What might a twenty-year old conference tell us about the future of our profession?

George W. Cobb
Mount Holyoke College

Journal of Statistics Education Volume 21, Number 2 (2013),
www.amstat.org/publications/jse/v21n2/cobb.pdf

Key Words: Statistics education; NSF; Assessment; Activities; Simulation; TQM; Randomization; Bayes.

Abstract

The 1993 inaugural issue of the Journal of Statistics Education (JSE) published an article about a small conference for Principal Investigators (PIs) and co-PIs of twelve projects in statistics education funded by the National Science Foundation (NSF). This twenty-year retrospective (1) offers some personal memories related to the founding of JSE, (2) offers some thoughts about the legacies of the twelve funded projects, (3) sets out a sense of how the conference themes have fared over the last twenty years, and (4) indicates what this might suggest about the future of our profession. In conclusion, I argue (briefly) that at this moment in its history, statistics education faces the biggest opportunity and challenge of its last 40 years.

1. Introduction: Return to the electronic Jurassic?

I knew I had officially become a dinosaur when I was invited to look back and write about something I did two decades ago. Truth to tell, although I'm quite happy to munch contentedly on the various appealing Lycopodiceae of my retirement, I'm honored by the invitation. This opportunity arose because twenty years ago, I wrote an article about a conference that Joan Garfield and I organized for Principal Investigators and co-PIs of projects in statistics education that had been funded by the National Science Foundation (Cobb 1993). (It is a pleasure to acknowledge here the help of National Science Foundation (NSF) Program Officers William Haver and Neil Sedransk.)

While I was preparing this article, I found it growing in length as I became increasingly persuaded that for all of us, as we think about the future of our profession, it can be useful to look back on the projects that the NSF chose to fund between1990 and 1992, reflect on their similarities and differences, and detail their relation to the current state of statistics education. It is my goal in what follows to encourage readers to integrate my account of what was in progress twenty years ago with their own sense of where we are now and where we may be headed.

As context for this thinking, I begin with an eight point excavation of the fossil record from 1992:

(1) By today's standards, **personal computing** in 1992 was embryonic: noisy dot matrix printers, only green on the screen, 3.5" floppies with a capacity of 128k, and equally primitive software. For an example of where things stood, Francis Anscombe's pioneering 1981 book *Computing in Statistical Science through APL* about real-time interactive statistical computing was only a decade old, and had not taken the world by storm. In 1992, using any kind of software for teaching statistics was still considered cutting edge. Worse yet, APL in 1992, like R in 2000, appealed mainly to a select group of enthusiasts. Using software came with high up-front costs, including capital expenses for the machines, licensing for the software package, and in-class time to teach how to use them. Colleges were not used to paying for software, and were reluctant to open a new line in their budgets. Any faculty adventurous enough to use software in their teaching had to rely on packages that had originally been designed for main-frame computers.

(2) The **Internet** was in its toddlerhood. As just one example, the Mathematical Association of America submitted a "STATS" workshop proposal (Statistical Thinking and Teaching Statistics) to NSF that called for creating what would later come to be known as "online communities" for teachers of statistics. Back then, NSF's reviewers were skeptical that this e-mail component was either workable or worthwhile. Fortunately, the project had enough other strengths that it was funded anyway. As we expected, most participants did not yet have e-mail accounts, so my co-PI Mary Parker had to work patiently to help participants sign up for accounts with AOL (America On Line) and learn how to use them. The NSF grant had to pay for the accounts. Today, of course, we take it for granted that everyone has an account and knows how to use it.

(3) If the Internet was in its toddlerhood in 1992, **the Web** was still crawling on all fours wearing diapers. Mosaic, the first popular browser, was developed late that year and first released to the public in 1993 (see http://www.ncsa.illinois.edu/Projects/mosaic.html for more details). As the main way to explore the web, it was clunky: at best it was a "browser" in the sense of an ambling ruminant, definitely not yet worthy of being called a search engine.

(4) **On-line resources** were non-existent or nascent in 1992: no DASL (Data and Story Library, http://lib.stat.cmu.edu/DASL/), no ISOSTAT (Isolated Statisticians) list serve, no CAUSE (Consortium for the Advancement of Undergraduate Statistics Education, http://www.causeweb.org), no SERJ (Statistical Education Research Journal, http://tinyurl.com/nnl5uax). Carnegie Mellon University's StatLib

(http://lib.stat.cmu.edu/) was only 3 years old, and Chance News (http://test.causeweb.org/wiki/chance/index.php/Main_Page) was first published in September, 1993.

