

# New Pedagogy and New Content: The Case of Statistics

David S. Moore

*Department of Statistics, Purdue University, West Lafayette, IN 47907 USA*

## Summary

Statistical education now takes place in a new social context. It is influenced by a movement to reform the teaching of the mathematical sciences in general. At the same time, the changing nature of our discipline demands revised content for introductory instruction, and technology strongly influences both what we teach and how we teach. The case for substantial change in statistics instruction is built on strong synergies between content, pedagogy, and technology. Statisticians who teach beginners should become more familiar with research on teaching and learning and with changes in educational technology. The spirit of contemporary introductions to statistics should be very different from the traditional emphasis on lectures and on probability and inference.

*Key words:* Statistical education; Technology for teaching; Pedagogy.

## 1 Introduction

No one concerned about the teaching of the mathematical sciences can have missed the movement to reform teaching at all levels. In the United States, new standards for school mathematics (NCTM, 1989), calculus reform (e.g., Steen, 1988), and broader manifestos by the National Research Council (1990, 1991) have all called for change in traditional approaches. The teaching of any active branch of knowledge, like the church, is of course “reforming and ever to be reformed”. Calls to modernize what we offer students are always with us. What is striking about the current reform movement is not only its momentum but the fact that it centers on pedagogy as much as content. We ought, say the reformers, to radically alter our style of teaching. In this paper I examine critically the reform thesis as it affects statistics. I am particularly concerned with the interaction between new content and new pedagogy in the teaching of statistics to beginners at the university level.

Sections 2, 3, and 4 provide some background. They briefly summarize the context of reform, the elements of reformed pedagogy, and trends in the content of introductory statistics teaching. I offer some opinions about both pedagogy and content, too briefly to do more than (I hope) provoke thoughtful reaction but fortified with references for those who want to read further. Section 5 announces my main thesis: that changes in content, pedagogy, and technology reinforce each other in a way that strengthens the case for change in our teaching. The remaining sections of the paper apply the method of considering all three domains to comment on several specific developments in statistics teaching, in particular on several new and old technologies.

## 2 The Context of Reform

The call for change takes place in a context, the first element of which is what David Vere-Jones (1995) calls *the democratization of mathematics*. Mathematics is no longer embedded in a deliberately

elite curriculum, but is seen as a necessary part of the general education of all citizens. The gap between the level and quantity of school mathematics studied by men and women, by students from various social strata, and by cultural minorities and dominant cultural groups, is steadily shrinking in most nations. The proportion of secondary school graduates who go on to post-secondary education has increased sharply. Even in the United States, where this proportion had already reached 50% in the early 1960s and remained there through the early 1980s, it surpassed 60% in the 1990s. We therefore teach a more diverse and less specialized student population than in the past. In some countries, such as South Africa, the change is revolutionary in scope and rapidity.

Democratization tends to move mathematical studies away from the esoteric toward the immediately useful. University faculty rightly want to educate, but most of our students—also rightly—seek career preparation as well. We argue that we are preparing students for a career, not simply for their first year on the job. We argue that larger principles and deeper understanding will carry students further than a few specific skills. All true. Yet it is difficult for us who have been socialized into the peculiar culture of university faculty to recognize how esoteric we have allowed university mathematics in particular to become. We imagine (incorrectly) that dominance of the abstract over the concrete, absence of ties to applications, and an emphasis on rigor over fluency of use are inherent in the discipline. We value, in Richard Feynman's words, precise language over clear language. Reformers urge a change of culture toward the concrete, toward applications, toward ability to *use* mathematical concepts and tools over rigor of detail. They offer pedagogical reasons, but they are also responding to the pressures of democratization. This is an opportunity for statistics: as mathematical studies shift toward a more utilitarian approach, a larger place for statistics (understood broadly as dealing with data and chance) opens up.

Democratization is driven in part by the *quantization of society*. Employment increasingly requires analytical, quantitative, and computing skills, and these requirements put pressure on educational systems. Note carefully that it is not at all clear that *statistical* skills in the traditional sense are required. Few people will need to interpret ANOVAs, fewer will need to carry them out, and still fewer will need to understand the details behind the ANOVA software. This is the counterpoint to the larger place for statistics that is a consequence of democratization. Our teaching must therefore avoid the “professional's fallacy” of imagining that our first courses are a step in the training of statisticians. We should ask whether traditional introductions to statistics for general students are too narrow.

The quantization of society is in its turn driven by the implacable advance of *technology*. Changes in computing, communications, and multimedia come so rapidly that comments in a printed journal are out of date before publication. Technology changes how we teach as well as creating demands for teaching new content.

My thesis is that the most effective learning takes place when *content* (what we want students to learn), *pedagogy* (what we do to help them learn), and *technology* reinforce each other in a balanced manner. Specialists in each of these three areas tend to underestimate the importance of the others—witness the hostility of many content experts to advice from those who do research on teaching and learning. I am of course a content expert writing for other content experts. *Caveat lector*.

### 3 The New Pedagogy

How can we best help students learn? Figure 1 summarizes the reform diagnosis and prescription. Both the diagnosis and the prescription are based on research on teaching and learning. Research that bears directly on the teaching and learning of statistics and probability is summarized in Garfield (1995), Garfield & Ahlgren (1988), Kapadia & Borovcnik (1991) and Shaughnessy (1992). The central idea of the new pedagogy is the abandonment of an “information transfer” model in favor of a “constructivist” view of learning: Students are not empty vessels to be filled with knowledge

- **Goals:** Higher-order thinking, problem solving, flexible skills applicable to unfamiliar settings.
- **The old model:** Students learn by absorbing information; a good teacher transfers information clearly and at the right rate.
- **The new model:** Students learn through their own activities; a good teacher encourages and guides their learning.
- **What helps learning:** Group work in and out of class; explaining and communicating; frequent rapid feedback; work on problem formulation and open-ended problems.

Figure 1. *The Reform of Pedagogy*

poured in by teachers; they inevitably construct their own knowledge by combining their present experiences with their existing conceptions.

In practice, the new pedagogy asks us to change what students do from listening and reading to active participation. We may replace or supplement traditional expository texts by new texts that structure student activities, e.g., Pearl & Stasny (1992), Rossman (1996), Scheaffer *et al.* (1996), Spurrier *et al.* (1995), Tanner (1990). The abundance of these recent books suggests the changing nature of introductory instruction in statistics. We may retain an expository text and continue to do some presentation in the classroom, while moving in the direction of more interaction and more student activity. As one of the best expositions of a moderate reforming position (National Research Council, 1991) puts it:

*What is needed is a variety of activities, including discussion among pupils, practical work, practice of important techniques, problem solving, application to everyday situations, investigational work, and exposition by the teacher.*

“Variety” is the key word in this summary. My own classroom style is now more varied. I ask more questions—if students should see the next step, I ask them rather than telling them. Students bring something to class every day—an attempt at a problem, output from a template computer program, an item of data about themselves. Students break occasionally into small groups to discuss an example or attempt a problem. I give shorter but more frequent examinations. I insist that problem solutions state a conclusion in the context from which the data come—a number, a graph, or “Reject  $H_0$ ” are not adequate solutions. And so on.

**Opinions.** In the large, the reformers are right. My anecdotal experience conforms to the systematic studies: although we may “cover” somewhat less material when we increase interaction in our classrooms, students appear to emerge with a greater store of usable knowledge. This overall conclusion should not be forgotten amidst the qualifications to which I now turn.

Reformers often slight the genuine usefulness of “telling”, both by lectures and by texts. Students profit from a systematic overview of an academic subject, and they are unlikely to “construct” or “discover” the big picture for themselves. Moreover, learning how to learn is one goal of education. We want our students to learn to take a lecture home and interact with it in an exploratory and constructive fashion without explicit guidance—in effect, to learn to build their own program of active learning on an efficient but passive transfer of information. The reformers’ antipathy to lectures reflects in part the origins of the reform movement in studies of school mathematics, and requires some moderation at the university level.

It is nonetheless more important for teachers to remember that *we overvalue lectures*. We overvalue them in part because they worked for us when we were students. What worked for us, however, is not necessarily effective for our students. We are unusual. We are the survivors, the fittest by quite esoteric standards of fitness. The reformers are right: most of us should lecture less most of the time.

This is a fundamental change in the nature of much university teaching. It moves our teaching of beginners in the direction of the collaborative mentor-apprentice model that we have long preferred for post-graduate instruction.

