Reconsidering Statistics Education: A National Science Foundation Conference

George W. Cobb

Reconsidering Statistics Education: A National Science Foundation Conference

George W. Cobb
Mount Holyoke College

Abstract

1 Recent survey data demonstrate an acute need for curricular resources in statistics. The first half of this paper summarizes and compares a dozen current or recent NSF projects, most of which are developing such resources. (As an aid to interested instructors, an appendix gives more detail on the individual projects, along with a list of available files that provide even more detail.) Nearly all these projects involve activities for statistical laboratories, at least implicitly, although the labs are used in a variety of ways: for analysis of archival data sets, for hands-on production of data for analysis, and for simulation-based learning. These three kinds of labs are compared in terms of their complementary sets of advantages.

2 This paper grows out of a small conference which brought together NSF Program Officers, Principal Investigators and Co-PIs of the projects, and a half-dozen other teachers of statistics. The second half of the paper develops four themes from the conference: (1) Questioning standard assumptions, (2) Resistances to change, (3) Total Quality Management, and (4) Educational Assessment. These themes are (a) offered (modestly) as useful guides to thinking about teaching statistics, then (b) exploited (shamelessly) to argue for a scorched-earth assault on our introductory courses.

1. Introduction

```
3      ||   The need for curricular resources in statistics is     ||
       ||   acute, arguably more acute (at the college level)      ||
       ||   than in any other subject.  The reason:  Of all        ||
       ||   subjects taught as often as statistics, surely no      ||
       ||   other subject is so often taught by faculty with so    ||
       ||   little formal training in the subject.                 ||
```

4 In keeping with the tenets of our subject, we should insist on data to support statements like the pair above. Accordingly, here are two clusters of data, corresponding to how often statistics is taught, and by whom. First, how often. In the fall of 1990, according to a recent survey by the Conference Board of the Mathematical

Sciences, departments of mathematics and statistics at four-year colleges and universities taught over 3000 sections of elementary statistics, to almost 120,000 students (Albers et al. 1992, pp. 22, 24). These 1990 enrollments are the result of decades of steady growth. At four-year colleges and universities, enrollments in elementary statistics doubled between 1970 and 1990 (*ibid*., p. 127), while at two-year colleges enrollments more than tripled (*ibid*., p. 86). One particularly striking view of the growth comes from comparing enrollments in statistics and calculus courses at two-year colleges, as in Table 1. In 1966, there was only one section of statistics for every ten sections of calculus; by 1990 there were five.

```
TABLE 1.  Statistics Enrollments as % of Calculus
    Enrollments at U.S. Two-Year Colleges
    ==================================================================
    Year         1966      1970      1975      1980      1985      1990
    Percent       10        19        37        27        36        52
    ==================================================================
    (Computed from enrollment figures reported in Albers, et al., 1992, p. 85.)
```

5 Who are the teachers of all those thousands of sections of statistics taught each year? There is considerable evidence to suggest that only a comparatively small fraction of the sections are taught by statisticians, or even by others with substantial and recent training in the subject. According to the CBMS survey, for every section of elementary statistics taught in a statistics department, there were almost five sections taught in mathematics departments: In fall 1990, only 494 sections were taught in departments of statistics, compared with 2601 sections in departments of mathematics (*ibid*., p. 24). These data are for universities and four-year colleges; had data from two-year colleges been available to include in these totals, the imbalance would be even more extreme. Of course one must not assume that all sections taught in departments of mathematics are taught by non-statisticians. The CBMS survey, unfortunately, provides no data on the statistics backgrounds of faculty in departments of mathematics. But an older, less systematic survey reports that among 80 responding liberal arts colleges, only 25 could claim a Ph.D. statistician on their faculty (Moore and Roberts 1989). Additional data come from the 1993 applicants to a series of NSF-funded statistics workshops for mathematics faculty. (The workshops are part of the Mathematical Association of America's Project STATS: Statistical Thinking and Teaching Statistics. See Landwehr 1993.) These week-long workshops, explicitly designed for faculty without training in statistics, drew over 150 applications for 1993, almost all from mathematicians who are called on by their departments to teach statistics. Roughly half the applicants list no post-baccalaureate work in statistics. Among the others, a typical background would include one or two graduate courses, more often theoretical than applied, and even in the case of an applied course, often taken prior to the sweeping changes that cheap computing has brought to statistics in the last decade.

6 Although these data do not address the teaching of statistics in the social sciences and engineering, nevertheless it seems a safe extrapolation to suppose that, for the foreseeable future at least, a very substantial portion of beginning statistics instruction will remain in the hands of mathematicians. Statistics, however, is fundamentally different from mathematics, and can be taught appropriately only by someone who recognizes and understands that difference (Moore 1988, 1992). To the extent that teachers of statistics have not been trained as statisticians, they must rely on the work of others if they are to do justice to the subject: hence the acute need for curricular resources.

```
7     ||   Since 1986, the National Science Foundation has      ||
      ||   made more than 190 awards, totaling $9.6 million,     ||
      ||   in support of statistics education, primarily at      ||
      ||   the post-secondary level.                             ||
```

8 All twelve of the curricular projects described in this article offer useful ways of thinking about statistics

teaching, and almost all are developing resources (data sets, hands-on laboratory activities, computer simulations) which any instructor could use to make the introductory course more effective. Some of the projects are developing resources which are also appropriate for more advanced or specialized courses. (These projects are being conducted with support from the Division of Undergraduate Education with some participation from the Division of Mathematical Sciences.)

9 Experience with projects in other subject areas has suggested that it is valuable to bring project directors together to a conference where they can share their ideas, plans and experiences at a stage when the exchange can still have an effect on the developing projects. This article is in large part a result of such a conference, which was held June 28-29, 1992, in Washington, DC, with three dozen statisticians and NSF officers in attendance. Most of the participants are (or were) Principal Investigators or co-PIs for the dozen NSF projects listed and discussed later in this report. The conference also included another half-dozen teachers of statistics, each the inventor of an unusual course or distinctive approach. The conference was organized by Joan Garfield (University of Minnesota) and myself, with assistance from Nell Sedransk and William Haver, who were at the time both NSF Program Officers.

10 To maximize the chance for genuine exchange, the conference schedule included few formal presentations, but many small-group discussions. These later discussions were planned as variations on a sequence of four main themes: (1) Questioning the standard assumptions, (2) Anticipating resistances to change, (3) Total Quality Management, and (4) Assessment of student learning. Happily, despite the orchestration written into the score, the discussions themselves evinced healthy improvisational and aleatory elements.

11 With almost no formal presentations at the conference, there is nothing that might be serve as the basis for the usual sort of conference proceedings, but this article is offered as a substitute of sorts. In what follows, I shall first (Section 2) describe and summarize the projects themselves, then (Section 3) discuss the four conference themes. Section 2 is intended to provide general information about some of the work on curriculum that the NSF has been funding, while at the same time serving the practical purpose of helping readers to decide which of the various sets of resources now being developed are most relevant to their own teaching. Some of the projects already have materials available for field testing, and others are planning eventually to distribute their materials to interested teachers. For readers who want more detail than Section 2 provides, an appendix presents paragraph-length descriptions of the individual projects, with information on how to reach a director of the project. For many of the projects, still more detail is available in supplementary files; the appendix lists these files and gives instructions for obtaining them.

12 Section 3, on the conference themes, is presented not as a report of our discussions, but rather as one set of interrelated approaches for thinking about the elementary curriculum. Two of the themes, total quality management and educational assessment, have become notoriously fashionable, and together they are already guilty of having incited dozens of conferences and tens of thousands of published pages. It is my hope that by focusing narrowly on their use for thinking about curriculum, the present contribution to their growing burden of guilt will be a modest one.

2. Summary of the Projects

```
13      ||   Nearly all the NSF-funded projects involve         ||
        ||   statistical laboratories:  for analysis of archival  ||
        ||   data, or for hands-on production of data, or for    ||
        ||   simulation-based learning.  The prominence of the   ||
        ||   lab approach accords with the movement of           ||
        ||   statistics back towards its roots in science, and   ||
```

```
||   with research in education that demonstrates the      ||
||   importance of active learning.                        ||
```

14 The projects are summarized and compared in Tables 2 and 3. Of the dozen projects represented at the conference, ten are similar enough to be discussed together. By happy coincidence, the other two projects have already been presented in detail elsewhere, so I shall discuss each only briefly before turning to the remaining ten.