(5) **Assessment** was just beginning to be recognized as important, and had yet to be developed specifically for use in statistics: no ARTIST (Assessment Resource Tools for Improving Statistical Thinking, https://apps3.cehd.umn.edu/artist/), and no CAOS (Comprehensive Assessment of Statistics) test, for example (see delMas, Garfield, Ooms, and Chance 2007, for more details).

(6) **Conferences** with a focus on statistics education were few and rare. There was an annual conference on improving the teaching of statistics in schools of business, and Bob Hogg's famous ASA (American Statistical Association) Louisville winter conference on teaching statistics had just taken place in 1992. The first three International Conference on Teaching Statistics (ICOTS) gatherings had been held in 1982, 1986, and 1990, but the International Association for Statistical Education (IASE) was not formed until 1991. We had no USCOTS (United States Conference on Teaching Statistics), no Chaska conference (held in 1995) to explore and establish regional meetings of isolated statisticians, no regional ISMs (Isolated Statisticians' Meetings, begun in 1996), and no SRTL (Statistical Reasoning, Thinking and Literacy) conferences.

(7) **Organizations**. By 1992, ASA had long had a Section on Statistics Education, largely focused on graduate programs at universities. In addition ASA supported the Joint American Statistical Association (ASA)/National Council of Teachers of Mathematics (NCTM) Committee, focused on the Promethean task of bringing statistics to K-12. The Joint ASA/Mathematical Association of America (MAA) Committee on Undergraduate Statistics was brand new and without official status; there was no ISOSTAT group, no SIGMAA (Special Interest Group of the Mathematical Association of America).

(8) **Enrollments**. By 1992, although statistics enrollments had been growing steadily for more than a decade, the numbers were still puny by today's standards. For example, in 1965, at two-year colleges, there had been only one section of statistics for every ten sections of calculus. By 1990, the statistics-to-calculus ratio had grown five-fold, from 1:10 to 5:10 (Albers, Loftsgaarden, Rung, and Watkins 1992). Although the ratio reflects rapid growth, the absolute numbers were small in 1990, and have more than doubled since then. At U.S. four-year colleges and universities, fall semester enrollments in elementary statistics have been (Blair, Kirkman, and Maxwell 2013; Lutzer, Rodi, and Maxwell 2007):

1990	169,000
1995	208,000
2000	245,000
2005	260,000
2010	371,000

> The Advanced Placement (AP) statistics examination would not come into being until 1999, when only 7,500 students took the exam (Franklin, Hartlaub, Peck, Scheaffer, Thiel, and Tranbarger 2011). In 2013, the number was more than 23 times as large, at 175,000 (Rossman 2013).
>
> What message should we take from these numbers? To my saurian eye, if we judge by today's standards, statistics education of the 1990s was barely getting itself unstuck from decades in the tar pits of intellectual neglect. PIs who could foresee our evolution into a dominant force in the emerging science of data were indeed the Darwins of that era.

In that 1990 context of primitive technology, scarce resources, modest interest, and organizational weakness, the founding of JSE had the potential to make a huge difference, as in fact it has done.

In this look back, I present my thoughts under three headings: personal memories, funded projects, and conference themes. I first offer some vignettes related to the founding of JSE, next offer some thoughts about the legacies of the twelve funded projects, and then set out my sense of how the four themes of the conference have fared over the last twenty years, and what this might suggest about the future.

2. Personal memories: Three salutes

Daniel Solomon. Dan had the vision and follow-through to seek and find funding for the organizational meeting that led to the creation of JSE; Dan issued the invitations that brought us together; Dan chaired our meetings and somehow managed to keep us all efficiently on task. At the same time, Dan also managed somehow to keep himself in the background. No wonder North Carolina State University made him Department Chair and then later promoted him to Dean. Thanks, Dan.

E. Jacquelin Dietz. Dan, with his characteristic mix of determination, diplomacy, shrewdness, and modesty, has long kept his eye out to recognize talented women statisticians and to use the advantages of his administrative positions to help ensure that they get fair consideration. Jackie Dietz was a perfect choice to become founding editor of JSE. Like Dan, Jackie was (and is) knowledgeable, accomplished, thoughtful, organized, and efficient. Like Dan, Jackie has a remarkable ability to get appropriate people to get good things done, all without putting herself forward. To cast myself as a mule (metaphorically, and only momentarily): During her editorship, Jackie never raised a stick, and she dangled the carrot so skillfully that I was always eager to follow her lead without ever seeing orange. It is largely thanks to Jackie's vision, leadership and effort that JSE has become so firmly established and so valuable a resource. Thanks, Jackie.