A second qualification concerns the harm that can come from taking the constructivist position to extremes. In the world of mathematics education research, the realization that active learning is essential is sometimes called “naive constructivism”. It is academically more respectable to espouse “radical constructivism” or “social constructivism”. The careful language of several leaders in the field suggests the direction. Here is Ernst von Glasersfeld (1990) on the central tenet of radical constructivism:

*The function of cognition is adaptive, in the biological sense of the term, tending towards fit or viability, and cognition serves the subject's organization of the experiential world, not the discovery of an objective ontological reality.*

Paul Cobb *et al.* (1992) express the social constructivist view:

*Mathematical truth is accounted for in terms of taken-as-shared mathematical interpretations, meanings, and practices institutionalized by wider society. . . . We do not dwell on the traditional question of whether or not something is true in an ahistorical, culture-free sense but instead treat mathematical interpretations or solutions that are considered to be true by members of a community as being practically true in particular situations.*

Both viewpoints attack the relevance (and sometimes the existence) of an objective “reality” that constrains the experiences that learners organize. Mathematicians usually ask rhetorically in which cultures  $2 + 2$  (in whatever notation) does not equal 4, and the debate over the nature of mathematics is relaunched. The claim that knowledge is socially (or even individually) constructed seems more applicable to statistics than to mathematics. My real concern is that extreme versions—some might call them parodies, but I have met educators who expound them—of these positions have paralyzing effects on attempts to help students learn. Taken to its limits, radical constructivism suggests that because we all construct our own knowledge, teaching is essentially impossible. No one’s preconceptions can be said to be “wrong”, even if the preconception in question is that the logarithm acts like a linear function. Social constructivism suggests that because knowledge is socially constructed, with no necessary correspondence to any truth “out there”, teaching in any active sense is akin to indoctrination.

That way lies madness. Teachers of the mathematical sciences would, I think, be wise not to start in the direction of radical or social constructivism. These academic positions give us no practical aid. We do better to remain naive. Even in a thoroughly reformed classroom, the teacher has a special place earned by training, maturity, and, yes, knowledge. Guide, facilitator, moderator, provoker rather than lecturer she may now be, but she still wants her students to move in specific directions, toward specific mathematical and statistical competence.

#### 4 Changing Content in Statistics

What we want beginners to learn about statistics has changed dramatically in the past generation. Older introductions to the discipline were dominated by probability-based inference. Students learned to carry out specific procedures to gain answers to well-posed questions under somewhat restrictive assumptions. A modern first course will (in my opinion) offer a more balanced introduction to data analysis, data production, and inference. Exploring data, designing data production, using diagnostic tools to ask whether a proposed method of inference is appropriate have a “back and forth” flavor quite unlike the “straight ahead” nature of traditional statistical calculations. Why have we changed?

1. **Emphasize the elements of statistical thinking:**
  - (a) the need for data,
  - (b) the importance of data production,
  - (c) the omnipresence of variability,
  - (d) the measuring and modeling of variability.
2. **Incorporate more data and concepts, fewer recipes and derivations. Wherever possible, automate computations and graphics.** An introductory course should:
  - (a) rely heavily on *real* (not merely realistic) data,
  - (b) emphasize *statistical* concepts, e.g., causation vs. association, experimental vs. observational and longitudinal vs. cross-sectional studies,
  - (c) rely on computers rather than computational recipes,
  - (d) treat formal derivations as secondary in importance.
3. **Foster active learning**, through the following alternatives to lecturing:
  - (a) group problem solving and discussion,
  - (b) laboratory exercises,
  - (c) demonstrations based on class-generated data
  - (d) written and oral presentations,
  - (e) projects, either group or individual.

Figure 2. Recommendations of the ASA/MAA Joint Curriculum Committee

The chain of influence begins once again with technology. The continuing revolution in computing first changed the practice of statistics, then changed our tastes for what constitutes interesting research in statistics. More slowly, the combined weight of technology, professional practice, and research tastes have influenced introductory instruction (at least when the instructor is a statistician).

An account of the changes now underway, in an informal style with amusing comments from many individuals, appears in Cobb (1992). This is the report of the joint curriculum committee of the American Statistical Association (ASA) and the Mathematical Association of America (MAA). The Board of Directors of ASA has approved the brief form of the committee's recommendations that appears in Figure 2. The third heading in that display endorses active learning. I will use the first two headings to structure my brief comments on content.

*More data and concepts:* First courses should offer experience working with data from real problem settings. They should concentrate on the things that can't (at least yet) be automated, such as interpretation of graphics, strategies for effective exploration of data, basic diagnostics as preliminaries to inference, and the conceptual meaning of "*P*-value", "confidence", and "statistical significance". The other side of this coin is *fewer recipes and derivations*: Grasp of the reasoning of inference is more important than how many individual procedures we touch on, and derivations should only be done if they convince our students. Formal proofs and algebraic derivations convince us, but recall once more that what worked for us is not necessarily effective for our students.

*Emphasize statistical concepts:* "Statistics has its own substance, its own distinctive concepts and modes of reasoning. These should be the heart of the teaching of statistics to beginners at any level of mathematical sophistication". Those words appear in the opening paragraph of Moore (1992), where I elaborate on that theme. There are few principles I hold more strongly. A student who emerges from a first statistics course without an appreciation of the distinction between observation and experiment and of the importance of randomized comparative experiments, for example, has been cheated. Those specific examples are instructive, for they point to core *statistical* ideas that are not *mathematical* in nature.

*Automate computations and graphics:* After perhaps a short example done by hand for pedagogical purposes, means and standard deviations are buttons on a calculator and scatterplots with regression lines are menu items in software. Automating computations is controversial among mathematicians. It is much less controversial among statisticians, as it reflects the practice of our discipline. Some consequences will appear in the following sections.

**Opinions.** It is already clear that I think the trend toward emphasizing data and concepts is healthy. Increased emphasis on data analysis and data production yields a broader and more broadly relevant introduction to statistics. Here are some further suggestions.

We would do well, in response to the opportunities created by democratization, to further broaden the scope of our introductory instruction. It is, recall, users of statistics and citizens who sit before us, not future professional statisticians. Issues of *data ethics*, for example, are important and are of interest to a broad range of students. Perhaps “informed consent” and “confidentiality” belong alongside “scatterplots” and “simple random samples” in our syllabus. Perhaps even a brief discussion of the nature of *official statistics* and the organization of *national statistical offices* is in order. Here is a major aspect of our profession that is invisible in our teaching. Ethics and official statistics (for the United States) both receive substantial exposition in Moore (1997). This text is unusual in having students in the liberal arts as its target audience; instruction for other groups of students might also consider these and other innovative topics.

If we devote more attention to hands-on data analysis and the design of data production, and perhaps mention data ethics and official statistics, we shall have to reduce the attention paid to some traditional topics. My candidate for the guillotine is formal probability.

Experienced teachers recognize that students find probability difficult. Research on learning confirms our experience. Garfield & Ahlgren (1988) document the fragility of probability concepts even among students who can work formal probability problems. They conclude that “teaching a conceptual grasp of probability still appears to be a very difficult task, fraught with ambiguity and illusion”. They recommend exploring “how useful ideas of statistical inference can be taught independently of technically correct probability”.

In my opinion, only an informal grasp of probability is needed to follow the reasoning of standard statistical inference. That reasoning is based on consistently asking the question, “What would happen if we did this many times?” The sampling distribution of a statistic answers that question in full, and leads to the more specific answers summarized by confidence levels and  $P$ -values. Sampling distributions can be demonstrated by simulation and studied by the tools of data analysis. The distribution of a variable, not probability in any formal sense, is the conceptual basis of inference. We use the language of probability to describe “What would happen if we did this many times?” but the formal machinery is a barrier rather than a help at this level. After all, the main use of  $P(A^c) = 1 - P(A)$  in a first statistics course is to note that if area 0.95 lies to the left of a point on a  $t$  curve, then area 0.05 lies to the right of that point.

An informal introduction to probability should include experience with chance behavior, usually starting with physical chance devices and moving to computer simulations. Important facts such as the law of large numbers and the central limit theorem can be demonstrated and made convincing in this manner. Many instructors will want to clarify the distinction between the long-run regularity that probability describes and the short-run irregularity of random phenomena. Psychologists (e.g., Tversky & Kahneman 1971) have noted that most people accept an incorrect “law of small numbers” that asserts that even short runs of random phenomena should be regular; if they are not, we look for an explanation other than chance behavior. Our intuitive judgments of probability are in general (Tversky & Kahneman, 1983) “not likely to be coherent, that is to satisfy the constraints of probability theory”. Informal probability can clarify issues such as popular assessment of risks—see e.g., Slovic, Fischhoff & Lichtenstein (1982) and Zeckhauser & Viscusi (1990) for comments on this fascinating mixture of probability and psychology. Informal probability lays a foundation of experience and

concepts for more formal study, as well as providing the basis for understanding the reasoning of statistical inference.