15 *The Hogg Workshop*. The first of the two exceptions differs by virtue of being a meta-project: a conference about statistics education in general rather than a curricular project tied to a particular course. The conference was organized by Robert Hogg of the University of Iowa, and brought together 39 statistics educators for three days of work in teams on various aspects of statistics teaching. The appendix contains a brief summary of the conference report; a longer summary appeared in *Amstat News* (Hogg 1990), and the full report has been included in a recent volume of the MAA Notes series (Hogg 1992). The report includes four hortatory imperatives for beginning statistics courses: (1) state the course goals, (2) analyze data and do projects, (3) use computers (for most courses), and (4) lecture less, teach more.

16 *A course called CHANCE*. The second of the exceptions, the Chance project of J. Laurie Snell from Dartmouth and a number of colleagues at other institutions, is like the other ten projects in that it does deal with a particular course, but the course itself is unlike any of the others in that it has no particular technical agenda. "We do not intend for Chance to replace any statistics or probability course; its aim is rather to encourage students to think more rationally about chance events and to make them more informed readers of the daily press. ... We try to choose topics that are currently in the news and are likely to remain so. What is the evidence that HIV causes AIDS? Should you lower your cholesterol? Should you believe opinion polls? How reliable is DNA fingerprinting? How should we adjust the 1990 census?" (Snell and Finn 1992, pp. 12-14).

17 The other ten projects, in one way or another, all seek to improve the teaching not just of statistical thinking but of statistical methods as well, by involving students actively in the practice of statistics as science. Despite this common thrust, however, the projects exhibit a lot of variety. Some have a particular applied focus (Ikem -- economics; Gilliland, Nelson -- engineering). Some are linked to particular software or hardware (Magel, Spurrier -- Minitab; Notz/Pearl/Stasny/Velleman -- Macintosh). Some are largely local in focus, with emphasis on establishing computer facilities (Ikem, Magel, Tichenor) or on a particular course (Nelson). Others are quite ambitious in their plans for dissemination on a national scale (Notz/Pearl/Stasny /Velleman, Scheaffer, Spurrier, Trumbo), with substantial supporting materials under development, and in some cases elaborate field-testing already underway. Some projects are deliberately modest in their use of technology. For example, most of the activities in Scheaffer's and Spurrier's projects require no fancy equipment, and some can be completed without using a computer. Trumbo's data sets, though intended for computer analysis, will be available in a format that is independent of both hardware and software. Other projects, with equal deliberation, seek to take fullest advantage of the impressive technological resources of major state universities (Meeker -- workstations, high-resolution graphics, statistical programming languages at Iowa State; Notz/Pearl/Stasny/Velleman -- statistical software integrated with videotape and compact disk technology at Ohio State).

18 The best way, I think, to appreciate both the unity and variety of these ten projects is to look at what they do with data. In the rest of this section, I consider three different sources of data, and discuss the pedagogical advantages and disadvantages of each source. Although many of the individual projects have quite naturally chosen to emphasize one source over the others, it would be unfair, and certainly is not my intent, to associate the shortcomings of a particular data source with the projects I mention in connection with that source. Rather,

I hope to illustrate that an ideal curriculum should rely on a variety of data sources in order to enjoy a mix of advantages.

```
TABLE 2.  NSF-Supported Projects in Statistics Education, 1990-92
=========================================================================
Project director/                             Stat. courses involved
Contact person          Principal focus       (Other depts. affected)
--------------------    ------------------    -----------------------
Dennis C. Gilliland     Computer simulations  Quality and
Michigan State Univ.    and lab activities for productivity
East Lansing, MI 48824  teaching concepts     course (Engineering)

Robert V. Hogg          Workshop on           Introductory course,
University of Iowa       Statistical Education statistics teaching
Iowa City, IA 52242                           in general

Fidelis Ikem            Multimedia classroom, Introductory course,
Virginia State Univ.    computer laboratory,  econometrics
Petersburg, VA 23806    minority students     (Economics)

Rhonda C. Magel         Undergraduate         Introductory,
North Dakota State      computer laboratory   regression, design
University                                    & ANOVA, nonpara-
Fargo, ND 58105-5075                          metrics

William Q. Meeker       Instructional         Time series, research
Iowa State University   simulation-based      methods, multivariate,
Ames, IA 50011          software modules on   quality control,
                        basic concepts        design

Peter Nelson            Enriching engineering Existing engineering
Clemson University      labs by incorporating laboratory courses
Clemson, SC 29634-1907  statistical concepts  (Engineering)
                        and methods

William I. Notz         Integrating software  Introductory courses
Ohio State University   and multimedia tech.
Columbus, OH 43210      into encyclopedia of
                        examples and exercises

Richard L. Scheaffer    Curriculum based      Introductory courses
Univ. of Florida        on lab activities
Gainesville, FL 32611   and exercises

J. Laurie Snell         Course based on       Mathematics for general
Dartmouth College       statistics and        education
Hanover, NH 03755-1890  probability in the
                        news

John D. Spurrier        Lab activities and    Introductory courses
University of           exercises for a one-
South Carolina          credit statistics lab
Columbia, SC 29208      course

Dolores M. Tichenor     Microcomputer         Upper division probability
Tri-State University    laboratory,           and statistics sequence
Angola, IN 46703        real datasets         (Engineering)

Bruce E. Trumbo         Collection of real    Introductory courses,
California State        datasets and          statistics courses in
```

University   supporting materials   general
Hayward CA 94542-3087
===

TABLE 3a.   Features of the Projects: Main Source of Data

Director	Hands-on activities	Real-world studies	Computer simulations
Gilliland			YES
Hogg			
Ikem		YES	
Magel		YES	
Meeker			YES
Nelson	----Engineering Labs------		
Notz		YES	
Scheaffer	YES		
Snell		YES	
Spurrier	YES		
Tichenor		YES	
Trumbo		YES	

===

TABLE 3b.   Features of the Projects: Special Concerns

Director	Student assessment	Controlled comparison	Minority students	Engineering	Large classes
Gilliland				YES	
Hogg					
Ikem			YES		
Magel				YES	YES
Meeker					
Nelson				YES	
Notz		YES	YES		YES
Scheaffer	YES				
Snell					
Spurrier	YES		YES		
Tichenor				YES	
Trumbo					

===

TABLE 3c.   Features of the Projects: Teaching Resources Produced

Director	Datasets	Teacher's manual	Workbook/ lab manual	Computer software	Electronic news serv.	Extensive biblio.
Gilliland			YES	YES		
Hogg	----------conference report and recommendations-----------------					
Ikem						
Magel	YES					
Meeker			YES	YES		
Nelson						
Notz	YES	YES	YES	YES		YES
Scheaffer		YES	YES			
Snell					YES	YES

```
Spurrier                   YES        YES
Tichenor
Trumbo      YES      YES     YES                                    YES
==========================================================================
```

2.1 Hands-on activities

19 Richard Scheaffer's project is developing an entire activity-based introductory curriculum, one which emphasizes statistics as a lab science. Examples of activities include analysis of survey data from the class, simulating capture-mark-recapture with goldfish crackers, randomized response surveys, sampling activities with strings and rectangles, the Gunter-Ortiz funnel experiment (response surface design), and many others. Where appropriate, the activities come with suggested extensions of the ideas to the real world, such as, for example, the application of capture-mark-recapture methods to the census undercount.

20 Although John Spurrier's project relies on computers for data analysis, an inventive variety of schemes for hands-on data production uses equipment no fancier than inexpensive measuring devices. Pulse rates are used to teach descriptive statistics and variability. Traffic counts illustrate time series, taste comparisons provide binomial data, and perceived versus actual distances provide data for scatterplotting. Other labs deal with the absorbency of paper towels, the breaking strength of string, the flight distance of paper airplanes, and the prediction of hickory nut weights.

21 Data collected by a statistics class about themselves might be regarded as intermediate between a hands-on activity and an archival data set. Rhonda Magel's project began with the expectation of relying mainly on archival data from a variety of client disciplines, but has shifted toward greater reliance on data collected in class, which "is more relevant to the students, such as the amount of time they spend studying, watching TV, listening to music, the type of soda pop they prefer, or whether or not they go to a fast food restaurant more than once a week."

22 *Advantages of hands-on activities*:

- 23 The resulting data are fresh, not someone else's leftovers.
- 24 Students are actively involved in data production, which is otherwise likely to be slighted in a beginning course. Potentially, at least, there are four distinct advantages here:
- 25 The student who has a hand in creating a data set is nearly always motivated to analyze it.
- 26 Experience with variability is immediate and concrete.
- 27 Some studies force students to deal with coding issues. ("What do I do with a Yes/No question answered `Maybe'?")
- 28 Sometimes -- though not often enough -- the activities can involve students with design issues.

