropriate methods of analysis, interpret summary measures and graphs of data, and draw reasonable conclusions based on their analysis. As an assessment method, it helps me see in which areas students are weak and may need additional work. Usually I find some consistent errors that then lead me to discuss these in class or structure additional activities for students to work on in groups.

3. Two Kinds of Introductory Courses (Gudmund Iversen)

Not all students have the same statistical needs. Some students will be actively collecting data in laboratory experiments or in small sample surveys and analyzing these data for course projects and possible theses. Other students will be analyzing secondary data for similar purposes. A large group of other students will be exposed to statistical results and conclusions in their readings of books and journals for a variety of different courses as well as in their reading of newspapers and news magazines. But these students will never be asked to collect or analyze data on their own.

The premise of this dichotomy leads to the construction of two kinds of (non-calculus level) introductory courses. A possible name for the first type of course is Statistical Methods, and this is essentially the type of introductory course commonly taught across the land. A possible name for the second type of course is Statistical Thinking, and this is a much more unusual course, even though varieties of this course are beginning to appear.

Having tried this distinction for the last five years, my own evaluation is that it has been very successful. When only Statistical Methods was offered, the students were much more mixed in their expectations, their interest in the material, and in their abilities to do the work. Now that the students have a choice, the Statistical Methods course draws students who are motivated to learn statistics because they know they need these methods for their work in other courses. Sometimes this motivation comes late, and they take the course as late as their senior year. But by that time they are often involved with their own data, and they really see the need for knowing statistics. The enrollment in this course is now larger than it used to be, and it is now typically taken by about 20% of any graduating class.

The goal of the Statistical Thinking course is for students to understand and critically evaluate statistical results they are exposed to both in their academic work and in their non-academic lives. This means they have to know about data collection, analysis, and inference, but they do not need to have the technical knowledge necessary to do any of these themselves. In practical terms this means the students do not have to see and study formulas.

For example, instead of ever seeing the formula for the correlation coefficient r, students are shown scatterplots with different degrees of correlation and are told that we can compute a number which ranges from zero to one, the value depending on how close the points are to a line through the points. Giving the r-value for each plot and seeing the degree of scatter that goes with each r teaches non-mathematically inclined students more about the strength of relationships than any study of the formula for r can do.

In the study of data collection the need is stressed for proper statistical random samples and randomized experiments. The students quickly catch on to the fact that self-selected samples and convenience samples cannot be used for generalizations. In the data analysis part of the course the focus is on simplifying the data so we can understand what they tell us without the loss of too much information. Graphical methods, with the use of overheads showing good and bad graphs, is a popular topic. Computational methods are harder to deal with, but it is still possible to explain about mean and standard deviation without writing down formulas.

The inference part centers around p-values, both for tests for single variables and tests for relationships between variables. The technical details of whether we need to compute z, t, χ-square, or F are interesting for statisticians but only of minimal interest and concern for a person with interests in the substantive matter at hand. These quantities are only tools used to find the p-value, which tells the important story about the data. Now that computer software routinely supply us with p-values, a student needs only to understand the proper interpretation of this concept and to know where to look on the output to find the actual value.

The last section of the course consists of a discussion of the role of statistics in today's society. It stresses the need for information and the role it plays for an individual as well as for society in its attempt to determine public policy.

An important part of the Statistical Thinking course consists of writing papers. The first paper deals with some aspect of statistics, based on what has been discussed in the course up to that point. A popular topic consists of taking an issue of a magazine like *Newsweek* and discussing the quality of some of the graphs found in that issue. Other topics include the role of statistics in some particular field. The second paper deals with the study of the relationship between two variables; how strong the relationship is, whether it is statistically significant, and whether it is causal or not.

The hope is that the course gives an overview of the role of statistics without getting bogged down in technical details. Some students have taken this course first and gone on to the Statistical Methods course or a mathematical statistics course. They report that they find the second course particularly interesting because they already have some sense of the underlying issues, and they see the need for the particular statistical methods in a new and informative light.

4. Project NABS: Nuts and Bolts Study (Dick Gunst)

This classroom experiment is an introduction to statistical problem solving. It encompasses concepts of quality improvement, the need for data (information), variability, group/team solutions, experimental design, graphical displays, and the scientific method. It requires approximately 30 minutes but can take up an entire 1 1/2 hour lecture depending on how much the instructor wishes to discuss the individual issues raised by the experiment.

Requirements.

1. At least four identical jars are used for the first phase of the experiment. The jars are approximately 7 inches tall and are not wide enough for an adult's hand to fit through the mouths. Each jar is filled with nuts, bolts, and washers. Most of the nuts in the jars fit the bolts, but there is a wide assortment of sizes in each of the jars and there are some nuts that do not fit any of the bolts. Refer to these jars as the "mixed" jars.
2. At least four stop watches, the number equal to the number of mixed jars.
3. At least four additional jars, the number again equal to the number of mixed jars. Half of the jars are identical to the mixed jars. The other half of the jars are smaller, with mouths the same size as the tall ones. Approximately half of each size jars are filled with an assortment of large bolts and their corresponding nuts. The remaining jars of each size are filled with an assortment of small bolts and their corresponding nuts. No washers are included in any of these jars and each nut fits a bolt. Refer to these jars as the "sorted" jars.