Mathematical probability is of course a noble and useful subject. It is essential for mathematical modeling and for the mathematical theory that underlies some parts of statistics. Attempting to present a substantial introduction to probability in a data-oriented statistics course, however, is in my opinion unwise. Formal probability does not help these students master the ideas of inference (at least not as much as we teachers imagine), and it depletes reserves of mental energy that might better be applied to essentially statistical ideas.

## 5 Content, Pedagogy, and Technology in Synergy

Discussions of one of the triad content–pedagogy–technology are often partial. The argument for changing introductory instruction in statistics is strengthened by reinforcing relationships among the three domains. Figure 3 summarizes some of these synergies.

The double arrows suggest that reinforcement flows in both directions. Exploratory analysis of data, for example, is an important topic that we want to present in a first course. That requires that students actually explore data. Education researchers tell us that active hands-on work helps learning, so that active data analysis is an effective entry into statistics independent of its importance as separate content. Hence the entry

Data analysis  $\iff$  Hands-on work

in Figure 3. Here are brief remarks on the other synergies in Figure 3.

*Content  $\iff$  Pedagogy*: Asking students to work cooperatively and to communicate their findings orally and in writing aids learning. It is an aspect of good pedagogy. More emphasis on communication and cooperation helps prepare our students for employment, and so should be part of our broadened content. To cite just one example, I recently spent a week at Motorola learning how this successful company does TQM. Motorola wanted the assembled academics to hear this message: “All of our work is done cooperatively in teams. Why do you persist in sending us students whose only experience is individual and competitive?”

- **Content  $\iff$  Pedagogy**

- Data analysis  $\iff$  Hands-on work

- Statistics in practice  $\iff$  Communicate, cooperate

- More concepts  $\iff$  Less proof

- **Pedagogy  $\iff$  Technology**

- Visualization  $\iff$  Automate graphics

- Problem-solving  $\iff$  Automate calculations

- Active learning  $\iff$  Multimedia

- **Technology  $\iff$  Content**

- Computing  $\iff$  Data analysis, diagnostics, bootstrap, . . .

- Automation  $\iff$  More concepts

- Simulation  $\iff$  Less proof

**Figure 3.** Synergy in Statistical Education

The final entry under this head reminds us that on the content side we have always wanted to emphasize conceptual understanding; good pedagogy urges us to choose conceptual explanations over proofs that are not convincing to most of our students.

*Pedagogy*  $\iff$  *Technology*: Carefully structured visualization (“multiple representations” in the language of education research) aids learning. So does work on open-ended problems that require multiple attempts and combination of several tools rather than a single path to the unique right answer (Garfield, 1995, p. 30). Choosing a model for somewhat complex data—a dialog among the data, candidate models, and various diagnostic tools—is a typical example of an open-ended statistical problem. Consistent emphasis on visualization and problem-solving are hardly possible if graphics and calculations must be done by hand. In the other direction, what is technically easy is almost always done, so that lots of graphs and multiple analyses are a consequence of the technology. Good teachers now pay more attention to effective strategies for e.g. regression diagnostics, and less attention to the details of how to calculate regression fits. Good strategies will (we hope) head off the obvious pitfalls of essentially free and immediate multiple analyses.

Multimedia educational systems, the new version of computer-assisted learning, may at last enable genuinely active learning (the core of the new pedagogy) on a technological platform. Multimedia systems deserve separate comment, and receive it in Section 8.

*Technology*  $\iff$  *Content*: New content in statistics reflects the computing-intensive nature of statistical practice. Exploratory data analysis is characterized (Velleman & Hoaglin, 1992) by an “iterative process of describing patterns, subtracting them, and searching anew for pattern in the residuals [that] continues until the data analyst decides to stop”. This iterative process is tedious indeed if the details are not automated. We may wish to expose even beginners to newer topics such as regression diagnostics and the bootstrap. The bootstrap is a nice example of synergy between technology and content: it is a conceptually simple extension of the core idea of a sampling distribution that is widely useful—and impossible without fast and cheap computing.

Automation of routine operations both allows and demands that teaching lay more stress on larger concepts and strategies. Automation forces teachers and learners to pay more attention to what isn’t automated, if only at the level of deciding which item in the “plot” menu is appropriate for the present problem. This allows the conceptual emphasis that teachers have always preferred, but that was impeded by students’ struggles to implement routine recipes.

Simulation offers an alternative to proofs and algebraic derivations as a way of convincing students of the truth of important facts. The central limit theorem, always a fact we could not prove to beginners, is both more comprehensible and more convincing when we actually see it at work via simulation and graphics.

## 6 Old Technologies: Video and Computing

Television has changed the world, yet has had only peripheral impact on education. The reason is not (or not only) that educators are immovably attached to neolithic traditions. Video is in fact not effective as the primary medium of instruction. To see why, let us look at this old technology from the point of view of pedagogy and content. More details, and references, appear in Moore (1993).

Video’s most obvious strength as a teaching medium is that it compresses time and space to focus on essentials. We can view a lengthy process briefly, aided by microscopy and telescoping as needed. Careful editing and polished technique often make video more convincing than being there yourself. There is another, more subtle but equally important, aspect to video: it operates subliminally as well as rationally, changing the attitudes of viewers at a subconscious level.

The weaknesses of video parallel its strengths. Video *shows* rather than *tells*. It is a very poor medium for exposition. This is particularly true in subjects where much important content is not highly visual. Video exposition must choose between bad television (a talking head) and skewing

the content in a search for visually appealing topics. It is no accident that “science television” concentrates on furry animals. More seriously, video leaves its viewers passive. The talking head cannot interact with learners. Even excellent video presentations have limited cognitive impact—students have acquired a certain visual sophistication from thousands of hours of television; they have learned to remain disengaged from a world they assume to be the surreal creation of a clever producer.

The conclusion seems clear. Video has value as a means of changing attitudes by taking students out of the classroom to see real people using statistics in real settings. Video can help motivate students, and can on occasion show a phenomenon otherwise hidden from our view. These are supplementary roles. Video is not suited as even a primary medium for exposition, let alone for the more active components of the varied style recommended by current pedagogical thinking.

It may seem odd to class computing with video as an old technology. We have, however, asked our students to do computing as part of their statistics training for at least a generation. Our viewpoint as statisticians has been shaped by the *content* we want to present: computing allows realistic problems, serious statistical methods, and emulation of actual statistical practice.

Adding *pedagogy* to the mix of content and technology broadens our thinking about computing. Teachers should consider computing as a tool for *learning* statistics, not simply for *doing* statistics. Because graphics and manipulations aid learning, we should encourage students to use software to explore, visualize, and interact with data and simulations, not simply to automate calculations.

There is another aspect to computing as a teaching tool: it improves students’ ability to complete problems. A student who understands conceptually what must be done to solve a problem is much more likely to be able to implement that understanding when software is at hand. Although computing can be frustrating, that frustration is in my experience less annoying than the frustration of being unable to complete a problem due to some minor error in a long but routine calculation. Good software reduces the students’ cognitive load, replacing complex algorithmic procedures by simpler commands, thus allowing learners to focus on higher-level understanding. These pedagogical advantages strengthen the case for using computers (or in some settings, advanced calculators) in beginning instruction.

Software designed for doing statistics is not necessarily well structured for learning statistics. Rolf Biehler (1993, 1995) has interesting things to say about the technical features—both existing and envisioned—that make software good for learning as well as for doing statistics. Statisticians interested in software design for teaching will profit from the second paper in particular.

## 7 New Technology: Graphing Calculators

One of the most striking developments in statistical education in the past generation has been the arrival, in at least most English-speaking countries, of working with data as a standard strand in school mathematics curricula. The *Standards* of the U.S. National Council of Teachers of Mathematics (NCTM, 1989) specify “increased attention” to probability and statistics at all grade levels. In the United Kingdom, “data handling” is to make up 20% of the mathematics curriculum for ages 11–16.

This development is due in part to the movement toward the democratization of mathematics. In the United States, at least, school mathematics has been influenced by a concern arising from the quality management movement that students who do not enter university should have stronger quantitative preparation for the work force. The presence of an early and excellent model in the ASA/NCTM *Quantitative Literacy* series no doubt contributed. (The books in the QL series are Gnanadesikan, Schaeffer & Swift (1986); Landwehr & Watkins (1986); Landwehr, Watkins & Swift (1987), and Newman, Obremski & Schaeffer (1986).)