29 *Disadvantage*:

- 30 Some data sets tend to be toy-like, both in import, and in complexity, at least in comparison with some archival data sets.

2.2 Real-world data sets

31 Four of the projects rely mainly on archival data, as opposed to data the students create themselves. Fidelis

Ikem's project uses a micro-economic data base (CitiBase) from CitiBank Corporation. Dolores Tichenor's project relies on local data, from Tri-State University's Department of Biology and Analytical Testing Laboratory, from the Steuben County health department, and from state environmental agencies. Both of these projects are most immediately concerned with establishing a computer laboratory on campus, which statistics students can use for the analysis of data sets. Thus neither of these projects is able to focus exclusively on exportability. Bruce Trumbo's project, on the other hand, includes an explicit emphasis on eventually making his data archive available nationwide. Materials "will include (1) a collection of carefully documented data sets from recent and important scientific research ready for computer use, (2) an instructor's manual, and (3) student workbooks." Data will be available in a variety of formats, including flat ASCII files, which are almost universally readable. The project of William Notz, Dennis Pearl, Elizabeth Stasny, and Paul Velleman relies on a mixture of student-generated and archival data sets. The latter are integrated into an electronic encyclopedia, with data sets linked to descriptions, questions, projects and film clips from the Annenburg/CPB tapes*Against All Odds*, and software (e.g., Data Desk and Minitab) that can be launched from within the encyclopedia to analyze the data.

32 *Advantages of archival data*:

- 33 Many data sets have unmistakable real-world import. No student who spends time with the data like the General Social Survey or the Center for Disease Control's AIDS data will come away thinking of statistics as just a collection of empty numerical rituals. Data sets like these give students a taste of "big time" statistics. An instructor might also choose to flavor a course with a sprinkling of small data sets of a more piquant sort, for example, data on the Transylvania effect (Larsen and Stroup, p. 59), or on law-and-order voters as scofflaws (*ibid*., p. 138).
- 34 Many of the data sets have a complex structure and are rich with statistically interesting features.
- 35 Materials to support particular curricular uses can be prepared in advance.

36 *Disadvantages*:

- 37 Students are not involved in any aspect of data production.
- 38 Analysis of the data takes students over ground that has already been covered by others (Bentley 1993).
- 39 Random variability remains implicit, hidden in the invisible past. (Student-generated and simulated data make variability more palpable.)

2.3 Simulation-based learning

40 Dennis Gilliland's project uses simulation to teach the statistics of quality and productivity to engineers. One module "illustrates the fundamentals of elementary control charting. Students choose from several production processes, make runs, and perform capability analyses. Some processes are in control, some have trends in mean, others have trends in variance, and others are chaotic. [Another] module illustrates the importance of organized experimentation. ... Students choose from several processes with three predictor variables and are asked to explore the response surface. ... [Initially] they generally do inefficient, one-at-a-time experimentation. ... After lectures on the design of experiments, students interact with the processes more intelligently and see the immediate impact of their education."

41 William Meeker's project is similar in spirit, but planned to affect a great range of undergraduate and graduate courses, with a more explicit focus on simulation to help the student learn concepts, especially those "that require visualizing in more than two or three dimensions e.g., multiple regression surfaces) or that

require theory beyond the mathematical abilities of the students."

42 *Advantages of simulation-based learning*:

- 43 The instructor can control the structure of the data, and thus the conceptual focus of the exercise.
- 44 Repeated simulations give students direct experience with variability.

45 *Disadvantage*:

- 46 The data, no matter how realistic, are not real.

2.4 The best of two worlds?

47 *Dr. Nelson's statistical house call*. Uniquely, among all the projects, that of Peter Nelson embeds statistical learning within a course in a client discipline. His project teaches statistical thinking and methods to engineering students at Clemson University in the context of their labs in an existing engineering course. This approach combines the advantages of hands-on activities (2.1) and real-world data sets (2.2). Students create their own data sets, which gives them added motivation to do the analyses. Moreover, because the data come from an experiment whose main purpose is to illustrate important engineering principles, the data carry a sense of real import that statistics courses ordinarily can achieve only with archival data. In view of the way Nelson's curricular visit to the client's home turf brings with it these two sets of advantages, it is unfortunate that the departmentally-based organization of colleges and universities makes this kind of collaboration so rare.

48 It is surprising that none of the NSF-funded curricular projects involve student data-gathering-and-analysis projects. As many recent authors point out, projects can combine the advantages of hands-on planning and production with the possibility to gather more complex data and investigate questions of greater import than is realistically possible in a single afternoon lab. (See Halvorsen and Moore 1992, Roberts 1992, Witmer 1992, Bryce 1993, McKenzie 1993, Sylwester and Mee 1993, and Zahn 1993b.) I shall return to projects briefly in a later section.

49 Curriculum reform, like politics, involves an uncertain mix of pragmatism and idealism. To make substantial progress, we need both a vehicle that runs and a distant objective to aim for. We must be realistic in judging how fast we can change, but at the same time ambitious in our choice of direction and ultimate destination. The projects I have been discussing are innovative, but at the same time appropriately practical: they are all constrained, to a greater or lesser extent, by the quaint expectation that they should actually deliver something concrete during the finite term of their funding. Moreover, all ten of the projects that promise to teach technical content are based at universities, nine of them public, most of them large -- places where restrictions tend to be greater, and the climate less hospitable to the more adventurous kinds of curricular experimentation that are often encouraged at small liberal arts colleges. In the next section, starting from some examples of (unfunded) projects at such colleges, I suggest a variety of devices we can use to think ambitiously about the long-term direction of reform.

3. Four Approaches to Rethinking the Beginning Course

3.1 Questioning the standard assumptions

```
50      ||   The usual introduction to statistics is (a) a      ||
        ||   survey course, (b) organized by statistical methods ||
        ||   and concepts, which are (c) presented in a standard  ||
        ||   order, (d) with the instructor doing almost all the  ||
        ||   talking.  Although these four features are           ||
        ||   characteristic of introductory courses in the        ||
        ||   sciences, there are now many courses whose           ||
        ||   existence demonstrates that not one of the four is   ||
        ||   essential.                                            ||
```

51 (a) *Introductory statistics need not be taught as a survey course.* As David Moore has written, "If I use regression 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" (quoted in Cobb 1992). In fact, Robin Lock has for several years now been teaching applied time series analysis as a first course in statistics at St. Lawrence University (Lock 1990). At Carleton College, Frank Wolf developed an introduction based on multivariate descriptive statistics (Wolf 1990). At Mount Holyoke College, we have two introductory courses: one on experimental design and analysis of variance, with emphasis on experimental data in psychology and biology, and an alternative introduction based on applied regression analysis, with emphasis on observational data. The idea for a pair of introductory courses along these lines is one we borrowed from Gudmund Iversen at Swarthmore College.

52 (b) *A first course need not be organized by statistical topic*. Instead, the structure of the course can come from a series of applied questions, as in the Chance course. Courses organized by statistical topic, with examples trimmed and tailored to fit the method *de jour*, risk encouraging students to drift into a Cinderella approach to data analysis: each method takes on the mystique of a glass slipper, which the data must be forced to fit lest the kingdom be lost. Arthur Dempster distinguishes two attitudes that are relevant here, one that regards a given method as fixed, and passes several data sets through its sieve, and a second attitude that takes the data sets as fixed, and uses several analyses to view the data from a variety of perspectives (Dempster 1977). In teaching, one wants data sets to illustrate the methods, of course, but ultimately the correct emphasis should be that a set of methods is used to illuminate each data set, not that the data sets are there to serve the methods. An effective way to instill this attitude is to organize the course as a series of applied problems. This organization is more faithful to the practice of statistics, and can be effective at any of a variety of levels: an introduction to quantitative reasoning for first-year students (Cobb 1990), an introduction to applied statistics at the junior-senior level (Bentley 1993), a data analysis adjunct to a mathematical statistics course (Witmer 1990), or a graduate level research methods course (Willett and Singer 1992).

53 (c) *A first course need not present topics in the standard order*. By "standard order" I mean the familiar sequence that starts with descriptive statistics and leads, more or less directly, first to tests and confidence intervals based on z-statistics, and then to their counterparts based on t-statistics. Fortunately, textbook authors are more and more often choosing to include important side excursions away from the well-worn path. Nevertheless, if one could superimpose maps of the routes taken by all elementary books, the resulting picture would look much like a time-lapse night photograph of car taillights all moving along the same busy highway. For one example of a road less traveled, consider Robert Wardrop's *Statistics: Learning in the Presence of Variation* (Wardrop 1993). This book does not call on teachers to eschew the traditional topics and methods, but by presenting those topics first for the case of dichotomous variables, and postponing until late in the semester the modifications needed to handle multi-outcome and measured variables, his book is able to introduce the main concepts of statistics in their technically simplest form. Moreover, the reorganization makes it possible to involve students from the start in the active planning and collection of data

to answer questions of real interest. Students know enough after the first chapter to plan and execute their own completely randomized study, and enough after the second chapter to use their data to test a simple hypothesis.