Robert Hogg. Finally, and more personally, I want to salute Bob Hogg. Readers of JSE who have had the privilege to spend time with Bob may not learn much new from this salute: It's vintage Hogg. (No doubt Bob himself would want me to turn "vintage Hogg" into a joke about aged pork.) For those readers who have not had the privilege of time with Bob, I want to offer a personal memory. One dinner during our JSE meeting took place at an Asian restaurant. It was a social occasion, not a working meal, with spouses present, and alone among all the grown-ups, there was one four-year-old girl, my daughter. She had every reason to feel overwhelmed by the adult presence, to feel bored, restless, and entitled to behave badly. But: I'll always remember how Bob had her fully engaged, watching with him while they talked together about the carp in a big tank. As I say, and as those of you who know Bob will recognize, it was vintage Hogg. Thanks, Bob.

3. The funded projects: How have they fared?

The original article about the NSF projects (Cobb 1993) described each project in considerable detail, with even more detail about each project in an appendix, and links there for readers to follow to get still more detail. Because of the diversity of projects, my 1993 report included some tables of the "compare and contrast" variety. I won't repeat the descriptions or those tables here, because I hope to avoid duplicating what can be found elsewhere. At the same time, I want to include enough background to allow a reader to make sense of what I say without having to refer back or read the older article. (However, I do urge readers to read or reread David Moore's interview (Moore 1993) with Fred Mosteller in that first issue of JSE.)

In this section, I first give some general comments about the funded projects as a whole under four sub-headings that correspond to some of the comparisons in the tables of the 1993 article: sources of data, assessment, minority students, and large classes. I end the section with some thoughts about the legacies of the projects and what this may suggest about the future.

3.1 Three sources of data

Taken together, the projects relied on three different sources of data: archives, activities, and simulation. Archived data can have unmistakable real-world import and a rich complexity that is not possible with either class activities or simulated data, but students are not involved with producing the data. Simulated data sets give the instructor control over the conceptual focus, and repetition gives students experience with variability, but of course the data and context are not real. Hands-on activities, despite their toy-like quality, do involve students with data production; their experience with variability is immediate and concrete; and most students are motivated to analyze data they have helped create.

Although the advantages and disadvantages of the various kinds of data have not changed much since 1992, the priorities have changed.

Over the last 20 years, since 1992, archived data sets from real studies have become all but obligatory in textbooks and courses. This near-universality of real data for examples and exercises makes it hard to remember that this is a comparatively recent development. In a focus group on statistics, David Moore urged us all "to automate calculations and graphics" (as cited in Cobb 1992, p. 7). That he felt the need to say so is a mark of where things stood a mere twenty years ago.

Using activities in class to generate data to engage students has become much more common, largely, in my opinion, thanks to the leadership of Richard Scheaffer, who assembled a team of statisticians determined to change the student attitude that, in the words of Bob Hogg, in 2004, “If it moves, it’s biology; if it changes color, it’s chemistry; if it breaks, it’s physics; if it puts you to sleep, it’s statistics,” (as quoted in CHANCE news at http://www.dartmouth.edu/~chance/chance_news/current_news/current.html).

Simulation-based data has had a more mixed history since 1992. On one hand, the use of simulation has declined as a way to illustrate abstract results, especially results from probability theory, in parallel with the decline of the role of probability in the introductory statistics course. On the other hand, the use of simulation as an engine for inference has been gaining momentum, first in statistical practice (the bootstrap, exact inference, and Markov Chain Monte Carlo) and more recently in teaching of the introductory course. See, for example, the randomization-based approaches of NSF-funded projects led by Rossman, Chance, Cobb and Holcomb (DUE-CCLI 0633349; “Concepts of Statistical Inference: A randomization-based curriculum”), Garfield, delMas, and Zieffler (“Adapting Innovative Materials in Statistics,” http://www.tc.umn.edu/~aims/), Tintle, VanderStoep, Holmes, Quisenberry, and Swanson (2011), and West and Woodard (DUE 0817262; “INCIST: Improving National Acceptance of Computing Intensive Statistical Techniques”). See also Tabor and Franklin (2013).

3.2 Assessment

In 1992, assessment was a major and explicit focus in just two of the twelve projects. Over the last two decades there has been steady growth in our recognition of the importance of assessment in shaping what we teach, in what NSF decides to fund, and in how we persuade colleagues to consider making or supporting changes. Today, the chance of funding for a proposal without serious attention to assessment would surely be far less than 10 in 12, almost surely far less than 2 in 12. I have no doubt that this growth in understanding owes much to the leadership of my co-PI Joan Garfield, and I am inclined to think of the conference as a turning point because Joan made assessment one of its four themes.