In my opinion, however, the strongest influence in the movement of data analysis into school mathematics is synergy with the new active learning pedagogy and, perhaps less obviously, synergy

- **Whole-number counting**  $\iff$  **Bar graph**  
 “How often do you wear a hat?”  
 Bars of stick figures count responses: Often, Sometimes, Seldom, Never.
  
- **Number line**  $\iff$  **Dot plot**  
 “What is the date on your coin?”  
 Interesting left-skewed distribution.
  
- **Place notation**  $\iff$  **Stem-and-leaf plot**  
 “How many pages does your favorite book have?”
  
- **Betweenness and grouping**  $\iff$  **Histogram**  
 “How long does it take you to get to school?”

Figure 4. *Primary School Synergy*

with the core content of primary school mathematics. Figure 4 shows how closely a natural sequence of basic data graphics parallels a natural sequence of number concepts and skills. The examples cited are from a recent collection of classroom activities (University of North Carolina, 1997ab), but similar examples appear in many other recent texts and supplements. Once established in primary school mathematics, data analysis remains present in the later school years for both pedagogical and utilitarian reasons.

Graphing calculators have become the standard technology for advanced secondary school mathematics in the United States. They are allowed (and so in effect required) for the popular Advanced Placement examinations that offer university credit in calculus and (starting in 1997) in statistics. The current generation of graphing calculators, such as the Texas Instruments TI-83, offer most of the inference procedures taught in a first statistics course, in addition to the statistical graphics and simulation functions of previous models. Asked to carry out a  $t$  test from keyed-in data, the TI-83 can draw a  $t$  density curve on its screen and shade the area corresponding to the  $P$ -value, while printing the  $t$  statistic and its  $P$ -value. An alternative screen provides more numerical output, including  $\bar{x}$  and  $s$ .

From the viewpoint of *pedagogy*, graphing calculators have striking advantages. Portability combined with moderate cost allows students to carry their technology with them. Calculators can be used in places where computers are unavailable, for example in a project in rural schools in South Africa (Laridon, 1996). Even when computers are available, calculators in the hands of all the students in a classroom encourage active participation. Secondary school teachers report that the “sense of ownership” felt by students extends from the calculator to the mathematics and statistics being learned.

From the *content* side, however, calculators remain more suitable for mathematics than for statistics. The difficulty of entering more than small amounts of data and the limits imposed on statistical graphics by small screens are the most important weaknesses. Against these must be set, in addition to the pedagogical advantages just mentioned, good simulation capabilities and continued rapid technical improvement that may yet partially overcome the data entry and graphics limits. The likely path is a closer link to computers, so that a single classroom computer can function as “server” for many calculators. The calculators would download data (e.g., via infra-red signals) from the computer. The reverse link would process calculator output for display at higher resolution on the computer’s monitor.

Graphing calculators are an example of tension between pedagogy and content. They are now

adequate computing-and-graphing tools for a first statistics course, but not the preferred tool. The suitability of advanced calculators for implementing the reform emphasis on constant student involvement give them an advantage over computers in most classroom settings. If substantial numbers of students arrive in our courses already familiar with graphing calculators from secondary school, the argument may tip towards adopting this technology.

## 8 New Technology: Multimedia

“Multimedia” (Figure 5) is one of the most prominent technological buzzwords, made still more prominent by a vision of the ability to transmit full multimedia across long distances electronically in real time. Like all buzzwords, “multimedia” is used loosely. Alleged multimedia educational products differ greatly in the depth and accuracy of their *content* and in the effectiveness of their *pedagogy*. We should surely examine these critically, and not be overly impressed by mere technology. Velleman & Moore (1996) discuss some of the pedagogical challenges in the design of multimedia systems that present a structured, hierarchical subject such as introductory statistics. There are also pedagogical challenges for the instructor who uses multimedia: given that multimedia can to some extent stand on its own, how shall we integrate the role of this much more powerful software with those of a text and a human teacher?

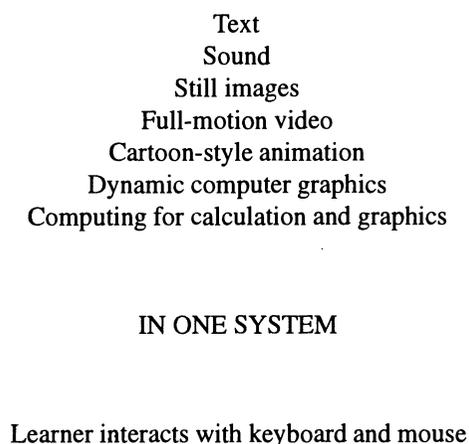


Figure 5. *Multimedia is . . .*

Our concern for content and pedagogy should, however, not lead us to tolerate weak technology. There are as yet few good multimedia instructional systems to set standards in subjects similar to statistics. (There are attractive supplements that do not pretend to offer consecutive instruction, and also attractive systems to aid learning of less hierarchical subjects such as art and history. The latter often use a loose “explore this world” structure that is less appropriate for statistics.) Early “multimedia” products in the mathematical sciences have tended to be text-based, enhancing the text with illustrations and student activities. In my opinion, alleged multimedia systems that are driven by text are weak uses of a promising technology, and we should reject them. Large blocks of text belong in print, where text is easier to read, easier to navigate in, and more portable.

That said, multimedia has clear advantages from the viewpoint of pedagogy. It can utilize the strengths and avoid the weaknesses of each medium. My comments on video, for example, point to a restricted but useful role for video in a multimedia environment. Most important, effective multimedia instruction can (and should) be highly interactive. The learner controls the pace and

launches each succeeding activity. She can manipulate video and animated computer graphics, so that teaching demonstrations are turned over to her for more exploration. Statistical software is always available for use, in an environment that can offer instruction on the software as each new capability is required. Embedded exercises with immediate feedback and unlimited ability to review the material just presented facilitate a “mastery learning” style in which the learner is satisfied that she has mastered each concept before going on.

As far as content is concerned, multimedia should do well for the types of content that I believe our introductory courses should stress. Hands-on work with data; conceptual understanding gained through demonstrations and simulations that the student can control, modify, and repeat; routine use of software to automate graphics and calculations—all this is natural in a multimedia setting.

The weaknesses of multimedia technology center on the social aspects of learning. “Communication and cooperation” is also on the list of reform recommendations for student activities. Moreover, students learn from each other and from a teacher who is sensitive to non-verbal signs and non-cognitive barriers to learning. Well-designed multimedia software will offer home pages, discussion groups, email to the teacher, and so on. I do not believe these are sufficient replacements for a personal presence. The analogy with distance learning is revealing: distance learning has proved effective for mature and motivated students; it works much less well for the relatively immature and less motivated students we often see in first courses.

There will remain an essential role for human teachers even when technology carries the primary teaching burden. I see no substitute for the *motivation and encouragement* that a teacher can provide. Nor do I believe that *assessment* can be taken out of human hands, at least if we genuinely want our students to learn those “higher order” skills applicable to unfamiliar problems. Assessment is a large and controversial topic that I have avoided here. Do note, however, that the first principle of assessment is what one overview (National Research Council, 1993) calls the *content principle*: assess what you value most. After all, students will concentrate on the content they expect us to assess. If we want “flexible problem-solving skills”, we must assess those skills. This will be difficult indeed to automate. Garfield (1994) is a good place for statisticians to start thinking about assessment of student learning.

We will still, in the multimedia future, interact regularly with our students. In fact, spared the burden of tasks that are now automated, we may meet students in discussion groups that are smaller than our present classes. Our assessment of student learning may well be more elaborate, and may hold students to the higher standard of completed work that good technology permits. We ought, in short, to welcome and explore the use of new technologies for which “multimedia” is a shorthand. We ought also to keep our understanding of valuable content and effective pedagogy clearly in mind as we view these and other developments with a critical eye.

## 9 Conclusions, and a Few Qualifications

Content and pedagogy—our understanding of what students should learn and of effective ways to help them learn—should drive our instruction. Technology should serve content and pedagogy. Yet technology has changed content and allows new forms of effective pedagogy. “Synergy” is thus my one-word summation.

The most effective teachers will have a substantial knowledge of pedagogy and technology, as well as comprehensive knowledge about and experience applying the content they present. Universities have always emphasized subject-matter knowledge, and have often assessed their staff’s contributions to knowledge by high standards. They have, however, generally assumed that a subject-matter Ph.D. equips faculty to teach well. That assumption, and the systems by which universities evaluate their staff, are slowly yielding to the same pressures that form the context of the movement to reform teaching. We would do well to broaden our own scholarly knowledge as well as the content of our

introductory courses. The references offer some starting points.