54 (d) *A course need not rely on lectures to present the material*. At Purdue University (one of the world's best-known liberal arts colleges!), approximately 200 students per semester elect a course designed by David Moore, based on "mastery learning." In Moore's words, "[the] student progresses by repeated attempts at problems of a certain type until consistent mastery is demonstrated. We use a modification of this method (in a resource room open 28 hours a week and staffed by undergraduate tutors); it is very popular and appears to work well for learning skills and for the challenge of learning what procedure is appropriate. ... I remain unconvinced that this popular system encourages the elusive `higher order learning', `conceptual understanding', etc." (Moore 1993). The Purdue course saves large amounts of faculty time (after a large initial investment), but apparently pays a price in the depth of understanding that students achieve. Other courses can afford to invest more faculty time on a weekly basis, and so can be more ambitious for the students. Joan Garfield describes her approach this way: "In my classes, I do not lecture at all. 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 ... 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" (quoted in Cobb 1992). The appendix gives instructions for obtaining a set of "Sample Course Handouts" that Garfield uses to structure and facilitate group work in her courses. Alan Rossman of Dickinson College is developing materials for a course that is very similar in spirit: "This `workshop' approach eliminates lectures and removes the distinction between lecture and laboratory sessions. Students spend class time working on activities carefully designed to enable them to discover statistical principles and apply statistical techniques for themselves. These activities ask students to analyze and to explore genuine data, some of which come from available sources as well as some collected through in-class surveys and experiments" (Rossman 1993).

3.2 Anticipating resistances to change

```
55      ||   Although we naturally think of teaching and      ||
        ||   learning in terms of the intellect, resistances to  ||
        ||   change tend to be more directly concerned with    ||
        ||   feelings.  To respond effectively, an instructor    ||
        ||   must be attentive to logistics, must be explicit    ||
        ||   about goals and standards, and must have regular,  ||
        ||   reliable feedback from students.                   ||
```

56 As prelude to any analysis of resistances, I first offer a compilation of typical student complaints, compressed into a single gigantic, generic whine: "I wasted three hours last night trying to do one stupid computer problem. It'll probably take me another three hours to track down the people for the survey. What a Mickey Mouse assignment. Why can't they just give us some data? It's too hard to find times for our group to meet outside of class, because Absentia is hardly ever on campus. And Bootless is taking the course pass/fail; that means he'll probably blow off the rest of the group. I need a decent grade to get into law school, so I'll get stuck with most of the work. And writing those reports? Forget it! I didn't deserve a C on the last one -- I knew how to do the formula. What's writing got to do with statistics anyway? I wish we could just have tests like in a regular course. I do OK on those."

57 I suggest that complaints such as these be regarded as symptoms of a problem, but that the "presenting complaint" is not always the same as the disease. (Taking a complaint seriously is not the same as taking it literally.) Most student concerns seem to be of three kinds: concerns about time, concerns about grading, and

concerns about competence. The first two are usually voiced explicitly ("The assignment took too long; the test wasn't fair."), but the third tends to be disguised, often as a concern about time or grading. ("I'm no good at that," is not something we generally say out loud, especially from a lower position in a hierarchy.) Sometimes "It took too long," really does mean "You expected too much work," but sometimes it means "I could have worked a lot more efficiently if I'd had better instructions," and sometimes it means, "There wasn't enough stuff I felt good at to balance all the time I spent feeling stupid and frustrated." Similarly, "It wasn't fair," can mean "You weren't clear in advance about what you expected," but could also mean, "You didn't give me enough chances to show the things I was good at."

58 To deal effectively with these obstacles, the instructor needs to be (1) attentive to logistics and (2) explicit about goals and standards. Many of the concerns can be alleviated or even preempted by careful attention to logistics. If students have not used a particular computer package before (or, more exigently, have never even used a computer before), a field-tested handout with step-by-step instructions will not only save them time, but will also give them a reassuring sense of control as they anticipate their first encounter with the system, and may even spare them the experience of incompetence and frustration that is so often a part of electronic toddlerhood. In this example, attention to logistics ameliorates concerns about time and concerns about competence. Some concerns about fairness can also be handled as logistical problems. For example, groups can be asked at the outset to decide what they will do if someone is not pulling his weight, and to write their decision into a contract (Halvorsen and Moore 1992).
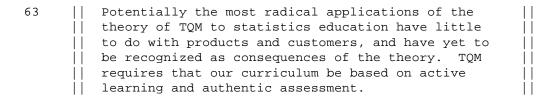
59 Many if not most logistical problems can be handled effectively once we are aware of them, but other problems are more fundamental. Learning necessarily involves doing some things we are not already good at, which makes concerns about competence both natural and unavoidable. To get beyond them, we must first be convinced that the gain is worth the pain. (See Snee 1993.) As the Hogg report (1992) emphasizes, the instructor needs to be clear about goals: Why is this worth learning, and why is this particular approach likely to be more effective than what the students may have been expecting? (What's in it for them?) In a similar spirit, many concerns about grading can be avoided if we are both clear and specific. (Ambiguity breeds anxiety.) Students will feel more in control, and more secure, if they know what we expect from them and how we will evaluate what we get.

60 Even so, no matter how assiduous we are about logistics, no matter how clear about goals and standards, there are sure to be problems we did not anticipate, especially if we are trying a new approach. It is easy, I think, to get misled by a spurious association: complaints tend to be more frequent with a new approach, which tempts us to conclude that students are resisting innovation; worse yet, maybe the new approach is not so good after all. I suggest instead that oversights and unanticipated problems tend to be much more frequent with a new approach, and that we ought to expect (and at any rate have to accept) a period of adjustment perhaps extending over the first several times we try something really different.

61 To smooth the process of settling in, an instructor needs a regular source of reliable information from students. A version of the minute paper (Mosteller 1988) works well for collecting written information in class. Electronic mail now offers other possibilities (Meeker 1993). With luck, information gathered regularly in these ways can alert an instructor to problems while there is still time to do something about them. In addition, for students to see their teacher making changes, in response to their concerns, increases their sense that the course and its innovations are intended to serve their interests.

62 In summary, then, our analysis suggests that very often what appear as resistances to innovation are instead symptoms of other problems often associated with change. Moreover, these problems can be effectively addressed by attending to logistics, by being clear about goals and standards, and by gathering data from students. To anyone familiar with Total Quality Management(TQM), these three strategies are already old

friends.

3.3 Total Quality Management

```
63      ||   Potentially the most radical applications of the      ||
        ||   theory of TQM to statistics education have little     ||
        ||   to do with products and customers, and have yet to    ||
        ||   be recognized as consequences of the theory.  TQM     ||
        ||   requires that our curriculum be based on active       ||
        ||   learning and authentic assessment.                    ||
```

64 I was tempted to sub-title this section on TQM "Rising tide of progress, or rising gorge of disgust?" in order to acknowledge the only sharp division of opinion at the conference. In general, those with liberal arts backgrounds tended to be the ones less interested in TQM and more skeptical about it; those with substantial experience teaching statistics to students of business and engineering knew a lot about TQM and were enthusiastic about it. Here is a summary of the skeptic's position: "If you try to apply the TQM vocabulary, you find that identifying the customer and the product is full of ambiguity. Is the student the customer (which makes the course the product), or is the student the product (which makes his future employer or next teacher the customer)? This ambiguity-of-fit is just a symptom of a larger problem. After all, TQM developed as a way to improve the quality of manufactured products. Why should we expect it to be of any use for truly deep thinking about teaching and learning? An education, unlike a car or a microwave, is supposed to change the person who receives it; ideally, both the education and its human host continue to develop over the course of a lifetime. To impose an industrial template is to caricature education as an assembly-line process that stamps out one class after another of little formed intellects." Such skepticism, to adherents of TQM, is largely a rationalization to cover faculty conservatism. As David Moore put it, "Faculty don't think quality management (especially reduction of unplanned variation) applies to them. Neither did auto plant managers. Both are wrong." The conference ended and we dispersed without finding a way to resolve our differences of opinion.

