



## Teaching Statistics Theory Through Applications

D. Nolan; T. P. Speed

*The American Statistician*, Vol. 53, No. 4. (Nov., 1999), pp. 370-375.

Stable URL:

<http://links.jstor.org/sici?sici=0003-1305%28199911%2953%3A4%3C370%3ATSTTA%3E2.0.CO%3B2-C>

*The American Statistician* is currently published by American Statistical Association.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/astata.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

# Teacher's Corner

## Teaching Statistics Theory Through Applications

D. NOLAN and T. P. SPEED

This article presents a model for developing case studies, or labs, for use in undergraduate mathematical statistics courses. The model proposed here is to design labs that are more in-depth than most examples in statistical texts by providing rich background material, investigations and analyses in the context of a scientific problem, and detailed theoretical development within the lab. An important goal of this approach is to encourage and develop statistical thinking. It is also advocated that the labs be made the centerpiece of the theoretical course. As a result, the curriculum, lectures, and assignments are significantly restructured. For example, the course work includes written assignments based on open-ended data analyses, and the lectures include group work and discussions of the case-studies.

**KEY WORDS:** Instruction; Quantitative literacy; Statistical thinking.

### 1. INTRODUCTION

Over the past 10 years we have developed a course that teaches undergraduate upper-level mathematical statistics through the use of in-depth case studies, which we call labs. It is through the labs that the theory of statistics is introduced, which leads to an integration of statistical theory and practice in a way not commonly encountered in an undergraduate course. The labs raise scientific questions that are interesting in their own right, and they contain datasets for use in addressing these questions. The context of the scientific question is the starting point for developing statistical theory.

This article presents a model that we have found to be successful for designing these labs. The model calls for the lab to be a substantial exercise with nontrivial solutions that leave room for different analyses, and for it to be a central part of the course. The lab should offer motivation and a framework for studying theoretical statistics, and it should give students experience with how statistics can be used to answer scientific questions. An important goal of this

approach is to encourage and develop statistical thinking while imparting knowledge in mathematical statistics.

It often can be difficult for students to bring the mathematical statistics learned in the classroom to an independent project in school or on the job. It can also be difficult for students to make the transition from reading and understanding a critical review of a statistical analysis to successfully working on their own problem. In our attempt to address these issues, we propose here to mix theory with substantial case studies. In addition to increasing student understanding of statistical theory, through the labs students develop their quantitative reasoning and problem-solving skills in a broad, multidisciplinary setting. They become practiced in communicating their ideas orally and in writing, and they become versed in the use of statistical software.

### 2. BACKGROUND

The labs were designed to accompany a typical theoretical statistics course for juniors and seniors. We have used these labs for two such courses—one for mathematics and statistics majors and one for engineering and computer science majors. Both courses require students to have studied calculus for two years and probability for one semester. There is no statistics prerequisite for the course.

The material covered in the course is outlined in the following; it is divided into five broad categories. In a typical semester, roughly one week is spent on the first topic, summary statistics, and about three weeks are spent on each of the remaining four topics: sampling, estimation and testing, regression and simple linear least squares, and analysis of variance and multiple linear least squares.

#### 1. Summarizing data

- a. *Univariate summaries*: histogram, box plots, bar charts, and quantile plots; measures of location and spread; normal approximation, skewness, kurtosis.
- b. *Bivariate summaries*: cross-tabulations, quantile-quantile plots, and scatterplots.

#### 2. Sampling

- a. *Simple random*: sample mean, percent, and total; expectation, variance, and estimator for variance; normal approximation and confidence intervals.
- b. *Stratified*: expectation, variance; proportional and optimal allocation; bootstrap.

#### 3. Parameter estimation, goodness-of-fit, and hypothesis testing.

D. Nolan is Professor, and T. P. Speed is Professor, Department of Statistics, University of California, 367 Evans Hall, Berkeley CA 94720-3860 (Email: nolan@stat.berkeley.edu). This work has been partially supported by NSF grant 9720490. The authors thank the reviewers for their comments which led to an improved and more informative article.

### 3. A LAB MODEL

- a. *Parameter estimation*: method of moments; maximum likelihood; mean square error, asymptotic bias, and variance; information and Cramer–Rao inequality.
  - b. *Goodness-of-fit and testing*: chi-square tests; two-sample tests; test statistic, null distribution, and Type I and II errors.
4. Regression and the simple linear model.
- a. *Regression*: correlation; least squares; residuals, confidence intervals; bivariate normal.
  - b. *Simple linear least squares*: linear model, RMS error, transformations, polynomial regression, *t*-tests and *F*-tests.
5. Extensions to linear models
- a. *Analysis of variance*: one-way and two-way classification, additive model, *F*-test; least squares with indicators; sum of squares decomposition.
  - b. *Multiple linear regression*: geometry of least squares; diagnostics, model checking.