3.3 Minority students

Of the twelve funded projects, only three had a major focus on recruiting minority students to statistics. Now, twenty years later, we do much better, but not enough better. We need more projects like Lisa Harlow’s NSF-supported QTUG (Quantitative Training for Under-represented Groups, see https://sites.google.com/site/qtugsmep/home for more details) which organizes training workshops in conjunction with meetings of the Statistics Division of the American Psychological Association. In addition to this work, Lisa is adviser to other NSF-funded projects, including Lisa Dierker's at Wesleyan University (Dierker, Kaparakis, Rose, Selya, and Beveridge 2012), and Nathan Tintle's at Dordt College (Tintle et al. 2011).

3.4 Large classes

Only two of the projects dealt explicitly with large class sizes. Class size remains an ongoing challenge, despite new developments such as the use of clickers, social media, and massive open on-line courses (MOOCs). It is a stretch, but one could regard MOOCs as an on-line resource, and in that sense regard the project of William Notz et al. (see Project 7 in the Appendix) as an ancestor of sorts. Also as a stretch, one could regard the use of clickers as a high-tech extension of Scheaffer's activities for teaching statistics (see Project 8 in the Appendix).

3.5 Legacies of the projects

All twelve projects had strong reviews at the time, and I have every reason to think that all succeeded in their local environments. Nevertheless, as David Moore reminded us at the 1992 conference, "… all innovations succeed in the hands of their innovators …." For lasting impact, changes must be institutionalized (Moore 1995). Now, with two decades of hindsight, we can see more clearly which projects have had the biggest impact on a national scale. Here are four:

The Hogg workshop (See Project 2 in the Appendix; see also Hogg (1990; 1992))
The Hogg conference brought together 29 statistics educators to spend time working toward a consensus about principles and recommendations to improve the teaching of statistics. It is my personal opinion first, that Hogg's conference has had a large and lasting impact, and second, that its impact has been not so much from the long list of principles and recommendations, but, rather, more as a model for successor conferences. In particular, four conferences strike me as particularly noteworthy; I name them here after their principal organizers:

(1) The (Daniel) Solomon conference at North Carolina State University in 1991. This is the meeting that led to the founding of JSE.

(2) The (Thomas) Moore conference in Chaska, Minnesota in 1995. This conference led to the formation of regional conferences for isolated statisticians, one of which has been held annually ever since.

(3) The (Richard) Scheaffer conference at ASA in Alexandria in 1999. This conference, which was part of ASA's Undergraduate Statistics Education Initiative (USEI), partially funded by the NSF, led to a much larger conference headed by David Moore, in advance of the Joint Statistical Meetings in Indianapolis.

(4) The (Deborah) Rumsey/ (Joan) Garfield conference at the Ohio State University which led to the formation in 2006 of the Consortium for the Advancement of Undergraduate Statistics Education (CAUSE).

Notz, *et al*: DASL. (See Project 7 in the Appendix; see also Velleman, Hutcheson, Meyer, and Walker 1996).
The Notz project was far ahead of its time in recognizing the potential value of on-line resources, and in exploiting that potential to make it easier for teachers of statistics worldwide to use compelling contexts and real data in their classes. One enduring product of the project is the Data and Story Library (DASL, http://lib.stat.cmu.edu/DASL/), which has served as a prototype for many successors, including JSE's regular section "Data Sets and Stories."

Snell *et al*: the Chance project (See Project 9 in the Appendix; see also Snell (1992))
Ever marching to his own drummer, Laurie Snell used his NSF grant to create a course like no

other. (John Kemeny, Mathematician and former President at Dartmouth, with his own mix of generosity and modesty, used to tell his students that Snell was the main source of the ideas behind their many collaborations.) Alone among the twelve projects, Snell's had no particular agenda with regard to statistical content. The main goal of his project can be condensed to a single word: motivation. Snell's plan was to use reports of current events and scientific studies of general interest to entice students into talking together and thinking actively about statistical logic and concepts and logic. With pioneering vision for what could be done with technology, Snell pushed the limits of what the Internet could support at the time. He was among the first to post videos of lectures and he started *Chance News*, which might be considered one of the first blogs. I see a link, conscious or not, between the Chance project in the 1990s and David Moore's 1979 *Statistics: Concepts and Controversies* (and its predecessors Moroney's 1951 *Facts from Figures*, Bross's 1953 *Design for Decision*, and, of course, Tanur and Mosteller's 1972 *Statistics: A Guide to the Unknown).*

Scheaffer *et al*.: Activity-Based Statistics (See Project 8 in the Appendix; also Scheaffer, Gnanadesikan, Watkins, and Witmer 1996).