65 For those unfamiliar with TQM, I first present a thumbnail characterization, then outline some of TQM's various connections to statistics, ending with a brief summary of ways TQM has been applied to improve the teaching of statistics. Against this background, I then argue that we have yet to explore fully the deeper implications of TQM for curricular reform.

66 What is TQM? At the conference, David Moore gave the best short summary I know, distilling TQM as practiced in industry to a tripartite essence:

- 67 Emphasis on data. (If you don't know what to measure, measure anyway: you'll learn what to measure.)
- 68 Process orientation. (Constancy of purpose is crucial; improvements must be institutionalized.)
- 69 Customer focus. (We must abandon the notion, "I know best, I'm the expert.")

70 I find it useful to distinguish four ways to think of TQM in relation to statistics: (a) as an application, (b) as a theoretical frame, (c) as a method for improving course presentation, and (d) as an approach to thinking about curriculum.

71 (a) *TQM as an application of statistics*. Statistical process control is a good illustration. In the present context of curricular reform, this view of TQM is neither controversial nor particularly relevant.

72 (b) *TQM as a theoretical frame for statistical practice*. By frame I mean something similar in scope and

spirit to the decision-theoretic frame provided by the Neyman-Pearson-Wald theory. Here is a one-sentence illustration that suggests the main idea, but without doing justice to the possibilities for depth: One can think of statistics as embedded within the Deming cycle of Plan-Do-Check-Act. For an extended and thoughtful treatment in an introductory statistics textbook, see Cryer and Miller (1991).

73 (c) *TQM as a method for improving the process of course presentation.* My phrase "course presentation" is meant to exclude curricular aspects of teaching, and might best be understood as that part of the teaching-and-learning process that takes the curriculum largely as given. In drawing this distinction, I do not mean to underestimate the importance of course presentation. As David Moore pointed out at the conference, "The typical college class can be greatly improved by caring about the outcome and by attention to detail, without a major overhaul of the process. (TQM people say this is true of most processes.)" More and more statisticians are using elements of TQM in this way, especially the Mosteller one-minute drill (Zahn 1993a), and many report that their courses are substantially more effective as a result. Some of the more mechanical but nonetheless critical improvements include such things as switching to a microphone to be heard better, and using thicker chalk in order to write with larger letters. Other improvements include clearer lectures as a result of daily feedback, and a shorter but more effective list of assigned readings. If these improvements strike you as too obvious to serve as an argument for TQM, I suggest an empirical test: try it yourself. (Sometimes the only way to find out if you live in a glass house is to hand out stones and listen for the tinkle of breaking panes.) The teachers who have written about these changes were already effective and popular, and they were often truly surprised at what they discovered. For helpful examples of things to try and first-hand accounts by those who have tried them, see Hau (1991), Bateman and Roberts (1992), and Zahn (1992).

74 (d) *TQM as an approach to thinking about curriculum.* If "course presentation" in (c) refers to how we teach, then "curriculum" here refers to what we teach. At the surface, we can think of the "what" as subject matter content, perhaps even a list of concepts and skills. In (c) above, we were able to fit TQM to the learning process by thinking of course presentation as the product, teacher as the producer, and student as the customer to be satisfied. One straightforward and increasingly common alternative identifies the student as the product, and the student's future employer (or teacher in the next course) as the customer. This view leads to the eminently sensible though not terribly deep notion that what we teach in our courses should take into account the needs of those who will be teaching or employing our students after they have left our course.

75 The two uses of TQM in (c) and (d) above are similar in that both rely on a model of customers and products. In this respect both uses are close to the origins of TQM in statistical process control, and both uses are open to the objection that manufacturing provides at best an ill-fitting and superficial model for teaching and learning. I think there is value to the objection, despite the persuasive evidence, from those who apply the manufacturing model, that it leads to substantial improvements in their courses. To me the objection is valuable because it pushes us to ask whether TQM can offer a better fit than we get using the manufacturing model. After all, as Snee (1990) and others remind us, TQM involves more than just process control.

76 To many of its practitioners, TQM is not just about manufacturing, but is more profoundly (1) an interpersonal theory, (2) of shared responsibility (3) for the quality of a process, (4) taking place in a traditionally hierarchical organization. Though some may brand me heretic, I consider these four elements essential to TQM, in a way that customers and products are not. As long as we can measure the quality of a process, we can apply TQM. (Customers and products just happen to combine, like functionals acting on functions, to provide useful measures of quality in the setting that originally gave rise to the theory.) On this view of TQM, the fit to education is a good one: Teaching-and-learning is an interpersonal activity, a process for which thinking about quality and trying to measure it are profoundly important, and one for which teacher and student share responsibility, in a setting which is traditionally, and to some extent unavoidably, hierarchical. (For a first-hand account of how the theory can change a teacher's relationship with her students,

see [Kinard (1992)](#). For a somewhat different view, see [Stinnett (1992).](#))

77 Applying this model of TQM leads to two principles that govern the student-teacher relationship:

- 78 Student and teacher share responsibility for the quality of a process -- the process of the student's learning (only indirectly and secondarily the quality of the teacher's teaching).
- 79 The core motivation, for both student and teacher, should be the satisfaction that derives from improving the quality of the student's learning.

80 Despite the shared responsibility and implied cooperation, a student-teacher relationship governed by these principles is still hierarchical. After all, the teacher is more experienced (and is the one getting paid). Moreover the focus of the cooperative efforts is -- asymmetrically -- on the student's learning, not the teacher's. Nevertheless, this model calls for a major change in the teacher's role. Under the old model, the teacher delivers knowledge, directs the details of student efforts, and scores received-knowledge-as-product. Under the alternative, the teacher must put two goals ahead of all others: (1) to lead students to claim responsibility for their own learning ([Rich 1979](#)), and (2) to lead students to derive their principal motivation from improving the quality of their learning process.

81 How can we as teachers do these two things? Surely not by becoming cheerleaders. (Neither quality nor teaching is about pom-poms and slogans.) Can we design our curriculum to serve these goals? I think we can, if we base that curriculum on two principles: active learning and authentic assessment.

82 *Learning must be active if it is to build a student's sense of responsibility for the process; lecture-based courses undermine the student's sense of responsibility for learning*. The teacher is neither producing a course for the student nor producing a student for an employer. Both these models make our students passive consumers rather than active constructors of their education. As teachers we should not think of teaching as like writing on a blank slate; instead, we should think of learning as like building with a set of Legos. If we tell our students where to put each block, is it any wonder they do not experience themselves as responsible for what goes on?

83 *Responsibility for the process of learning cannot be shared unless the way to assess its quality is authentic, public, clearly understood, and accepted by the students*. If quality is to improve, assessment must measure what matters most. If students are to share responsibility for quality, the standards for judging quality must be shared. In short, assessment is pivotal.

3.4 Educational Assessment

```
84     ||   Current thinking about educational assessment seeks   ||
       ||   to integrate the often-distinct processes of          ||
       ||   teaching and testing.  Both the timing and the        ||
       ||   tasks are in flux.  End-point testing is giving way   ||
       ||   to continuous assessment over the course of the       ||
       ||   semester.  Test questions designed just for grading   ||
       ||   are giving way to tasks like projects, oral           ||
       ||   reports, and open-ended writing assignments.          ||
```

85 It is an axiom of TQM that to improve quality, we must first have a way to measure quality; to the extent that our measure is superficial, improvements will be at the surface, rather than deep down. The same axiom is basic to educational assessment. In the words of psychologist Lauren Resnick, "We get what we assess, and if we don't assess it, we won't get it." (Quoted in [Wiggins 1992, p. 152.](#)) This points to another responsibility of the teacher: to choose those tasks that will be used to measure quality (and therefore structure the learning

process), and to provide clear, public standards of what constitutes high quality.

86 We need what Grant Wiggins calls "authentic assessment": ... composed of tasks we value ... the tasks and virtues at the heart of the subject -- its standards. (Wiggins 1992, p. 152.) We want to gather systematic information about how well a student is learning to do what it is that statisticians do, and to think the way statisticians think. Back before statistics became "a subject", the model for learning was apprenticeship -- watching, then assisting, then increasingly taking major responsibility for doing whatever the "master" statistician was called upon to do. Assessment was not based on a separate set of tasks designed for that purpose, but was built into all feedback to the apprentice from the master during the ordinary course of learning the craft by practicing it. Of course no statistician was ever assigned responsibility for 200 apprentices in a single semester, so the model has its limits if we take it too literally, but I consider it a useful guide -- perhaps what TQM people might call a "stretch objective" -- for thinking about assessment.