We started to develop labs for these courses in 1989. At that time, each lab consisted of a list of exercises designed to illustrate a particular statistical concept and to be applied to the data provided. These original labs were unsatisfactory for many reasons. They lacked context; the presentation was too much like an instruction manual; and they were separate from the development of the course material. Students viewed the labs as drudge work; they did not see them as important or relevant to the course. In response to these shortcomings, we began to expand the role of the labs. We kept those labs for which we had interesting scientific problems to accompany them, and we added to them detailed descriptions of the data, background material to put the problem in context, and suggestions for how to proceed in the analysis. In a further evolution, we incorporated the statistical theory into the lab. This was an important step in the lab development, because it made the labs the centerpiece of the theoretical course, not a secondary accompaniment.

As a result of this fundamental restructure of the labs, several changes took place in the course. The format of lectures changed. More time was spent on determining how to answer general scientific questions with statistical analyses and on deriving a statistical method from its application to a specific problem. Less time was spent covering many small examples constructed to illustrate a single statistical technique. Because the application serves as the motivation for working out the theoretical material, we still cover all the material traditionally covered in the course. Also, the course work changed; now lab assignments required written reports with accompanying graphs and charts, and exams contained questions on theoretical topics related to the labs. Finally, the material distributed to the students for each lab grew from a 2-page list of instructions to a 10- to 20-page in-depth case study. More details are provided in Section 4 on how the labs are integrated into the course.

Over time, we have settled on a model for the organization of a lab. Each lab is divided into five main parts: an introduction, data description, background material, investigations, and theory. Sometimes we include an extension section for a more advanced analysis and related theoretical material.

We describe the content and important features of each section in the following. The descriptions are accompanied by examples taken from a sample lab, one that introduces parameter estimation and hypothesis testing by asking students to search for unusual clusters of patterns in the DNA of a herpes virus. Despite the technical nature of the topic, this lab is one of the most popular among the students. The students derive hypothesis tests that study features of the DNA against a model for random scatter based on the homogeneous Poisson process.

#### 3.1 Introduction

Here a clear scientific question is stated, and the motivation for answering it is given. The question is presented in the context of the scientific problem, and not as a request to perform a particular statistical method. Typically the introduction ends with the question to be addressed by the data described in the following section. We do not address questions suggested by the data, instead we consider questions raised by researchers then use statistics to seek answers.

The introduction often begins with a current newspaper article on a related subject to convey the relevance of the question at hand. The DNA lab includes an excerpt from “Feat is Milestone in Study of Evolution” by Nicholas Wade, *New York Times* (08/01/95). The article describes the sequencing of a bacterium’s DNA. It includes a quote from Watson, one of the codiscoverer’s of the structure of DNA, explaining the importance of the event.

Life is a mystery, ineffable, unfathomable, the last thing on earth that might seem susceptible to exact description. Yet now, for the first time, a free-living organism has been precisely defined by the chemical identification of its complete genetic blueprint.

The creature is just a humble bacterium known as *Hemophilus influenzae*, but it nonetheless possesses all the tools and tricks required for independent existence. For the first time, biologists can begin to see the entire parts list, as it were, of what a living cell needs to grow, survive and reproduce itself. . . .

Following the article, students are provided a brief introduction to the problem: how to complete the “parts list” for the human cytomegalovirus (CMV) by locating regions of DNA that may contain information on how the virus reproduces. To further motivate them, they are told that CMV is a herpes virus which is potentially life threatening for people with suppressed or deficient immune systems, and to develop strategies for combating the virus, scientists study the way in which it replicates. To study how the virus reproduces itself scientists need to find the site on the DNA that contains the instructions for starting the process of replication, called the origin of replication. DNA can be thought of as a long coded message made from the four-letter alphabet A, C, G, and T, and for other herpes viruses, the origin of replication is specially marked with a

Table 1. Locations for the First and Last 30 Palindromes in the Cytomegalovirus DNA

177	1321	1433	1477	3248	3255
3286	7263	9023	9084	9333	10884
11754	12863	14263	14719	16013	16425
16752	16812	18009	19176	19325	19415
20030	20832	22027	22739	22910	23241
...	...	204548	205503	206000	207527
207788	207898	208572	209876	210469	215802
216190	216292	216539	217076	220549	221527
221949	222159	222573	222819	223001	223544
224994	225812	226936	227238	227249	227316
228424	228953				

Source: Chee et al. (1990)

pattern of letters in their DNA code. Students are told that they are to identify any unusual patterns in the 229,354-letter-long CMV DNA. The search is restricted to one type of pattern—clusters of complementary palindromes. In the DNA alphabet, A is complementary to T, C is complementary to G, and a complementary palindrome is a sequence of letters that reads in reverse as the complement of the forward sequence—for example, GGGCATGCCC.