Now for the qualifications. My opinions are shaped by some strong convictions. It would be easy to carry these convictions to unfortunate extremes, and I hope that readers will not extrapolate too far.

I feel strongly, for example, that statistics is not a subfield of mathematics, and that in consequence, beginning instruction that is primarily mathematical, or even structured according to an underlying mathematical theory, is misguided. Such instruction will inevitably understate the role of exploratory analysis and the design of data production, and may ignore essential distinctions, such as that between observational and experimental data, that are not captured by the theory of inference. It is nonetheless true that statistics makes heavy and essential use of mathematics, that advanced training in statistics requires considerable exposure to mathematics, and that elaborate mathematical theories underlie some parts of statistics. Bullock (1994) is wrong in claiming that “Many statisticians now insist that their subject is something quite apart from mathematics, so that statistics courses do not require any preparation in mathematics”.

Statisticians should eschew the contempt for mathematics that finds classic expression in the Royal Statistical Society’s discussion of Kiefer’s (1959) exposition of the principles of optimal experimental design. The discussants argue in effect that because not all aspects of statistical practice can be made precise, there is no virtue in any attempt to achieve precision. They express their views, moreover, with that articulate rudeness in which the British upper classes once specialized. This makes amusing reading decades later, but it reflects badly on both the intellectual breadth and the common courtesy of the discussants.

I also feel strongly that heavy use of computing technology is essential for realistic learning of practical statistics, and that automating anything that is “just a rule” is good pedagogy as well. I contend, for example, that

$$b = \frac{\sum xy - \frac{1}{n}(\sum x)(\sum y)}{\sum x^2 - \frac{1}{n}(\sum x)^2}$$

is just a rule, that it communicates no conceptual understanding to our students. It ought to be a button or a menu item. On the other hand,

$$b = r \frac{s_y}{s_x}$$

tells the algebraically literate a great deal about the regression slope  $b$ . This is a formula our students should know.

Yet I recognize that computing can frustrate students. “Real statisticians use SAS” (to quote a colleague) is not a reason to use SAS in a first course. The choice of software—and even the decision to use graphing calculators or spreadsheets rather than statistical software—depends heavily on our judgment of accessibility to students. We are teaching our subject, not the tool. I also recognize that not all teachers agree on what is “just a rule”, and therefore should be automated. Some may even feel that sums of squares speak to the souls of students. Steven Krantz, for example, in a generally sensible book on *How to Teach Mathematics* (Krantz, 1993) takes a more negative view of automating student work than I do. These are matters for informed judgment in the light of local circumstances.

Informed judgment in the light of local circumstances is, in fact, needed in every aspect of the teaching of statistics to beginners. I have tried to be clear about the *directions* in which I believe we should move. An attempt to be precise about the distance we should move in these directions would founder on the fact that my circumstances and my students differ from yours. It would also litter this paper with “on the other hand”, thus obscuring the strength of my convictions. I hope that most

readers will agree on our direction, and that they will use both their judgment and their ingenuity in arranging the journey. Let a thousand flowers bloom.

## References

- Biehler, R. (1993). Software tools and mathematics education: The case of statistics. In *Learning from Computers: Mathematics Education and Technology*, Eds. C. Keitel & K. Ruthven, pp. 68–100. Berlin: Springer.
- Biehler, R. (1995). Towards requirements for more adequate software tools that support both learning and doing statistics. Occasional paper 157, Institut für Didaktik der Mathematik, Universität Bielefeld.
- Bullock, J.O. (1994). Literacy in the language of mathematics. *American Mathematical Monthly* **101**, 735–743.
- Cobb, G. (1992). Teaching statistics. In *Heeding the Call for Change: Suggestions for Curricular Action*, Ed. L. A. Steen, pp. 3–43. Washington, D.C.: Mathematical Association of America.
- Cobb, P., Yackel, E. & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education* **23**, 2–33.
- Garfield, J. (1994). Beyond testing and grading: Using assessment to improve student learning. *Journal of Statistics Education* (electronic journal) **2**, no. 1.
- Garfield, J. (1995). How students learn statistics. *International Statistical Review* **63**, 25–34.
- Garfield, J. & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education* **19**, 44–63.
- von Glasersfeld, E. (1990). An exposition of constructivism: why some like it radical. In *Constructivist Views on the Teaching and Learning of Mathematics*, Eds. R. B. Davis, C. A. Maher & N. Noddings. Reston, Va.: National Council of Teachers of Mathematics.
- Gnanadesikan, M., Schaeffer, R. & Swift, J. (1986). *The Art and Techniques of Simulation*. Palo Alto, Calif.: Dale Seymour.
- Kapadia, R. & Borovcnik, M. (Eds.) (1991). *Chance Encounters: Probability in Education*. Dordrecht: Kluwer.
- Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society* **B21**, 272–319.
- Krantz, S.G. (1993). *How to Teach Mathematics: A Personal Perspective*. Providence, R.I.: American Mathematical Society.
- Landwehr, J. & Watkins, A. (1986). *Exploring Data*. Palo Alto, Calif.: Dale Seymour.
- Landwehr, J., Watkins, A. & Swift, J. (1987). *Exploring Surveys and Information From Samples*. Palo Alto, Calif.: Dale Seymour.
- Laridon, P. (1996). Graphics calculators in South Africa. Paper presented at the Eighth International Conference on Mathematics Education, Seville, Spain.
- Moore, D.S. (1992). Teaching statistics as a respectable subject. In *Statistics For the Twenty-First Century*, Eds. F.S. Gordon & S. P. Gordon, pp. 14–25. Washington, D.C.: Mathematical Association of America.
- Moore, D.S. (1993). The place of video in new styles of teaching and learning statistics. *American Statistician* **47**, 172–176.
- Moore, D.S. (1997). *Statistics: Concepts and Controversies*, 4th ed. New York: Freeman.
- National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards For School Mathematics*. Reston, Va.: NCTM.
- National Research Council (1990). *Reshaping School Mathematics: A Philosophy and Framework for Curriculum*. Washington, D.C.: National Academy Press.
- National Research Council (1991). *Moving Beyond Myths: Revitalizing Undergraduate Mathematics*. Washington, D.C.: National Academy Press.
- National Research Council (1993). *Measuring What Counts: A Conceptual Guide for Mathematical Assessment*. Washington, D.C.: National Academy Press.
- Newman, C., Obremski, T. & Schaeffer, R. (1986). *Exploring Probability*. Palo Alto, Calif.: Dale Seymour.
- Pearl, D.K. & Stasny, E.A. (1992). *Experiments in Statistical Concepts*. Dubuque, Iowa: Kendall/Hunt.
- Rossman, A.J. (1996). *Workshop Statistics: Discovery With Data*. New York: Springer.
- Schaeffer R.L., Gnanadesikan, M., Watkins, A. & Witmer, J. (1996). *Activity-Based Statistics: Student Guide*. New York: Springer.
- Shaughnessy, J.M. (1992). Research in probability and statistics: reflections and directions. In *Handbook on Research on Mathematics Teaching and Learning*, Ed. D.A. Grouws, pp. 465–494. New York: Macmillan.
- Steen, L.A. (Ed.) (1988). *Calculus for A New Century*. Washington, D.C.: Mathematical Association of America.
- Slovic, P., Fischhoff, B. & Lichtenstein, S. (1982). Facts vs. fears: Understanding perceived risk. In *Judgment under Uncertainty: Heuristics and Biases*, Eds. D. Kahneman, P. Slovic & A. Tversky, pp. 463–489. New York: Cambridge University Press.
- Spurrier, J.D., Edwards, D.G. & Thombs, L.A. (1995). *Elementary Statistics Laboratory Manual*. Belmont, Mass.: Wadsworth.
- Tanner, M.A. (1990). *Investigations for a Course in Statistics*. New York: Macmillan.
- Tversky, A. & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin* **76**, 105–110.
- Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* **90**, 293–315.
- University of North Carolina (1997a). *Teach-Stat Activities: Statistics Investigations for Grades 1–3*. Palo Alto, Calif.: Dale Seymour.
- University of North Carolina (1997b). *Teach-Stat Activities: Statistics Investigations for Grades 3–6*. Palo Alto, Calif.: Dale Seymour.
- Velleman, P.F. & Hoaglin, D.C. (1992). Data analysis. In *Perspectives on Contemporary Statistics*, Eds. D.C. Hoaglin & D.S.