87 To illustrate some of the current ideas about educational assessment, consider a multiple-choice test and a data-gathering-and-analysis project as alternative vehicles for assessment. The test is taken at the end of the semester (or at the end of some other "unit" of learning); it is designed and intended purely to measure what has already been learned, and not as a learning experience in itself. Testing (like end-point inspection) evaluates a process solely in terms of the rate of defects in the final product. In contrast, the project stretches out over a large chunk of the semester, and is appropriately regarded as an important learning experience in itself (quite apart from its role in assigning a grade). Many teachers who assign projects collect written work, provide feedback (and perhaps assign numerical grades) at several stages during the project (Halvorsen and Moore 1992). If we compare the project and the test against the apprenticeship model, clearly the project comes closer. (For a summary of current thinking about assessment, see "Notes on Assessment and Teaching Statistics," which Joan Garfield prepared for the conference. The appendix gives instructions for obtaining this file.)

88 As a metaphor, we can usefully regard assessment as a kind of composite function, first from statistician space (where we live) to task space (where our students live), then from task space to evaluation space (where the grades are). With a test, there is a clear and direct route from what it is the student does (i.e., provides a set of answers) to the grade the instructor assigns. In contrast, with a project, the scoring is much more subjective. Scoring a test has very high inter-rater reliability; grading a project not nearly so high. BUT -- the scoring is only half the path. The other half of the connection runs back from the task (project or test) to what it is that we want out students to learn. For the project, the connection is so direct that no one even raises the issue. Objections to projects are always logistical ("Students don't know enough to plan a decent project until late in the semester." "Dealing with the interpersonal aspects of teams is such a hassle." "Grading is too subjective, and besides, I don't have the time to read all those papers."); objections never take the form "When you come right down to it, answering test questions is closer to what statisticians actually do in practice." For the test, the connection to statistical practice is an unexamined article of faith, or, absent the faith, simply unexamined.

89 Authentic assessment should serve as our goal, but not just for a handful of tasks inserted at intervals throughout the semester. Once we accept that assessment must be authentic, the most radical implication of TQM is that the entire course should be built of assessment tasks. Certainly the whole course should be built of tasks we value, which means that any part of the course could be used for authentic assessment. For a course built this way, the only things that distinguish those tasks singled out for assessment are that we ask students explicitly to attend to the quality of their efforts, we provide them with clear standards for judging high quality work, and we give them feedback. If we want students to feel continuously responsible for the quality of their learning, we should be doing those three things all the time.

90 What kinds of activities lend themselves to "doing these three things all the time"? In principle, anything

that statisticians do in a professional capacity is a candidate for such an activity. In practice, one might start from the following two symmetric pairs of suggestions. (The structure within pairs follows one of the assessment principles of Grant Williams, that we should assess students not just on what they do, but also on how well they can assess what they and others do.)

91 1. Design for data production: (a) Plan an experiment or survey or other data-gathering project. (b) Evaluate someone else's plan.

92 2. Analysis: (a) Analyze a data set and produce a written or oral report of the analysis. (b) Evaluate a report of someone else's analysis.

93 Taken as is, the suggested activities might be used for a small number of major assignments. In addition, one needs smaller, shorter tasks, and for that purpose the pair of suggestions can be used to generate variations. As a time-compressed metonymic variant of (2), an assessment activity could substitute the outline of a proposed analysis for the analysis itself. Similarly, in the spirit of synecdoche, an activity might be based on just a representative part of (1) or (2), e.g., "Propose a strategy (or evaluate a given strategy) for protecting against the potential confounding effects of teacher expectation in a comparative study of the value of projects in teaching statistics", or "Here are some summary statistics and plots from the early stages of an analysis of possible age-bias in employee layoffs. Write a page telling what you would do next in trying to understand the data."

94 Taken together, assessment theory and TQM call on us to invert the usual order and priorities when we plan our courses. Instead of asking first "What will I teach them?" then "How will I deliver it?" and finally, almost as an afterthought, "How will I measure how much they got?" we should turn things around. Start with "What tasks, done well, would convince me that a student has learned the most important elements of how to think and work like a statistician?" Then build a course from those tasks.

4. Conclusion

95 The mystery writer Dorothy Sayers once wrote, "Facts are like cows. If you look them in the face hard enough they generally run away" (Sayers 1927, p. 68). In our particular field, the usual herd of cud-chewing introductory courses has been a standing fact of life for much too long. However, the various curricular projects summarized in Section 2 represent a hard and encouraging look forward. If more and more of us stare harder and harder, I think it is reasonable to hope that most of the old curricular ruminants will eventually turn tail and run. In Section 3.1, I have argued by example that, of the usual supposed facts about the beginning course, neither its content, nor its organization, nor its mode of delivery is essential for effective learning about statistics. We are actually much freer than we often think to rebuild our curriculum from the ground up. In Section 3.2, I argued that if we stare analytically into the face of what appear to be student resistances to innovation, these also go away: Although student complaints reflect legitimate concerns, they are best understood as symptoms either of correctable logistical problems or correctable lack of clarity about goals and standards. Solving these problems requires better communication of two kinds: regular feedback from students to the instructor about their concerns, and more effective statements from the instructor about what is important and why, and about what is expected and how to judge its quality. These are essential elements of TQM. In Section 3.3 I have argued that taking TQM seriously means teachers must lead students to claim responsibility for the quality of their learning, and to derive their principal satisfaction from improving that quality. Doing this requires that we base our curriculum on active learning and authentic assessment. Finally, in Section 3.4 I borrow from Grant Wiggins the idea that authentic assessment is composed of tasks we value.

It is this idea, ultimately, that tells us how to build our curriculum. We first ask "What are the things we do that are most basic to being a statistician?" We then build a course from those activities, and use the activities themselves as the basis for assessing quality.

Appendix
NSF-Supported Projects in Statistics Education, 1990-1992

1. Improving Education in Statistics for Engineers

Dennis C. Gilliland, Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824

Phone: 517-353-7820 or 517-355-9589 dept office, FAX: 517-336-1405

E-mail:
20974dcg@msu.edu

Supplementary information

A computer laboratory has been created, consisting of an instructor's workstation, slave monitors to display the instructor's screen, and sixteen workstations each accommodating two students. The goal is to allow faculty to demonstrate statistical ideas on computers, with interactive lab exercises as follow-up. Concepts to be demonstrated include variability, effect of sample size, central limit theorem, exact confidence intervals, and the bootstrap. A module on control charting will ask students to chart a variety of simulated manufacturing processes, some of them in control, others with drifting mean or variance, and others chaotic. Another module, on response surface analysis, will allow students to design experiments to investigate the effect of control variables such as temperature, concentration of reactants, and duration of the reaction on the yield. The statistics course for quality and productivity will provide the key test for the proposed teaching laboratory.

2. Workshop on Statistical Education

Robert V. Hogg, Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA 52242

Phone: 319-335-0828 x 0607 dept office, FAX: 319-335-0627

E-mail:
bhogg@stat.uiowa.edu

Supplementary information

Thirty-nine statisticians gathered for a workshop on statistical education in Iowa City, IA, on June 18-20, 1990. The workshop was sponsored by the University of Iowa and the American Statistical Association

(ASA) with financial assistance provided by the NSF, the Alcoa Foundation, the Ott Foundation, the Statistics Division of the American Society for Quality Control, and the Quality and Productivity Section of the ASA. Topics included (1) The nature of the problem, (2) Poor characteristics of science and mathematics education generally, (3) Initiatives for universities, (4) Problems with the introductory courses, and (5) Initiatives for the profession. Many of the main suggestions resulting from this workshop were listed on pp. 19-20 in the November 1990 issue of *Amstat News*. (See also pp. 34-43 of *Heeding the Call for Change*, ed. Lynn Steen, MAA Notes No. 22, 1992.)

3. Development of a Multi-Media-Based Statistics Classroom

Fidelis Ikem, Department of Information Systems and Decision Sciences, Virginia State University, Petersburg, VA 23806

Phone: 804-524-5110, FAX: 804-524-5110

The goal is to utilize advances in technology to improve the teaching of statistics generally, and in particular to improve the learning environment so as to attract more minority students to quantitative subjects. The project will involve two introductory statistics courses and an econometrics course, which together affect 30% of the student body. IBM PCs will provide access to the university's mainframe. The courses will use SPSS PC and the time series package Micro TSP to analyze data from the microeconomic data base CitiBase.