The main goal of Scheaffer's project, as with Snell's, was to motivate students to want to learn statistics. Snell used current issues and controversies, with no particular technical agenda. Scheaffer used hands-on activities, most of them designed with particular topics or lessons in mind. (For example, the "random rectangles" activity was designed to show how random samples beat judgment samples.) Both Scheaffer's and Snell's projects continue to influence the teaching of statistics for the same two reasons. First, both projects created lasting resources. Snell's on-line bulletin *Chance News* continues still, thanks largely to William Peterson and Jeanne Albert, as a resource for teachers and as an ongoing tribute to the memory of Snell and his influence on the teaching of statistics. The book *Activity-Based Statistics* by Scheaffer and colleagues (1996) continues to be a resource that myriad teachers rely on, both as a source of specific activities, and for the approach to teaching that those activities illustrate and inspire. Second, both Scheaffer and Snell recruited broadly based teams of statistics educators from across the country to serve as advisers and advocates, and gave talks at national meetings and held workshops to publicize their approaches. Snell's Chance course, for example, was developed jointly and then taught by faculty from Dartmouth, Grinnell, King, Middlebury, Princeton, and Spelman.

What distinguishes these four projects?

Two features stand out. All four of the projects just described were expensive in comparison with most of the others, and all four were correspondingly much more ambitious in terms of their plans for dissemination. Their lasting nationwide impact seems due in no small part to their higher cost and systematic efforts to publicize their work.

All the same, I remain convinced that there is also value in less expensive projects that are more local in scope.

As I look back, I think history strongly supports the evolution of NSF's grant programs since 1992 toward a distinction between Phase I (proof of concept) and Phase II (dissemination).

4. Statistics education past, present, and future: The four conference themes.

Our plan for the 1992 conference was to keep formal presentations to a minimum in order to give the PIs many chances to talk among themselves in small groups. To provide some structure for these conversations, we chose four themes:

(a) challenging the usual assumptions, (b) anticipating resistances to change,
(c) total quality management, and (d) assessment.

4.1 Challenging the usual assumptions

As I wrote in 1993, "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" (Cobb 1993).

Snell's project violated all four assumptions, making his Chance course by definition an extreme outlier among the usual introductory courses. Sadly, most courses today remain faithful to (a) – (c). Only (d) is being systematically varied by a rebellion now known as the "flipped classroom," but it started in the 1990s by (among others) the innovative book by Rossman (1992). As I urged orally in 2005 at USCOTS and electronically in Cobb (2007), we should commit ourselves to finding workable alternatives to (a) – (c). Fortunately, a number of NSF-supported efforts are underway, and a pioneering book (Lock, Lock, Morgan, Lock, and Lock 2013) has been published.

4.2 Anticipating resistances to change

Today, as I look back on the conference, I consider my 1992 view of possible resistances to have been naïve. Back then, I was thinking about resistance more from individuals than from institutions. Now, I have a much better appreciation for David Moore's dictum that "… all innovations succeed in the hands of the innovator …" and that to endure, changes must be built into the structure of institutions. (The theme of the 2013 U.S. Conference on Teaching Statistics – Making Change Happen – is a mark of how resistance to change continues to resist change.)

I wish we had spent more time at the conference thinking about institutional resistances. Two in particular come to mind: inertia and economics.

Inertia: Whenever you have even a moderately large group of decision makers, that body tends to stay at rest. With a variety of opinions, a consensus for change is hard to achieve, the status quo attracts with gravitational force, and the energy required to achieve escape velocity is correspondingly high.

Economics: Instead of "economics," I might have said "turf wars." Most of the time any decision about faculty hiring devolves into a zero sum game. Suppose, as background, and as my former Dean asserted, that liberal arts colleges are the primary educational institutions that meet two criteria:

- the teaching load is light enough to allow faculty to stay on top of research in their field and to experiment with curriculum; and

- the reward system is flexible enough to offer an incentive for such curricular experimentation.

It follows, says my former Dean, that liberal arts colleges are the primary places where cutting edge research and thinking gets translated into the undergraduate curriculum. As just one example, in the 1950s, when my father was studying for his Ph.D. in statistics, linear algebra was considered a graduate level course. At the same time that he was in graduate school, John Kemeny and Laurie Snell and their colleagues were developing a curriculum that made linear algebra part of their 1959 book *Finite Mathematical Structures* to be used for a second year undergraduate course at Dartmouth College.

With this premise about liberal arts colleges as background, here is the institutional resistance: At liberal arts colleges, statisticians are housed in departments of mathematics. As a consequence of this housing arrangement, any new faculty position in statistics comes at the cost of a position in mathematics. Is it any wonder that enrollments in statistics courses taught in mathematics departments have been rising much more rapidly than numbers of faculty positions in statistics? There is nothing malicious about this growing imbalance. In a perverse way, it can be taken as a sign of progress, in that mathematicians are coming to recognize that statistics is not a branch of mathematics.