- Moore, pp. 19–39. Washington, D.C.: Mathematical Association of America.
- Velleman, P.F. & Moore, D.S. (1996). Multimedia for teaching statistics: promises and pitfalls. *American Statistician* **50**, 217–225.
- Vere-Jones, D. (1995). The coming of age of statistical education. *International Statistical Review* **63**, 3–23.
- Zeckhauser, R.J. & Viscusi, W.K. (1990). Risk within reason. *Science* **248**, 559–564.

## Résumé

L'éducation statistique se déroule maintenant dans un nouveau contexte social. Elle se trouve influencée par un mouvement de réforme de l'enseignement général des sciences mathématiques. De plus, la nature changeante de notre discipline exige un contenu révisé au niveau des cours d'introduction et la technologie influence fortement aussi bien ce qui est enseigné que la façon dont on l'enseigne. Les arguments en faveur de changements substantiels dans l'enseignement de la statistique reposent sur de fortes synergies entre le contenu, la pédagogie et la technologie. Les statisticiens qui enseignent aux débutants devraient se familiariser davantage avec les recherches sur l'enseignement et l'apprentissage et avec les changements dans les technologies d'éducation. L'esprit de l'introduction à la statistique contemporaine devrait se détacher de l'importance traditionnellement accordée au cours magistral et à la probabilité et l'inférence.

# Discussion

## Joan Garfield

*University of Minnesota, Department of Educational Psychology University of Minnesota, 332 Burton Hall, 178 Pillsbury Drive S.E., Minneapolis, MN 55455, USA*

### 1 Introduction

I enjoyed reading David Moore's article which provides an eloquent and logical rationale for the current reform effort in statistics education. As I am in agreement with most of the points made in his paper, my comments will focus on the current extent of this movement and identify obstacles that challenge reform efforts. In doing so, I will question a few of Moore's views on teaching and learning statistics. I conclude with some suggestions to help accelerate the progress of this reform.

### 2 Successful Outcomes of the Reform Movement

The article begins with a description of forces contributing to the current reform movement, including changes in the discipline, advances in technology, and new theories of how students learn. Some of the recommended changes in statistics education may be summarized as: "less theory, more data"; "less lecture, more active learning"; and "less calculation, more technology". The statistics education community has already accomplished quite a lot in working toward these changes. We seem to be most successful in the following areas:

*Instructional materials:* Many improved instructional materials are now available that provide exercises and examples involving real data and reflect changes in the practice of statistics. For example, innovative textbooks (e.g., Moore, 1995; Rossman, 1996), collections of activities (e.g., Scheaffer et al., 1996; Spurrier, et al., 1995), educational software (e.g., Wackerly, 1996), and internet resources (e.g., Snell, 1996) have been developed and are being disseminated.

*Technology.* The amount and quality of technology now available in most colleges and universities continues to improve, as students have increased access to computers, graphing calculators, Internet resources, and multimedia. Excellent software programs are available for exploring data, illustrating statistical concepts, and even tutoring students.

*Literature and guidelines on teaching statistics.* There is an impressive body of literature focused on statistics education. Resources include the electronic *Journal of Statistics Education (JSE)*, the British journal, *Teaching Statistics*, electronic discussion groups and newsletters, and proceedings volumes from recent international and national conferences on statistics education. These publications discuss philosophical issues related to teaching statistics and provide recommendations for using active learning strategies, student projects, and technology resources. Student assessment and psychological issues related to teaching and learning statistics are also addressed.

### 3 What Hasn't Changed

Despite the quality of the available materials and resources and the enthusiasm of education reformers, most statistics courses taught in institutions of higher education have changed very little. What hasn't seemed to change in many of these courses are:

*Teaching methods.* In many institutions students sit passively in classes while a professor lectures to them using an overhead projector or blackboard to write out formulas and work out problems. As Hogg stated: "Statistics teaching is often stagnant; statistics teachers resist change" (1992, p. 36).

*The stigma associated with "statistics".* Many students express anxiety or fear about taking a course in statistics which is a requirement for many fields of study. Those who complete an introductory course are often dissatisfied and make negative comments about their statistics courses.

*Instructor frustration.* Many instructors of statistics courses remain frustrated as they attempt to teach students who often lack motivation, have poor mathematics skills, do not attend class regularly, and perform poorly on exams. Instructors of higher level courses often feel dissatisfied with the quality of students' learning and their inability to apply skills in other contexts.

### 4 Faculty Beliefs about Teaching

These three areas that have resisted change suggest a need to examine why the educational reform has not had more of an impact on college statistics classes. I speculate that this lack of widespread change is largely due to beliefs held by faculty which are very stubborn and resistant to change, as well to external constraints. The faculty beliefs, or perhaps myths, are described below.

#### **"Good teaching means good lectures."**

Many faculty view teaching as logically laying out material for students to absorb. They view good teaching as being well prepared for lectures, being knowledgeable about their subject area, speaking clearly, using good overheads, and adding a touch of humor to their lectures to keep students engaged. However, as Moore reports, current theories of learning contradict the image of students as empty vessels or blank slates, and suggest that lecturing is not an effective learning method. Cobb (1992) states: "Shorn of all subtlety and led naked out of the protective fold of educational research literature, there comes a sheepish little fact: lectures don't work nearly as well as many of us would like to think" (p. 9).

Despite Moore's explanation of the widely accepted constructivist theory of learning, he seems to overestimate the efficiency and utility of lectures. He remembers that they "worked for us as students" and he views them as "an efficient but passive transfer of information". Lectures are indeed a transfer: of numbers, formulas, definitions, and examples; usually from the instructor's notes or overhead transparencies to students' notebooks (or sometimes, tape recorders). This information is often transferred back again to exam papers, where if it does not appear in the exact manner it was delivered, the student is not awarded full credit. Not surprisingly, many educators have found that students who produce correct answers on exams often are unable to explain their understanding nor apply what they have learned to a problem or real-world context (Garfield, 1995). Therefore, I believe that this "transfer" of information reinforces the myth of "teaching is telling and learning is

remembering”.

New pedagogical methods being advocated by reformers encourage students to develop and express their own understanding of statistics, rather than copy and recite their teachers' views. A wonderful example of this approach can be found in the recently published text, *Workshop Statistics* (Rossman, 1996).

Eliminating or reducing lecturing changes the role of teacher to that of facilitator and guide, rather than “source of knowledge”, a role that makes many professors uncomfortable. A teacher/facilitator spends more time constructing learning activities, monitoring class time, listening to students, and assessing student learning, than in traditional “teaching” activities. Allowing students to work collaboratively in class, solving problems and communicating their understanding better prepares them to apply their learning in future courses and careers (Garfield, 1993).

Many faculty feel that unless they lecture, they will not “cover” material. Moore talks about the need for lectures to provide a “big picture” for students. There doesn't seem to be evidence that this is actually true. Perhaps one way to find out is to ask students what they understand after having listened to a lecture that provided “the big picture”. I predict that most students will provide surprising answers to this question. Anyone who is still convinced that students are learning well from their lectures should try giving minute papers at the end of class, where students are asked to describe their understanding of an important concept introduced or discussed that day. Reading students' descriptions of what they have learned from even a “brilliant” lecture is often quite humbling, and can make faculty more aware of the difference between what was “told” and what was “learned”.

#### **“Students won't learn as much with the new methods”**

Moore acknowledges that implementing a variety of learning strategies in a course will result in a decrease in the content students learn. I disagree with this statement. Although faculty may believe they cover more material in lecture classes, alternative learning methods are often more effective. For example, requiring students to read assigned sections of their textbooks outside of class, allows in-class time to be spent on small group discussions, problem solving, and hands-on activities. I have found that I am able to cover more content using this approach than if I spent all four hours of class time each week lecturing to students.

#### **“These new methods won't transfer to my classes”**

Faculty interested in making changes in how they teach statistics often are concerned about how well they will be able to adapt recommended techniques to their particular courses and student populations. Many fear that these instructional strategies will not work well in their unique setting. For example, faculty who teach large lecture classes tend to believe that they are unable to use cooperative learning activities or student projects in their courses, which is not necessarily true.

#### **“Technology merely replaces calculations and graphs done by hand”**

Moore discusses the advantages of using technology to decrease hand computations, to produce graphics and to explore data. He lists the advantages and disadvantages of the new multimedia materials. However, he does not address new problems that arise when students interact with technology, leading to unanswered questions about the appropriate role of technology in statistics education. Unlike the research on outcomes related to different instructional methods (e.g., cooperative group learning), there is currently a scarcity of good empirical research to help statistics instructors better understand and monitor how students interact with technology and to choose strategies that best promote understanding.