4. Undergraduate Statistics Laboratory

Rhonda C. Magel, North Dakota State University, 300 Minard Hall, Fargo, ND 58105-5075

Phone: 701-252-5363, 701-237-7177 7492 dept office, FAX: 701-237-8562

Supplementary information

The focus of the project is an undergraduate statistics laboratory that will include a network of one server and 20 color Macintosh microcomputers equipped with MINITAB. The goal is to incorporate experience with real data into the teaching of statistics. To that end, data sets from a variety of client disciplines will be collected: agriculture, business, computer science, home economics, pharmacy, psychology, and transportation. In addition, the project will rely on class-generated survey data on issues of current interest to students. Six courses with a combined annual enrollment of 1420 will use the lab: two introductory courses (one non-calculus, one calculus-based), two regression courses, an introduction to experimental design, and a nonparametrics course.

5. Development of a Modern Computing and Graphics-Based Method for Teaching Important Concepts in Undergraduate Statistics Courses

William Q. Meeker, Department of Statistics, Iowa State University, 326 Snedecor Hall, Ames, IA 50011

Phone: 515-294-5336, FAX: 515-294-4040

E-mail:

wqmeeker@iastate.edu

[Supplementary information](#)

The goal is to develop 30 easy-to-use instructional software modules to illustrate fundamental statistical concepts. The modules will use a combination of state-of-the-art workstations, statistical programming languages (S-plus and Lisp-Stat), high resolution color graphics, and a highly interactive user interface. One module, for example, will focus on the interpretation of probability plots, more specifically, on distinguishing signal from noise and on the effect of sample size. Undergraduate courses to be affected include: applied time series, statistical research methods (regression analysis and analysis of variance), multivariate analysis, quality control, and experimental design, as well as some service courses.

6. Improving the Undergraduate Statistical Education of Engineers

Peter Nelson, Department of Mathematical Sciences, Clemson University, Clemson, SC 29634-1907

Phone: 803-656-2882, FAX: 803-656-5230

E-mail:
npeter@clemson.bitnet

[Supplementary information](#)

The goal is to introduce concepts and practice of statistics into existing engineering courses through their labs, rather than by teaching a separate statistics course for engineers. For example, an existing engineering lab exercise involves using a heat exchanger to gather data from which to estimate various heat transfer coefficients. The exercise will be revised and enriched to address a number of statistical issues: using replicate observations to check an assumption of linearity, attaching confidence limits to the estimates, and using ideas from experimental design to choose an efficient set of combinations of conditions for the observations. This project requires close collaboration between the statisticians and the engineers.

7. Technology-Based Learning: Exploring Statistical Concepts and Methods

William I. Notz (et al.), Department of Statistics, Ohio State University, Cockins Hall, 1958 Neil Avenue, Columbus, OH 43210

Phone: 614-292-3154 x2866 dept., FAX: 614-292-2096

E-mail:
win@osustat.mps.ohio-state.edu

[Supplementary information](#)

The goal is to develop an extensive electronic encyclopedia of examples, exercises, and data sets which integrates statistics software (Data Desk and Minitab), videotape materials (clips from the Annenburg/CPB tapes *Against All Odds*), and affordable multimedia technology (e.g., Apple's Quicktime CD technology). The encyclopedia will be easy to use and will be keyed both to topics and to statistical concepts. An important

feature will be the built-in ability to launch statistics software from within the encyclopedia in order to answer questions or analyze data sets being viewed. The materials will be used at Ohio State, where 30,000 undergraduates are required to take statistics as part of the general education curriculum, and at Cornell University.

8. An Activity-Based Introductory Statistics Course for All Undergraduates

Richard L. Scheaffer, Department of Statistics, University of Florida, Gainesville, FL 32611

Phone: 904-392-1941, FAX: 904-392-5175

E-mail:
scheaffe@stat.ufl.edu

Supplementary information

The goal is to produce a cohesive collection of laboratory units for a one-semester course in modern statistics for the introductory-level education of all undergraduates. The units will (1) focus on modern graphical and exploratory approaches to data analysis, (2) build understanding of probability through simulation experiences, (3) introduce the basic ideas of statistical inference using an exploratory approach, (4) consider elementary model fitting, (5) emphasize hands-on activities with real data, including student projects, in a "laboratory" setting rather than a traditional lecture setting, and (6) use calculators and computers throughout to enhance understanding of concepts. The project will develop: (1) a detailed outline of topics for a one-semester activity-based course, (2) a set of laboratory units covering the material in the outline, including complete examples from a variety of disciplines, (3) a comprehensive list of references, (4) a teacher's manual on alternative methods of assessment, including advice on how to grade group activities, lab participation, and student projects, and (5) a teacher's manual on implementation strategies, covering large and small classes and a variety of computing facilities.

9. CHANCE: Case Studies of Current Chance Issues (An Introductory Mathematics Course)

J. Laurie Snell, Department of Mathematics, Dartmouth College, Hanover, NH 03755-1890

Phone: 603-646-2951 work, 603-646-3347 home

E-mail:
j.laurie.snell@dartmouth.edu

Supplementary information

This project is a cooperative effort among six colleges -- Dartmouth, Grinnell, King, Middlebury, Princeton, and Spelman -- to develop and teach a new introductory mathematics course called CHANCE. The course treats issues that are reported in the news: statistical problems related to AIDS, the effects of lowering serum cholesterol on heart attacks, the use of DNA fingerprinting in the courts, maintaining the quality of manufactured goods in the face of variation, informed patient decision making, reliability of political polls,

and so forth. The course does not aim to develop a canon of statistical methods, but rather to help students learn how to think about statistics and probability, and how to seek out for themselves the tools appropriate for a particular problem. Students study the newspaper accounts of these topics, articles in general science journals such as *Chance*, *Science*, *Nature*, and *Scientific American*, and finally the original research papers. The project is developing summary modules, student projects, classroom experiences, data sets, and bibliographies, and maintaining an electronic bulletin board, *CHANCE News*.

10. Elementary Statistics Laboratory Course Development

John D. Spurrier, Department of Statistics, University of South Carolina, Columbia, SC 29208

Phone: 803-777-5072, FAX: 803-777-4048

E-mail:
N410009@UNIVSCVM.bitnet

Supplementary information

The goal is to establish a prototype elementary statistics lab and to create a one semester hour lab course to be taken with or after the standard lower-division introductory course. The lab will guide the student through simple, but meaningful experiments that illustrate important points of applied statistics. In each session the student will discuss and perform an experiment, collect and analyze data, and write a report. The lab will differ from traditional science labs in that the emphasis will be on statistical concepts. Labs will be conducted in a room equipped with 12 Macintosh Classics running MINITAB. Student and teacher's manuals will be prepared. The lab topics are (1) Introduction to Macintosh and MINITAB, (2) Pulse rate [descriptive statistics], (3) Parking lot sampling, (4) Real and perceived distances [scatterplots and variation], (5) Traffic counts [time series], (6) Taste test [binomial, paired comparison], (7) Carpet tacks [variation, quality improvement], (8) Sampling distribution of mean and median [simulation, estimation, Central Limit Theorem], (9) Absorbency of paper towels (sampling, confidence interval for mean], (10) Breaking strength of string and fish line [confidence intervals, hypothesis testing], (11) Airplane flight distance [factorial designs], (12) Normal walking versus exaggerated arm movement [dependent sample comparison of means], (13) Plant experiment [factorial design, model building], and (14) Prediction of hickory nut weight [regression, correlation, plotting].

11. Statistical Computing Laboratory

Dolores M. Tichenor, Department of Mathematics, Tri-state University, Angola, IN 46703

Phone: 219-665-4242, FAX: 219-665-4292

Supplementary information

The goal is to equip a microcomputer laboratory so that the upper-division two-quarter probability and statistics sequence can incorporate individual and group projects involving real data. Each quarter the students will conduct four individual projects and one group "capstone" project. Projects will be based on data from local sources: Tri-State University's Biology Department and Analytical Testing Laboratory both gather environmental data on area lakes; similar data come from the County Health Department, Indiana Departments of Environmental Management and of Natural Resources. Students will also analyze data from

the University Fitness Center, and will serve as consultants to students enrolled in biology courses.

12. Materials for a Computer-Based Introductory Statistics Curriculum Using Actual Data

Bruce E. Trumbo, Department of Statistics, California State University - Hayward, Hayward, CA 94542-3087

Phone: 415-881-3435, FAX: 415-727-2035

Supplementary information

The goal is to produce (1) a collection of carefully documented data sets, from recent and important scientific research, in a form ready for computer analysis, (2) an instructor's manual, with general suggestions for using computers to teach statistics, specific suggestions for using the data sets to teach particular statistical concepts, and detailed background on the data sets, and (3) student workbooks, which lead the student through exploration and analysis of selected data sets using MINITAB. The data sets and materials will not constitute a course in themselves, but rather are intended to be used with a variety of books, courses, formats, fields of application, and computer environments. There will be 30 data sets in all: 10 large (ca. 1000 cases, 12 variables), 10 moderate sized, and 10 small.