Setting. The class members are employees of or consultants to a company that manufactures nuts, bolts, and washers. One very large group of its customers requires that the bolts it purchases be shipped with the corresponding nuts securely fastened. In recent years the company has been losing customers because its competitors have been able to satisfy these customer requirements at a lower cost. Management has mandated that costs be reduced.

Costs of the raw materials and machinery to manufacture the nuts and bolts are standard industry-wide. A cost analysis of the process indicates that the only manufacturing component in which the company's competitors could be achieving lower costs is in the assembly process, the fastening of the nuts to the bolts. In particular, the amount of time needed to fasten the nuts to the bolts has been determined to be the critical factor in the cost differential between this company's costs and those of its competitors. The employees (consultants) have been assigned to a task force to determine the most effective means of reducing the assembly times.

Assembly Process. At this juncture the students are unable to recommend any cost saving measures because nothing is known about the actual assembly process. Any recommendations made at this point are easily recognized as pure conjecture. A reasonable suggestion (made by the instructor) is to familiarize the task force with the assembly process and to collect some data on assembly times.

Industry-wide, nuts are fastened to the bolts by hand. Automation of this procedure has not been found to be cost effective because the nuts and bolts arrive at the assembly stations in large bins containing a wide assortment of sizes of nuts, bolts, and washers. In order to reduce accidents due to spillage, the nuts and bolts are extracted from the bins and placed into jars. These jars are delivered to the individual workers. The workers remove nuts and bolts from the jars and fasten them together, placing the assembled nuts and bolts into sorted bins in preparation for packaging. Data are needed on the time needed to fasten the nuts to the bolts. Stopwatches are deemed suitable for this purpose.

The students are divided into groups in any convenient way. There are at least two students in each group, with a maximum of about six per group. There are a sufficient number of mixed jars and stopwatches to distribute one of each to each group. The task group members are permitted to familiarize themselves with the operation of the stopwatches. They may also examine the jars of nuts and bolts. As they are doing so, they are informed of the specifics of the assembly procedure. Because of safety reasons, OSHA (The U.S. Occupational Safety and Health Administration) mandates the following procedure:

(a) The nuts and bolts may not be poured from the jars (danger of spillage and subsequent employee accidents). Some portion of the bottom of the jar must always touch the table. The jars may be tilted if the bottom remains touching the table.
(b) Nuts and bolts must be extracted from the jar one at a time. More than one nut or one bolt may be placed on the table or desk and remain out of the jar as matches are sought.

Initial Data Collection and Analysis. One individual in each group is selected as the assembler. Another person is the timer. A third individual is the data recorder for the group. Timing begins as the lid of the jar is being unfastened. It stops when four nuts have been completely fastened to four bolts. Once the time has been recorded, the procedure is repeated with different students acting as the assembler, timer, and recorder. No controls (e.g., the same timekeeper) are allowed so that everyone can participate and so the uncontrollable variability due to different assemblers and timers will be evident in the results. If there are at least four groups, it is not necessary to repeat the experiment further.

One of the main features of the analysis of data from this experiment is its simplicity. A point stressed is that often only simple data analyses are required to extract important information from experimental results. This is especially true if the data are collected in a well-designed experiment that includes satisfactory controls over extraneous variability.

As the data are reported, they are graphed using a simple point plot (dot plot) of the assembly times. This is done on an overhead projector or on a blackboard. Usually the times range from about 50 to 200 seconds, although occasionally an observation is higher than 200 seconds.

The data are usually widely scattered. Comments are made about how the experiment was conducted under controlled conditions (one classroom, same time of day, same instructions and training, etc.) and yet the data vary widely. The concept of variability is stressed in this discussion. Students are alerted to the potential difficulty in interpreting data that exhibit such variability, motivating their presence in the course.

Causes/Solutions. Next students are asked to critique the assembly process, with special attention to improvements that would reduce the assembly times. With very little encouragement, numerous ideas are suggested, occasionally peppered with a sense of frustration! In this "brainstorming" session, all ideas are sought and are written on an overhead transparency or on the blackboard. Time permitting, these ideas are organized (e.g., cause and effect diagram) and prioritized.

In virtually every session, three of the ideas for improvement are:

- sort the sizes of the nuts and bolts,
- select a different size for the jars, and
- eliminate the washers.

While all the ideas are praised, these three are selected as especially noteworthy. The remainder (e.g., can't pour nuts and bolts from the jars, training, time of day, differences in stopwatches, differences in assemblers and timers, heights of desks, etc.) are discussed in terms of continuous quality improvement and the scientific method, with the end result that some will be selected in further rounds of experimentation.