4.3 Total Quality Management (TQM) in teaching and learning?

Now that the U.S. automobile industry is in recovery, TQM has lost some of its former cachet. All the same, the principles remain sound. The best summary I know is a set of three points quoted in Cobb 1993:

(1) Emphasis on data. (If you don't know what to measure, measure anyway: you'll learn what to measure.)

(2) Process orientation. (Constancy of purpose is essential; improvements must be institutionalized.)

(3) Customer focus. (We must abandon the notion, "I know best; I'm the expert.")

As I see it, all three principles have been thoroughly embraced by statistics educators and are being effectively implemented. We don't call it TQM or "six sigma" in education, instead we call it …

4.4 Assessment

Of the four themes chosen for the conference, assessment has turned out to be the most prescient choice. Of the dozen projects represented at the conference, only two (!) explicitly mentioned assessment. These days, of course, it is hard to imagine *any* proposal being funded without careful attention to assessment. The last twenty years have seen systematic efforts to develop a variety of tests and questionnaires to measure the effectiveness of projects in statistics education.

4.5 Overlooked or underestimated trends

In my mind three things stand out now for having not stood out then. All three are linked to how rapidly the power and speed of computing would increase:

(1) Bayesian inference in statistical applications. At the time of the conference, the seminal paper by Gelfand and Smith (1990) was new, and Markov Chain Monte Carlo (MCMC) had not yet risen to prominence as a tool for computing Bayesian posterior distributions for hierarchical models. Our conference did not address the role of Bayesian logic in the introductory course.

(2) Randomization-based inference. Although the article by Diaconis and Efron (1983) on computer intensive methods in statistics had been published a decade before the conference and was well-known, we did not explicitly address either the bootstrap or permutation tests. Both methods have emerged as effective methods for teaching the logic of inference in a beginning course.

(3) Visualization of multivariate data. Two decades ago, it was a novelty to teach about residual plots and normal plots in a beginning course. Few courses managed to get to multiple regression at all, and things like added variable plots, color graphics capable of showing four variables at a time, or spinning 3D plots on desktop computers were cutting edge. The exciting kinds of animated displays made accessible by Hans Rosling (2009) were far in the future.

5. Conclusion: looking back, looking ahead

I conclude with eleven hortatory imperatives, grouped, somewhat arbitrarily, under four sub-headings.

5.1 Old lessons reinforced

(1) Use real data.

(2) Teach with technology. Use computers for graphics and number crunching.

(3) Ensure active learning. Use class activities to generate data.

5.2 Lessons learned

(4) Institutionalize change. Dissemination matters.

(5) Assess what you do. Data is important both for continuous improvement and for convincing others to institutionalize your changes.

(6) Put what you create on line. Good resources make life easier for individual teachers. On line resources help institutionalize your changes.

5.3 Ongoing challenge

(7) Accept responsibility for outreach. Our profession needs to do more for under-represented groups. Doing so helps the students and it helps us all.

5.4 Newer challenges

These correspond one-one with the three items in sub-section 4.5

(8) Accept Bayes but keep Fisher. David Moore (1997) has argued forcefully and persuasively that we should not be in a hurry to teach Bayesian methods in a beginning statistics course. At the time that Moore wrote his paper, I was convinced that he was right, and I remain convinced that at the time of his writing, he was right. Now, however, as I write this, I think times have changed, and that it is worthwhile to revisit the question. Here are Moore's four main objections, together with my sense of how computers have changed things during the intervening years:

(a) Bayes is little used in practice. (True then, but not now.)

(b) There is no consensus on standard approaches. (True then, but a consensus is emerging.)

(c) Conditional probability is too hard. (Simulation reduces posterior probabilities to fractions of the form #Yes/# Reps, where #Reps is the number of repetitions that meet the relevant condition.)

(d) Teaching Bayes spends too much time on probability at the expense of data production and analysis. (Simulation eliminates the need for formal probability.)

I think it is only a matter of time before we give in to the pressure to teach Bayesian posterior intervals in a beginning course. At the same time, I do not think that this change automatically requires that we abandon Fisher's use of p-values for inference. In this advocacy of both approaches, I follow Dempster (1971), who argues that we should use Fisher's approach for testing and model selection, then use Bayesian methods for estimation within the framework of the chosen model.