Two additional files distributed by Joan Garfield at the NSF Conference are also available: "Notes on Assessment and Teaching Statistics" and "Sample Course Handouts for Group Work".

References

Albers , D. J., Loftsgaarden, D. O., Rung, D. C., and Watkins, A. E. (1992), *Statistical Abstract of Undergraduate Programs in the Mathematical Sciences and Computer Science in the United States*, MAA Notes No. 23, Washington: Mathematical Association of America.

Bateman, G. R., and Roberts, H. V. (1992), "TQM for Professors and Students," Graduate School of Business, University of Chicago.

Bentley, D. L. (1993), "Investigational Statistics: A Data Driven Course," in *1992 Proceedings of the Section on Statistical Education, American Statistical Association*, 119-122.

Bryce, G. R. (1993), "Data Driven Experiences in an Introductory Statistics Course for Engineers Using Student Collected Data," in *1992 Proceedings of the Section on Statistical Education, American Statistical Association*, 155-160.

Cobb, G. W. (1990), "The Quantitative Reasoning Course at Mount Holyoke College", in *The New Liberal Arts Program: A 1990 Report*, ed. Samuel Goldberg, New York: Alfred P. Sloan Program, 14-31.

----- (1992), "Teaching Statistics" in *Heeding the Call for Change*, ed. Lynn Steen, MAA Notes No. 22, Washington: Mathematical Association of America.

Cross, K. P. (1986), "A Proposal to Improve Teaching, or What `Taking Teaching Seriously' Should Mean,"

AAHE Bulletin, 39, 9-15.

Cryer, J. D., and Miller, R. B. (1991), *Statistics for Business: Data Analysis and Modeling*, Boston: PWS-Kent Publishing Company.

Dempster , A. P. (1977), "The Robustness of Applied Inferences," in *Statistical Decision Theory and Related Topics II*, eds. S. S. Gupta and D. S. Moore, New York: Academic Press, 121-138.

Ewell, P. T. (1991), "Assessment and TQM: In Search of Convergence", *New Directions for Institutional Research*, No. 71, San Francisco: Jossey-Bass, Inc., 39-52.

Halvorsen , K. T., and Moore, T. L. (1992), "Motivating, Monitoring, and Evaluating Student Projects," in *1991 Proceedings of the Section on Statistical Education, American Statistical Association*, 20-25.

Hau, I. (1991), "Teaching Quality Improvement by Quality Improvement in Teaching", Technical Report No. 59, Madison, WI: Center for Quality and Productivity Improvement.

Hogg , R. V. (1990), "Statisticians Gather to Discuss Statistical Education," *Amstat News*, No. 169, 19-20.

----- (1991), "Statistical Education: Improvements Are Badly Needed," *The American Statistician*, 45, 342-343.

----- (1992), "Report of a Conference on Statistical Education," in *Heeding the Call for Change*, ed. Lynn Steen, MAA Notes No. 22, Washington: Mathematical Association of America.

Kinard, K. A. (1992), "Being on One Team With Your Students: Radical Idea? ... Radical Results!," in *1991 Proceedings of the Section on Statistical Education, American Statistical Association*, 141-146.

Landwehr , J. M. (1993), "Project STATS: Statistical Thinking and Teaching Statistics," *Amstat News*, No. 196, 25.

Larsen, R. J., and Stroup, D. F. (1976), *Statistics in the Real World: A Book of Examples*, New York: Macmillan Publishing Co.

Lock, R. (1990), "Forecasting/Time Series Analysis: An Introduction to Applied Statistics for Mathematics Students," SLAW Technical Report No. 90-001, Department of Mathematics, Pomona College, Claremont, CA.

McKenzie, J. D., Jr. (1993), "The Use of Projects in Applied Statistics Courses," in *1992 Proceedings of the Section on Statistical Education, American Statistical Association*, 142-146.

Meeker , W. Q. (1993), response to Moore, D. S. "Statistics Education Fin de Siecle," in preparation.

Moore, D. S. (1988), "Should Mathematicians Teach Statistics?" (with discussion), *The College Mathematics Journal*, 19, 3-25.

----- (1992), "Teaching Statistics as a Respectable Subject," in *Statistics for the Twenty-First Century*, eds. Florence Gordon and Sheldon Gordon, MAA Notes No. 26, Washington: Mathematical Association of America, 14-25.

----- (1993), personal communication.

Moore, T. L., and Roberts, R. A. (1989), "Statistics at Liberal Arts Colleges," *The American Statistician*, 43, 80-85.

Mosteller , F. (1988), "Broadening the Scope of Statistics and Statistical Education," *The American Statistician*, 42, 93-99.

Rich , A. (1979), "Claiming an Education," in *Lies, Secrets, and Silence*, New York: W. W. Norton & Co., 231-235.

Roberts , H. V. (1992), "Student-Conducted Projects in Introductory Statistics Courses," in *Statistics for the Twenty-First Century* , eds. Florence Gordon and Sheldon Gordon, MAA Notes No. 26, Washington: Mathematical Association of America, 109-121.

Rossman, A. J. (1993), "Introductory Statistics: The `Workshop' Approach," in *1992 Proceedings of the Section on Statistical Education, American Statistical Association*, 352-357.

Sayers, D. L. (1927), *Clouds of Witness*, New York: Harper & Row.

Snee , R. D. (1990), "Statistical Thinking and Its Contribution to Total Quality," *The American Statistician*, 44, 116-121.

----- (1993), "What's Missing in Statistical Education?" *The American Statistician*, 47, 149-154.

Snell, J. L. (1992), "CHANCE: Case Studies of Current Chance Issues," in *Statistics for the Twenty-First Century*, eds. Florence Gordon and Sheldon Gordon, MAA Notes No. 26, Washington: Mathematical Association of America, 281-297.

----- and Finn, J. (1992), "A Course Called `Chance'", *CHANCE: New Directions for Statistics and Computing*, 5, 12-17.

Stinnett, S. S. (1992), "Quality Improvement Procedures in Statistical Consulting Education," in *1991 Proceedings of the Section on Statistical Education, American Statistical Association*, 147-152.

Sylwester , D. L. and Mee, R. W. (1993), "Student Projects: An Important Element in the Beginning Statistics Course," in *1992 Proceedings of the Section on Statistical Education, American Statistical Association*, 137-141.

Walton, M. (1986), *The Deming Management Method*, New York: The Putnam Publishing Group.

Wardrop, R. L. (1992), "A Radically Different Approach to Introductory Statistics," Technical Report No. 889, Department of Statistics, University of Wisconsin.

----- (1993), *Statistics: Learning in the Presence of Variation*, William C. Brown Publishing (in final preparation).

Wiggins , G. (1992), "Toward Assessment Worthy of the Liberal Arts," in *Heeding the Call for Change*, ed. Lynn Steen, MAA Notes No. 22, Washington: Mathematical Association of America.

Willett , J. B. and Singer, J. (1992), "Providing a Statistical `Model': Teaching Applied Statistics Using Real-World Data," in *Statistics for the Twenty-First Century*, eds. Florence Gordon and Sheldon Gordon, MAA Notes No. 26, Washington: Mathematical Association of America, 83-98.

Witmer, J. (1990), "Data Analysis: An Adjunct to Mathematical Statistics at Oberlin College," SLAW Technical Report No. 90-003, Department of Mathematics, Pomona College, Claremont, CA.

----- (1992), "Using a Class Project to Teach Statistics," in *1991 Proceedings of the Section on Statistical Education, American Statistical Association*, 26-28.

Wolf, F. L. (1990), "Multivariate Descriptive Statistics: An Alternative Introduction to Statistics", SLAW Technical Report No. 90-004, Department of Mathematics, Pomona College, Claremont, CA.

Zahn, D. A. (1992), "Getting Started on Quality Improvement in Statistics Education," in *1991 Proceedings of the Section on Statistical Education, American Statistical Association*, 135-140.

----- (1993a), "Notes on the Use of Minute Papers in Teaching Statistics Courses," in *1992 Proceedings of the Section on Statistical Education, American Statistical Association*, 62-71.

----- (1993b), "Student Projects in a Large Lecture Introductory Business Statistics Course," in *1992 Proceedings of the Section on Statistical Education, American Statistical Association*, 147-154.

George W. Cobb
Department of Mathematics, Statistics, and Computation
Mount Holyoke College
South Hadley, MA 01075
gcobb@mhc.mtholyoke.edu