For example, although technology provides powerful tools that can replace many tedious exercises formerly done by hand, it is not clear at what stage these should be replaced (e.g., should students first create a histogram by hand before having the computer generate one?). We are also unsure of what students actually see and understand when interacting with technology. There is a need for research to provide a deeper understanding of statistics learning and thinking processes in technological settings and to identify new problems that arise when technology is integrated with instruction. Some of the

questions we need research to address include:

1. What constitutes appropriate interaction between students and technology? For example, is it just as effective to see a simulation of a sampling distribution demonstrated in class as it is to interact individually with a simulation program?
2. How should the role of the teacher change and how should instruction change as more technology is introduced in statistics classes?
3. What are appropriate assessment methods to use that will accurately measure and monitor changes in students' learning as they interact with technology?

In addition to the myths about lecturing, coverage of content, applicability of new teaching methods, and uses of technology; there are also external constraints that impede the impact of educational reform. Two of these seem to be:

*Time:* Making substantial changes in teaching takes time. Many statistics instructors are not allowed enough time to investigate, prepare for, and implement instructional change. Furthermore, time spent on teaching is often viewed as time taken away from research, and research is what leads to promotion and tenure.

*Lack of academic freedom:* Many faculty who want to make changes feel constrained by their departments to use certain texts or to cover certain topics (e.g., formal probability). New instructors are sometimes given a syllabus to follow and overhead transparencies to use. Despite their interest in making changes, many feel powerless to do so, especially those in untenured faculty positions.

## 5 What Should be Done to Help Facilitate Changes in Statistics Education?

I think that in order to bring about change in statistics education, three efforts are needed.

1. *Increase the visibility of the statistics reform movement.* Inform more faculty about improved methods of teaching statistics and current resources available. One way to do this is to expand the number of summer faculty development workshops currently being offered for statistics instructors. Another strategy is to establish or improve communications between statistics education groups and other professional groups which include teachers of statistics (e.g., mathematics and psychology organizations).

2. *Develop observable models.* Develop ways for instructors to see models of good teaching in action, teaching that exemplifies recommended techniques and theories of learning. These methods need to appear real and transferable to other classes and settings. They also need to be convincing in their connection to positive outcomes: improved student learning and performance, positive student attitudes, and instructor satisfaction. I think one possible way to provide these demonstrations would be a multimedia CD ROM or web site for statistics teachers.

In order to create this resource, examples are needed of instructors and departments who have successfully implemented aspects of the reform. These individuals and their settings would be studied to document what they have done and how well their students have learned statistics. The resulting materials would include video segments of statistics teachers who teach in a variety of settings (e.g., a large class at a public university, a small class at a liberal arts colleges, and a community college class). These video segments would show a composite of an instructor's classes, as well as interviews with the instructor about their teaching methods. It might also include interviews with students about their perceptions of the course and show examples of students' work, such as presentations or student projects. The teachers selected for the video would represent innovative instructors who have successfully incorporated new techniques in their classes such as computer labs or demonstrations, small group discussions, hands-on activities, journal writing, or student projects. The CD ROM or web site would include links to the resources used by these teachers such as, data gathering activities, data sets, software programs.

3. *Obtain visible support from outside.* Strong support from government agencies and university administrators is needed to encourage change in the faculty reward system that will allow faculty freedom to experiment with new instructional methods and revise their courses, and provide time focus on improving their teaching. Faculty should also be encouraged to collaborate in educational research related to their teaching, addressing issues such as those raised relating to the use of technology. Publication of this “educational” research should be counted the same as published statistical research.

## 6 The Need to Monitor and Assess Statistics Reform Efforts

Along with the need for change agents such as the ones listed above, I believe there is a need to find out just how much of an impact there has been in statistics education to date, and to monitor the extent of this impact over the next few years. I think the statistics education community would be well served by a careful study of the statistics reform movement, disseminated in a report along the lines of a recent evaluation of the calculus reform (Tucker & Leitzel, 1995). Such a study would provide empirical evidence on the extent and impact of this reform and examine the teaching of statistics across disciplines and departments (e.g., mathematics, statistics, psychology, education, sociology, biostatistics, public health, and business). The project should include a large-scale survey of these departments as well as case studies of a few selected departments that have fully embraced the statistics reform. These “model” departments will be described to illuminate what is involved in implementing the reform and what the benefits are to faculty, students, and departments. This project could also be used to suggest case studies for the multimedia resource project described earlier.

Perhaps if more examples of successful instructional change are made available to faculty in forms they respect (e.g., multimedia CD ROM, internet web sites, and a rigorous study of the status of statistics education), the statistics reform movement will be better supported in its goals of transforming education and providing successful learning experiences for all students.

# Discussion

## Anne Hawkins

*President - IASE, Director - RSS Centre for Statistical Education, The University of Nottingham, Nottingham, NG7 2RD, UK, ash@maths.nott.ac.uk*

Clearly David Moore is correct when he asserts that the teaching of statistics has undergone considerable reform. The all-pervasive nature of statistics has led to the development of many new statistical methods, and to the spread of quantification in ‘user’-disciplines such as geography, economics, and the behavioural, social and physical sciences. Consequently, more and more people require statistical understanding if they are to be competent in their own work, if they are to work effectively with specialist statisticians, and if they are to fulfil their rightful roles in the societies in which they live. Moore is particularly concerned with introductory level university students, who now comprise by far the largest of our college-level audiences. He is thinking of reforms that will affect the less-specialist among our students. Much of what he says is therefore also directly relevant to school-level courses. Personally, though, I feel that specialist courses can also benefit from this kind of thinking.

Many countries of the world now espouse *Statistics for All* policies embodied within their mathematics curricula. We should never forget, however, that *Statistics for All* is not yet a universally-held

premise. There remain parts of the world where statistics for all would be deemed to be politically inconvenient for the government authorities. There are also still countries where universal primary education is so tenuous that statistical education is not seen to be a priority, although mathematics education for all is treated as such. This, of course, begs the question of whether someone can be considered to have been mathematically educated if they have not also been statistically educated.

As I have indicated elsewhere (Hawkins, 1996b), I remain somewhat sceptical that we should be satisfied with *Statistics for All* policies. Our true objective should rather be *Statistical Literacy for All*, so I must qualify as one of David's 'reformers'. Indeed, I would probably qualify as a particularly radical reformer, although not (I think) in the way that Moore describes. My reforms would *first* involve a change of emphasis in our teaching objectives, *after which* the appropriate synergy of content, pedagogy and technology to achieve these should be determined. I would share Moore's concern about the nature of any reforms if their source is uni-dimensional, but I would go further than him in saying that we may not yet have determined the right framework for reform.

The problem is that statistical education has largely evolved from a background that was significantly different from the needs and possibilities of today—in terms of the nature and practice of statistics itself; the available technology; and the students' characteristics, requirements and limitations. Reform is certainly required, therefore, but I am reminded of a traveller who, upon asking a local resident for directions to another town, receives the reply, 'Well, to be sure now—I wouldn't be starting out from here at all'!

The reform that we need must have the momentum and energy to challenge even the most fundamental and widely-held ideas about statistical education, and the ways in which these are currently manifested. Nothing should be taken for granted. We might even re-consider (notwithstanding curriculum and time constraints) whether the future of statistical education is really best served by its remaining a part of the mathematics education for most students. The reform, however, must be a process that is more than belief-based. Existing and proposed practices must be open to empirical scrutiny that can sort the 'better' or 'best' from the 'good' or the 'bad', in order to find out when and why content or pedagogy or technology do and do not work. It must also have the wisdom and finesse to avoid 'throwing the baby out with the bath-water'. As Moore indicated, for example, there is still a place for explanation.

In fact, I strongly agree with Moore when he argues for a 'more moderate' reformist approach that encourages the use of a *broad* range of ways to approach the teaching and learning process. (If he is a 'content expert' then I am sure that I am a *moderate* radical!) There is 'no one right way' to teach statistics although there may be many 'wrong' ways! Variety of methods and materials will always benefit both teachers and students, provided we have insights into how to use the available resources. However, the development of methods and materials should start from the point of what we already know about the teaching/learning process, and should take account of the extent to which we are aiming to produce specialists or non-specialists.

If we concede that there is a dimension related to statistical competencies (see Figure 1) that ranges from *non-specialists*, maybe through *para-statisticians* (Peter Moore, 1990), to *specialists* then it is reasonable to suppose that specialists will receive more training and acquire more competencies than non-specialists. The first point that I would make is that *everyone* at the outset is a non-specialist. Some will progress further along the dimension, however, than others. The question is, what competencies do we want the extremes of this dimension to possess? I would argue in favour of 'Statistical Literacy for All', that emphasises understanding over facts and tools, with specialists acquiring progressively more 'Statistical Literacy Plus', where the 'plus' possibly relates to more sophisticated/specialised *techniques*.