Primary Experiment. The primary experiment consists of a factorial experiment in two factors: jar size and nut/bolt size. Four or more additional jars are introduced, as described above in Requirement 3. The jars are randomly distributed to the groups. Prior to this phase of the data collection, students are asked their expectations. Most state that the times will be reduced with the smaller jars. Some also believe the larger nuts and bolts will reduce assembly times.

As times are reported, they are again plotted. A single point plot is constructed beneath the point plot for the initial data. On the second plot, data from each of the four combinations of the two factors are plotted using different symbols. Two things typically occur. First, the four plotting symbols overlap, revealing that no combination is clearly preferable to any of the others. Second, contrary to any of the student conjectures, *all* the plots show reduced assembly times; i.e., there is a downward shift in the second set of plots from the initial assembly times plot.

Conclusions. So long as some sorting is performed and the washers are eliminated, the assembly times will typically be reduced. The company can accrue additional cost savings by selecting the jar size that minimizes cost. These benefits are the result of simple experimentation and simple data analysis, both of which are based on sound statistical principles.

Appendix C: Making It Happen

Interesting and Available Data (Robin Lock, St. Lawrence University)

The purpose of this column [in the *Statistical Computing and Statistical Graphics Newsletter*] is to provide a means for sharing the sorts of interesting data that stimulate, challenge, and enliven the practice of statistics. ... Contributed data sets will be made available through the StatLib file server at Carnegie Mellon University. For example, to obtain the baseball data send a message which contains only the line

```
send baseball.data from datasets
```

to the address

```
statlib@lib.stat.cmu.edu.
```

A list of all data sets currently available can be obtained by sending the message

```
send index from datasets
```

to the same address. General information about StatLib is available by sending the message

```
send index
```

to the StatLib address. (From *Statistical Computing and Statistical Graphics Newsletter*, April 1991.)

EDSTAT-L: Statistics Education Discussion List (Tim Arnold, North Carolina State)

Announcing a new electronic forum for discussion related to statistics education: the purpose of this list is to provide a forum for comments, techniques, and philosophies of teaching statistics. The primary focus is that of college level statistics education, both undergraduate and graduate studies.

This list exists because we can learn from each other. What techniques are we using in statistical instruction? What strategies should we use to prepare our students for the future? What part do or should computers play in instruction? Are computers a new tool or a new subject for students? What assets (e.g., public-domain programs, data sets) can we share among ourselves? There are many individuals working in these areas, each with knowledge valuable to the others. This list attempts to bring together every teacher, student, researcher, and specialist interested in improving statistical instruction.

You are encouraged to:

- Submit summaries of or commentary on published articles and books which address problems or solutions in teaching statistics.
- Share the results of your experiments in teaching statistics.
- Ask thought-provoking questions about the future of statistics education and the environments in which it takes place.
- Recruit interested parties to this list, regardless of their status in education, industry, or government.

Subscribing to EDSTAT-L

Bitnet Users: Send the following command as the first line of a mail file to `listserv@ncsuvm`:

```
subscribe EDSTAT-L Your full name
```

IBM VM users in Bitnet may add themselves to the mailing list with this command:

```
tell listserv at ncsuvm subscribe EDSTAT-L Your full name
```

VAX/VMS users in Bitnet can subscribe in a similar way:

```
send listserv@ncsuvm subscribe EDSTAT-L Your full name
```

Internet Users: Send the following command as the first line of a mail file to `listserv@ncsuvm.cc.ncsu.edu`:

```
subscribe EDSTAT-L Your full name
```

Appendix D: Report of Workshop on Statistical Education

by Robert V. Hogg, UNIVERSITY OF IOWA

On June 18-29, 1990 thirty-nine statisticians gathered for a workshop on statistical education at the University of Iowa in Iowa City, Iowa. These persons represented universities, colleges, consulting firms, business, and industry. This workshop was sponsored by the University of Iowa and the American Statistical Association (ASA). Financial assistance was provided by the National Science Foundation, the ALCOA Foundation, the Ott Foundation, the Statistics Division of the American Society of Quality Control, and the Quality and Productivity Section of ASA.

As we prepared for the workshop, most of the participants, and several others not attending, wrote position papers on some aspect of statistical education, the majority of which concerned a first course in statistics. As a group we recognized several poor characteristics of science and mathematical education, including statistical education:

- Improvements are needed in the K–12 curriculum as there is widespread science and mathematics illiteracy in the United States.
- There is little effort given to recruiting students to these areas, often because we want only students like ourselves.
- There is a lack of truly qualified teachers because the pool from which they come is drying up.
- College and university instructors are often required to teach large introductory classes that allow little or no interaction with students and severely limit student involvement with others.
- Frequently students view courses as being "hard" because the class periods are dull and it is difficult to get good grades despite spending long hours doing homework.

Reprint of a report of a workshop on statistics education, especially at the introductory level, organized by Roberg Hogg and held June 18-20, 1990.