### 3.2 Data Description

Documentation for the data collected to address the question is provided. It includes a detailed description of the study protocol, as appropriate. The data for the DNA lab are quite simple to describe; the data are the locations along the CMV DNA of any palindrome longer than 10 letters. See Table 1 for a list of the first 30 locations; the entire dataset is available on the Web site [www.stat.berkeley.edu/users/nolan/statlabs](http://www.stat.berkeley.edu/users/nolan/statlabs). Chee et al. (1990) published the DNA sequence of CMV, and Leung, Blaisdell, Burge, and Karlin (1991) implemented a hash-coding computer program to screen the CMV DNA in search of patterns. Altogether, they found 296 palindromes that were at least 10 letters long. Palindromes shorter than 10 letters were ignored as they can occur too frequently by chance. The longest ones found were 18 letters long.

### 3.3 Background

Material to put the problem in context is provided in this section. Information is gathered from a variety of sources, and it is presented in nontechnical language. The idea is to present a picture of the field of interest that is accessible to a broad college audience.

For example, in the DNA lab, the material is at the level of the *Cartoon Guide to Genetics* (Gonick and Wheelis 1991). The background information provided includes: a summary of the history of the discovery of DNA; a brief introduction to the structure of DNA; the definition of a complementary palindrome in the DNA language; a description of how a virus' DNA replicates; information on the prevalence and symptoms of the herpes *Cytomegalovirus*; and mention of the burgeoning science of genomics, the study of living things through their full DNA sequences.

### 3.4 Investigations

Suggestions for answering the question posed in the introduction are provided next. These suggestions are written in the context of the problem, using very little statistical terminology. The ideas behind the suggestions vary in difficulty, and are grouped to enable the assignment of subsets of investigations. Also included are suggestions on how to write up the results—for example, as an article for a widely read magazine; as a memo to the head of a research group; or as a pamphlet for consumers.

For the DNA lab, the students already have had experience with histograms and quantile plots, and it is expected that they will derive diagnostic tools and statistical tests for comparing the DNA data against a benchmark homogeneous Poisson process. Following is an excerpt:

How do we find clusters of palindromes?

How do we determine whether a cluster is just a chance occurrence, or a potential replication site?

To begin, pursue the point of view that structure in the data is indicated by departures from a uniform scatter of palindromes across the DNA. Of course, a random uniform scatter does not mean that the palindromes will be equally spaced as are milestones on a freeway. There will be some gaps on the DNA where no palindromes occur, and there will be some clumping together of palindromes. To look for structure, examine the locations of the palindromes, the spacings between palindromes, and the counts of palindromes in nonoverlapping regions of the DNA.

*Simulation*—One starting place might be to first see how random scatter looks by using the computer to simulate it. The computer can simulate 296 palindrome sites chosen at random along a DNA sequence of 229,354 letters, using a pseudo-random number generator.

*Locations and spacings*—Use graphical methods to examine the spacings between palindromes that are one apart, two apart, and so on. Compare what you find for the CMV DNA to what you would expect to see in a random scatter. Consider graphical techniques for examining the locations of the palindromes.

*Counts*—Use graphical displays and more formal statistical tests to investigate the counts of palindromes in various regions of the DNA. Split the DNA into nonoverlapping regions of equal length to compare the number of palindromes in an interval to the number that you would expect from uniform random scatter. The counts for shorter regions will be more variable than those for longer regions. Also consider classifying the regions according to their number of counts.

*The biggest cluster*—Does the interval with the greatest number of palindromes indicate a potential origin of replication?

How would you advise a biologist who is about to start experimentally searching for the origin of replication? Write your recommendations in the form of a memo to the head biologist of a research team of which you are a member.

### 3.5 Theory

The theoretical development of the statistical concepts and methodology appear after the problem is introduced, at the end of the lab handout. The material includes information on the general topics described in Section 2 of this article, as well as topics more specific to the individual lab, such as goodness-of-fit tests for the Poisson distribution and parameter estimation for the log-normal distribution.