(9) Embrace de-mathefication; rely on simulation. These days, we take it for granted that computers should relieve students of the tedium of drawing graphs and computing standard deviations or regression coefficients. Such relief is truly welcome, not least because it allows courses to pay more attention to diagnostics, but we should not stop where we are now. We should vigorously explore the alternative curricula made possible by the computing power to rely on randomization-based methods. These methods

provide a simpler and more direct introduction to the logic of formal inference, which we can now teach without reliance on the normal distribution and central limit theorem. Moreover, re-randomizing can be taught as a simulated repetition of the randomized method used to produce the observed data. The direct link between production and probability encourages students to regard statistics as a unified logical system rather than a mosaic of little fragments.

(10) Push visualization. My choice of the verb "push" is deliberate. As Martin Wattenberg has said, "Visualization is the gateway drug to statistics" (quoted in Aldhous 2011, p. 44). In a series of articles and presentations, Christopher Wild and his colleagues (e.g., Wild, Pfannkuch, Regan, and Parsonage 2013) have been pushing visualization as the best way to give students the satisfaction and excitement that can come from discovering meaningful patterns in rich data sets. Wild is clear that leading with visualization need not mean that we dispense with inference in the first course. (Even with visualization as the gateway drug, we needn't regard inference as the methadone that induces nausea in users.)

(11) Seek a new curricular integration. I am convinced now, in my retirement, that at this moment in time, statistics education faces a bigger curricular challenge and opportunity than at any previous time during last 40 years. We have made many big changes: we use real data; we automate graphics and number crunching; we use fewer formulas and recipes; we teach exploration and the "model-data dialogue;" we put more emphasis on data production, less on formal probability, and so on. Despite the enduring import of these changes, I regard them as "normal" in T.S. Kuhn's (1962) sense of staying within an established paradigm: We still teach a beginning curriculum that is driven and structured by the centrality of its reliance on the normal approximation. As I see it, computers have already opened the door to a paradigm shift in statistical *practice* where the primacy of the normal approximation has given way to the bootstrap and MCMC. I suggest that computers have simultaneously opened a second door, inviting us to make a parallel shift in the content and organization of our beginning course. It will take time, effort, experimentation, and settling out before we have answers. Currently, we have mainly questions, and even the questions are still being articulated. Here are just two: First: For introducing students to inference, what is the right mix of Fisher, Bayes, and Neyman-Pearson? (Fisher and Neyman had sharply differing views about inference, and yet we typically teach an amalgam that blurs the distinctions they fought over.) Second: For the beginning course, what is the right balance between inference, design, and graphical exploration of multivariate data? (Inference is truly difficult, especially in a first course; so is design -- properly taught -- but design is hard in a different way. By comparison, graphical data exploration is the dessert that makes the other two seem like spinach and liver. What is the best way to achieve a balanced diet?)

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

2. Workshop on Statistical Education
Robert V. Hogg, Department of Statistics and Actuarial Science, University of Iowa

3. Development of a Multi-Media-Based Statistics Classroom
Fidelis Ikem, Department of Information Systems and Decision Sciences, Virginia State University

4. Undergraduate Statistics Laboratory
Rhonda C. Magel, North Dakota State University

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

6. Improving the Undergraduate Statistical Education of Engineers
Peter Nelson, Department of Mathematical Sciences, Clemson University

7. Technology-Based Learning: Exploring Statistical Concepts and Methods
William I. Notz (et al.), Department of Statistics, Ohio State University

8. An Activity-Based Introductory Statistics Course for All Undergraduates
Richard L. Scheaffer, Department of Statistics, University of Florida

9. CHANCE: Case Studies of Current Chance Issues (An Introductory Mathematics Course)
J. Laurie Snell, Department of Mathematics, Dartmouth College

10. Elementary Statistics Laboratory Course Development
John D. Spurrier, Department of Statistics, University of South Carolina

11. Statistical Computing Laboratory
Dolores M. Tichenor, Department of Mathematics, Tri-state University

12. Materials for a Computer-Based Introductory Statistics Curriculum Using Actual Data
Bruce E. Trumbo, Department of Statistics, California State University - Hayward

---

Acknowledgements

In writing this article I have been powerfully reminded of the influence that five giants of statistics education have had on our profession generally and on my life personally: Robert Hogg, David Moore, Fred Mosteller, Richard Scheaffer, and J. Laurie Snell. I write with highest regard and enduring thanks to Bob, David, Fred, Dick, and Laurie.

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, D.C.: Mathematical Association of America.

Aldhous, P. (2011), "Culturelab: The Art of Data," *New Scientist*, 5, 44.

Blair, R., Kirkman, E. E., and Maxwell, J. W. (2013), *Statistical Abstract of Undergrtaduate Programs in the Mathematical Sciences in the United States: Fall 2010 CBMS Survey*, Washington, D.C.: American Mathematical Society.