In the DNA lab, the homogeneous Poisson process on a line is introduced. We discuss how the Poisson process seems a reasonable model for describing the scatter of palindromes on a sequence of DNA. That is, the CMV DNA has 229,354 base pairs and 296 complementary palindromes, and given the large number of base pairs, we can think of the location of a palindrome as a point on a line segment. If the palindromes are scattered randomly across the DNA, then the number of palindromes found in any small region should be independent of the number in another nonoverlapping region, and the chance that one small region has a palindrome in it should be the same for all regions of the same length.

In addressing the main question of the lab, the analysis centers around issues of how to test the fit of the model. To help students in this task, we explore various properties of the Poisson process—for example, the uniform distribution of the locations of hits, the exponential and gamma distributions for gaps between hits, and the Poisson distribution for counts of hits in intervals. Quantile plots and simulations are used as diagnostic tools. Hypothesis testing is introduced through the chi-square goodness-of-fit test for count data. We also derive a test statistic based on the maximum number of hits (palindromes) in a set of nonoverlapping intervals (segments of DNA). To perform the tests, we need to estimate the intensity parameter, and we derive method of moments and maximum likelihood estimators for the parameter. The students also need to address practical issues of how to select interval lengths, what to do if a cluster is split across intervals, and how to reconcile various results.

## 4. IMPLEMENTATION

We have incorporated the labs into two theoretical statistics courses—one for statistics and mathematics majors and one for engineering students. The typical enrollment is 20 students and 40 students, respectively. The classes meet three hours a week with the instructor and one or two hours a week, respectively, with a teaching assistant. (We have also used versions of these labs with simpler investigations and theoretical material in a course for social and life science majors).

We have created more than a dozen labs for these two courses, but we use only about eight in a semester. The labs are chosen by the instructor according to the topic (theoretical or practical) and the background of the students. These eight labs are divided between those labs that are discussed primarily in lecture and those that require students to do ex-

tensive analyses outside of class and to write short papers containing their observations and solutions.

Typically we ask students to write reports for four labs, about one for each of the five main topics in the course (see Section 2). Table 2 provides two example sets of assignments. Option A could be used in a class for students in the social and life sciences, and option B could be used for engineering students. In addition to learning the core material, many labs also include special topics, which are listed in the table. In the course, we cover all of the material that we formerly covered in the theoretical course (see Section 2) plus a few of these additional topics.

### 4.1 Lecture

Although both types of labs feature prominently in the course, they are handled somewhat differently in class. For both types of labs we found it optimal when lectures include: discussion of the background to a particular problem, where students who have taken courses in related fields can bring their own expertise to the discussion; and motivation of the theoretical material through discussion of how to address a problem from a lab. In addition, for labs on which students are to write reports, we hold regular question-and-answer periods where students raise concerns about their work. For the other labs, we have students work in groups as described in the following.

Roughly about one class period in three is spent on these types of activities. The remainder of time is spent in a more traditional presentation of theoretical results (with reference to the current lab).

### 4.2 Group Work

To facilitate group work in class, we supply handouts to the students. The handouts have one of two formats. In one case, they are abbreviated lists of investigations, and students are asked to come up with a plan of attack for addressing the questions, such as a way in which to summarize and present the data. Sometimes we prepare in advance a slew of possible summaries, and we sketch results on the blackboard as students discuss their plans. In the other case, students are given a set of charts and graphs and they are asked to further summarize and interpret the output in order to answer questions from the investigation. Groups write their solutions on the blackboard, and the instructor leads the class in a discussion of the analysis. It would be ideal to lead these discussions in a multimedia classroom where output can be projected onto a screen, or where students can be seated at display stations.

An alternative format for group work that we have tried, has students sign up for a lab in groups of three to five (one of the nonreport labs). Each group meets with the instructor outside of class for assistance in their investigations. They prepare a presentation of their results for the class with their own handouts and materials, and these presentations are synchronized with the curriculum and scheduled throughout the semester.

This second type of group work has the advantage over the former in that students receive individual attention from the instructor, and they have the opportunity to work on

Table 2. Sample Sets of Lab Assignments

Lab	Option		Topics	Additional material
	A	B		
Maternal smoking and infant health I		×	summary statistics	
Student use of video games	×		simple random sampling	ratio estimator
Household radon levels		×	stratified sampling	parametric bootstrap
Patterns in DNA		×	estimation and testing	Poisson process
Crab molting	×		regression	prediction
Hispanic voting behavior	×		regression	weighted regression
Snow gauge calibration		×	simple linear model	inverse regression
Maternal smoking and infant health II	×		multiple linear regression	indicator variables
Transgenic mice with down syndrome	×		analysis of variance	unbalanced design
Designing a paper helicopter		×	analysis of variance	2 × 2 factorial

NOTE: The data and investigations for these labs are available on the Web site: [www.stat.berkeley.edu/users/nolan/stallabs](http://www.stat.berkeley.edu/users/nolan/stallabs).

the lab in greater detail than if it were only discussed in class. However, this approach exposes students to a fewer number of labs. As for the first approach, students seem to really enjoy working on problems in class where they receive immediate feedback from the instructor and see the variety of ideas their classmates have come up with.

### 4.3 Reports

Students work outside of class on a lab for two to three weeks. Their time is spent analyzing data and preparing reports on their findings. The datasets are complex, the analysis is open ended, and they must synthesize their findings coherently on paper. The students find this work very challenging, and we sometimes allow them to work in groups of two or three on their assignments.

It can be difficult to grade the lab reports because the investigations allow students to be creative in their solutions to the problem. We usually break down the score into four equal parts: one part for composition and presentation, including statistically sound statements; one part for a basic minimum set of analyses; one for relevant, readable, and understandable graphs and tables; and finally we make a list of several advanced analyses and look for a subset of these in the write-up. Sometimes we also request an appendix to the report for technical material.

The lab reports typically constitute 40% to 50% of the course grade, the remaining being grades for homework and exams. We also think it is important to include questions related to the labs on exams, because it maintains consistency with the approach we have taken in teaching theoretical statistics through applications.

### 4.4 An Alternative

M. Hahn at Tuft University teaches an undergraduate course similar to ours. The course requires three semesters of calculus and one semester of probability, and there are 15 to 20 students in the class, including statistics majors, mathematics majors, and graduate students in economics, biology, and engineering. The class meets 3 hours a week with the instructor, and there is an optional 1.5-hour problem-and-computer session led by a teaching assistant.

Last semester, she adopted some of our labs for her course, which she ran somewhat differently than discussed so far. The students worked on three labs—maternal smok-

ing and infant health, video game usage, and patterns in DNA—plus a final project of their own choosing. Together, project and lab reports were worth about one-third of the course grade.

The course was run in a seminar style. Students were responsible for reading the background materials on their own. Class time had plenty of discussion and brainstorming on how to solve the problems in the three labs, and instruction focused on the nature of the statistical problems needed to be solved to motivate the theoretical material to be covered. Assignments were not set in advance, but made on a daily or weekly basis. This allowed the course to be somewhat flexible in meeting demands raised by the students. As a result, all the material traditionally included in the course was still covered, new mini-topics were interjected as the need arose, and students felt that class time was essential.

The inclusion of the research project has an advantage over sole reliance on the labs we have prepared in that students tend to have greater enthusiasm for projects of their own choosing, which can boost retention. Students also can present their projects in class and their projects can be the source for theoretical exam questions. Projects can be problematic, however, for it can be difficult to ensure that all students wind up with successful, challenging projects.

### 4.5 Computing

Bringing the computer into the theoretical course enables us to go far beyond the traditionally small, artificial examples found in textbooks. But care is taken to keep the demands made upon the student at an appropriate level. For example, the teaching assistant is mainly responsible for providing assistance in the weekly section meeting on how to use the statistical software. Often the section meets in a laboratory room, where students double up at workstations to work on the assignment and the teaching assistant provides advice as needed. Other times, the assistant distributes handouts with sample code. In a course without a teaching assistant, handouts and office hours may be able to meet the students' needs, or a mini-course on the software could be organized.

## 5. EVALUATION

The student feedback to our labs has been very positive. In anonymous end-of-term evaluations, students report that

the labs helped them better understand statistical theory. They liked the practicality and concreteness of the course and the relevance to the real world. However, not all of the response to the laboratory exercises has been favorable. Most negative reactions center around computer anxiety, and students have requested more help with using software packages. We are considering offering a short course in the use of statistical software, for most students have no prior experience with such software.

In rating the overall effectiveness of the course, the median score was 6 with a mean of 5.6 on a scale of 1 (poor) to 7 (excellent), with 143 students from 5 classes reporting. The departmental average for courses at this level is 5.2.