Cobb, G. W. (1992), "Teaching Statistics: More Data, Less Lecturing," in *Heeding the Call for Change*, ed. Lynn Steen, MAA Notes No. 22, p. 3-43, Washington, D.C.: Mathematical Association of America.

Cobb, G. W. (1993), "Reconsidering Statistics Education: A National Science Foundation conference," *Journal of Statistics Education*, 1. Available at: www.amstat.org/publications/jse/v1n1/cobb.html.

Cobb, G. W. (2007), "The Introductory Statistics Course: A Ptolemaic Curriculum?," *Technology Innovations in Statistics Education*, 1. Available at: http://repositories.cdlib.org/uclastat/cts/tise/.

delMas, R., Garfield, J., Ooms, A., and Chance, B. (2007), "Assessing Students' Conceptual Understanding After a First Course in Statistics," *Statistics Education Research Journal*, 6, 28-58. Available at: https://www.stat.auckland.ac.nz/~iase/serj/SERJ6%282%29_delMas.pdf.

Dempster, A. P. (1971), "Model searching and estimation in the logic of inference," in V. P.Godambe and D. A. Sprott, *Foundations of Statistical Inference*, pp. 56-78. Holt Rinehart and Winston.

Diaconis, P., and Efron, B. (1983), "Computer-intensive Methods in Statistics," *Scientific American*, 245, 116-130.

Dierker, L., Kaparakis, E., Rose, J., Selya, A., and Beveridge, D. (2012), "Strength in Numbers: A Multidisciplinary, Project-based Course in Introductory Statistics," *Journal of Effective Teaching*, 12, 4-14. Available at: http://uncw.edu/cte/ET/articles/Vol12_2/Dierker.pdf.

Franklin, C., Hartlaub, B., Peck, R., Scheaffer, R., Thiel, D., and Tranbarger, K. (2011), "AP Statistics: Building Bridges Between High School and College Statistics Education," *The American Statistician*, 65, 177-182.

Gelfand, A. E., and Smith, A. M. (1990), "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of The American Statistical Association*, 85, 398-409.

Hogg , R. V. (1990), "Statisticians Gather to Discuss Statistical Education," *Amstat News*, 169, 19-20.

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

Kuhn, T.S. (1962), *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

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

Lock, R.H., Lock, P.F., Morgan, K.L., Lock, E.F., and Lock, D.F. (2013), *Statistics: Unlocking the Power of Data*, Hoboken, NJ: Wiley.

Lutzer, J., Rodi, S. B., and Maxwell, J. W. (2007), *Conference Board of the Mathematical Sciences (CBMS) Survey Reports 2005*, Washington, D.C.: American Mathematical Society.

Moore, D. S. (1993), "Interview with Fred Mosteller," *Journal of Statistics Education* 1. Available at: http://www.amstat.org/publications/jse/v1n1/moore.html.

Moore, D. S. (1995), "The Craft of Teaching" *MAA Focus*, 15, 5-8.

Moore, D. S. (1997), "Bayes for Beginners? Some Reasons to Hesitate," *The American Statistician*, 51, 254-261.

Rosling, H. (2009), "200 years that changed the world." Available at: http://www.gapminder.org/videos/200-years-that-changed-the-world/.

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

Rossman, A. (2013), Personal communication.

Scheaffer, R. L., Gnanadesikan, M., Watkins, A. E., and Witmer, J.A. (1996), *Activity-based Statistics : Instructor resources*, New York : Springer.

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, p. 281-297, Washington, D.C.: Mathematical Association of America.

Tabor, J., and Franklin, C. (2013), *Statistical Reasoning in Sports*, New York: W. H. Freeman.

Tintle, N., VanderStoep, J., Holmes, V., Quisenberry, B., and Swanson, T. (2011), "Development and Assessment of a Preliminary Randomization-Based Introductory Statistics Curriculum," *Journal of Statistics Education*, 19. Available at: www.amstat.org/publications/jse/v19n1/tintle.pdf/.

Velleman, P. F., Hutcheson, M. C., Meyer, M. M., and Walker, J.H. (1996), "DASL, the Data and Story Library: An Online Resource for Teaching Statistics," in *The American Statistical Association Proceedings of the Section on Statistical Education*, 45-49, Alexandria, VA: American Statistical Association.

Wild, C.J., Pfannkuch, M., Regan, M., and Parsonage, R. (2013), "Next Steps in Accessible Conceptions of Statistical Inference: Pulling ourselves up by the bootstraps," *International Statistical Review* (in revision).

George W. Cobb
Robert L. Rooke Professor of Statistics, emeritus
Mount Holyoke College
GCobb@MtHolyoke.edu