We have also been pleased with feedback from students after the course. Many students report they have used their lab reports as writing samples for job applications, and that interviewers have focused on one of the lab projects in job interviews.

Hahn also met with success by including labs in her course. She found the students were motivated to learn the theoretical material necessary to understand and complete the lab assignments, and that they approached the topics with vigor since they saw them as relevant. She reports that

The students were extremely enthusiastic about the labs. They took them as a challenge and liked the fact that they were not cookbookish. The labs generated more discussion among the students than anything I have used in the past.

The change in the course also seemed to generate an atypical interest among students in continuing their study of statistics. Students requested advanced statistics courses to be offered, they asked for advice and assistance on studying statistical applications to engineering and biology, and several decided to apply for actuarial jobs.

Although some of our evidence is anecdotal, the success of the labs is encouraging.

## 6. CONCLUSION

Recently Cobb and Moore (1997) called for the design of a better one-semester statistics course for mathematics majors that both strengthens their mathematical skills and integrates data analysis into the curriculum. Others have called for similar courses (Foster and Smith 1969; Hogg et. al. 1985; Kempthorne 1980; Moore and Roberts 1989; Mosteller 1988; Petruccioli, Nandram, and Chen 1995; Whitney and Urquhart 1990) and many consider training in statistical thinking important (e.g., Daisley 1979; Joiner 1989; Schuyten 1991; Riffenburgh 1995; and Nash and

Quon 1996). But it is a challenge to bring statistical thinking into the advanced undergraduate mathematical statistics course. We advocate that we are better able to, as Mosteller puts it, affect broadly the minds and lives of our students if we include in the curriculum case studies that are derived from problems with important scientific questions that contain rich background information, and are integrated with development of the theoretical material.

[Received June 1998. Revised March 1999.]

## REFERENCES

- Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerny, R., Hosnell, T., Hutchinson III, C. A., Kourzarides, T., Martignetti, J. A., Preddie, E., Satchwell, S. C., Tomlinson, P., Weston, P. M., and Barell, B. G. (1990), "Analysis of the Protein Coding Content of Human Cytomegalovirus Strain ad169," *Current Topics in Microbiology and Immunology*, 154, 126-169.
- Cobb, G. W., and Moore, D. S. (1997), "Mathematics, Statistics, and Teaching," *American Mathematical Monthly*, 104, 801-823.
- Daisley, P. (1979), "Statistical Thinking Rather Than Statistical Methods," *The Statistician*, 28, 231-239.
- Foster, F. G., and Smith, T. M. F. (1969), "The Computer as an Aid in Teaching," *Applied Statistics*, 18, 264-269.
- Gonick, L., and Wheelis, M. (1991), *The Cartoon Guide to Genetics*, New York: Harper Perennial.
- Hogg, R. V. (1985), "Statistical Education for Engineers: An Initial Task Force Report," *The American Statistician*, 39, 21-24.
- Joiner, B. L. (1989), "Statistical Thinking: What to Teach and What Not to Teach Managers," *ASA Proceedings*, 150, Alexandria, VA: American Statistical Association, 448-461.
- Kempthorne, O. (1980), "The Teaching of Statistics: Content Versus Form," *The American Statistician*, 34, 17-21.
- Leung, M. Y., Blaisdell, B. E., Burge, C., and Karlin, S. (1991), "An Efficient Algorithm for Identifying Matches With Errors in Multiple Long Molecular Sequences," *Journal of Molecular Biology*, 221, 1367-1378.
- 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.
- Nash, J. C., and Quon, T. K. (1996), "Issues in Teaching Statistical Thinking With Spreadsheets," *Journal of Statistics Education* (online at <http://www.amstat.org/publications/jse/>), 4, 1.
- Petruccioli, J. D., Nandram, B., and Chen, M.-H. (1995), "Implementation of a Modular Laboratory and Project-Based Statistics Curriculum," in *ASA Proceedings of the Section on Statistical Education*, Alexandria, VA: American Statistical Association, pp. 165-170.
- Riffenburgh, R. H. (1995), "Infusing Statistical Thinking into Clinical Practice," in *ASA Proceedings of the Section on Statistical Education*, Alexandria, VA: American Statistical Association, pp. 5-8.
- Schuyten, G. (1991), "Statistical Thinking in Psychology and Education," in *Proceedings of the International Conference on Teaching Statistics*, pp. 486-489.
- Whitney, R. E., and Urquhart, N. S. (1990), "Microcomputers in the Mathematical Sciences: Effects on Courses, Students, and Instructors," *Academic Computing*, 4, 14.