*Statistics for All* policies (as currently implemented), however, tend to focus on giving people the 'plus', albeit in varying amounts because if they are not fluent in the language and methods of exposition they cannot process, and therefore assimilate, as much as others. There is a sort of

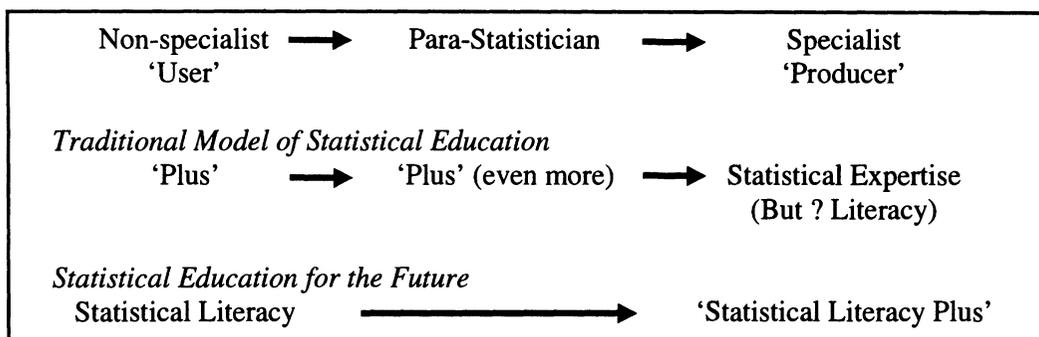


Figure 1.

tacit belief that a person will not be statistically literate unless he or she has acquired enough of the 'plus' type skills or competencies to become a 'statistical expert'. Moore also points to the *professional's fallacy*, whereby the traditional language and methods of exposition are seen as essential to the process—'What was good enough for me must be the (only) way to proceed!'. Alongside such beliefs is another assumption that 'statistical literacy' will *necessarily* emerge under such circumstances, i.e. that developing enough 'plus' is a *sufficient* condition for the acquisition of statistical literacy. 'Radical' reformer that I may be, I would argue that these assumptions can, and must, be challenged. If statistical literacy cannot be acquired prior to the 'plus', *Statistical Literacy for All* may be doomed. On the other hand, what good is a reduced collection of 'plus' without the statistical literacy to support its use? *Statistics for All*, in the absence of literacy, is worthless. If we cannot guarantee that the dosage of 'plus' administered to make 'statistical experts' will also make them statistically literate, what price their expertise? *Statistical Literacy for All* must be the bread on which some may spread butter, jam, or even caviar. Without the bread, however, there is nothing to support these optional toppings. Likewise, if we teach students a selection of the 'plus' tools and techniques rather than (or before) the principles and practical importance of statistics can be appreciated, we should not be surprised that '...by the time a scientist embarks on a research career, much of the reluctantly-learned and poorly understood material is forgotten. Often an aversion to statistics is all that remains!' (Employer's/trainer's comment to MEANS Project, 1997, see later.)

Unfortunately, there is not universal agreement, even within the statistical profession itself, about what statistical literacy is. At one level, there is a sense in which people 'would know it if they saw it'. However, this belies the very real, and on-going, debate about the nature of statistics and, more particularly, about the statistics that can and should be taught to all students. Statistics is not 'just sums', and it certainly is not just about algebraic derivations and proofs. It is clearly not a narrow area of specialism. It encompasses many different types of activity, requiring a range of skills and understanding applied in an infinite variety of contexts. These depend on different reasoning processes from others that comprise a mathematics education. In its simplest terms, statistical literacy can be interpreted as meaning an ability to interact effectively in an uncertain (non-deterministic) environment. It is not merely the possession of an ever-increasing collection of analytic tools and techniques, although this is the outcome that often results from present approaches to teaching statistics. A statistically literate person must understand the strategies for data collection and analysis, as well as the nature of chance processes and their relevance to data collection, and the assumptions that underlie statistical reasoning. The 'numeracy' with which mathematics educators concern themselves rarely includes thoughts of statistics and probability. The 'mathematics

of uncertainty' may crop up occasionally, but not really as a major issue. 'Graphicacy', too, has a rather specialised meaning to mathematics educators, one which seems to miss the mark as far as statistical literacy is concerned. Nor does *statistical* literacy feature prominently, if at all, in discussions about raising the *literacy* of the population. It seems that statistical literacy falls between a number of stools, and does not receive the widespread consideration that it should.

I will therefore coin the term 'informacy', because statistical literacy clearly incorporates the whole gamut of competencies needed for dealing with *information*; some that would otherwise be called 'literacy', together with numeracy, graphicacy, and the ability to visualise or perceive patterns, relationships and differences in data (presented in tabular or pictorial form). To be 'informate', one requires skills in summarising and representing information, be it qualitative or quantitative, for oneself and others. Modelling cannot therefore be the preserve of the specialist, or of those who progress ('educationally') to regression techniques and beyond. Even the simplest of charts has a function of representing or modelling reality. 'Informates' are able to select an appropriate model, and perceive its strengths and weaknesses for describing reality. They, therefore, have facility in *constructing*, not just in manipulating, models.

'Informates' have a good grasp of the nature of chance or random events, and are therefore less prey to superstitions and misconceptions that result from erroneous subjective intuitions. They are able to formulate probabilistic models and enumerate outcomes. Hence they can assess risks and likelihoods, and make reliable predictions and reasonable decisions in the face of uncertainty. At the same time, they are sensitive to, and able to learn from, feedback that indicates that their inferences or behaviour were less than optimal on a given occasion. They can operate in a multivariate environment, and are not constrained to uni- and bi-variate tools for describing it.

To be considered 'informate', people should be fluent in the language and principles of statistics and probability, and appreciate the inter-dependence of these two areas. They must be able to communicate, comprehend, and critically evaluate arguments that are couched in statistical or probabilistic terms. This is not a justification for continued, or increased, use of jargon in these areas. Rather, a move towards *statistical literacy for all* should be accompanied by a move towards *making statistical language intelligible to all*. If a specialist cannot talk to a non-specialist, it is all too easy to blame the non-specialist. However, in reality, failure to communicate may actually be a sign of the specialist's own statistical illiteracy, given that communication skills are by definition fundamental to (any type of) literacy. Training for 'informacy' also requires the development of a degree of computer literacy, because the use of technology is now an essential aspect of the practice of statistics by both specialists and non-specialists.

The above appears to be an extremely long and daunting 'shopping list' of skills and competencies for teachers to address. This is a challenge that statistical educators must face. It is much easier to develop a syllabus with a few statistical techniques listed (pre-defined to make for easy assessment) but this is not the way to teach *real* statistics. The tasks of the reformers must be to find ways of conveying 'informacy' to specialist and non-specialist students, and to convince the teachers of the merit and necessity of adopting these. David Moore argues for a balance between content, pedagogy and technology in the teaching/learning process, but points to the unwillingness of many specialists in these areas to recognise the importance of synergy between them. I agree. *Many* of the developments in statistical education in the past quarter of a century have been proposed from one particular perspective. *Some* have indeed evoked controversy. This is not to say that all such developments have been counter-productive. In general, the picture that emerges of present-day statistical education gives cause for optimism. However, there are certainly those, and I would count myself among them, who regret that such developments have largely been made in the absence of evidence-based understanding about the teaching/learning process.

The situation that exists in statistical education today can best be described as one of *Common Variance* (Hawkins, 1996a) in teachers' approaches to the subject. Different beliefs about what

constitutes 'basic' statistics, and about how the subject fits into the curriculum, serve to fuel the debate about how best to teach statistics. So too do differing views about the role of the 'professional' statisticians that education aims to produce. In some cases, research into statistical education has itself contributed to explaining the observed variance. For example, Piaget and Inhelder's work (1951, tr 1975) and that of Fischbein (1975) have led their followers to have very different ideas about what would constitute the appropriate treatment of probability early in the school curriculum.

The whole issue of *what is intellectually hard* and therefore *what is academically meritorious* is something to which many, raised in Moore's 'peculiar culture of university faculty', pay only lip-service. This is borne out by the content and style of the many texts entitled *Basic Statistics*. What is 'basic' to one author is clearly not necessarily 'basic' to another. A great deal of variation exists with respect to the role that mathematical explanations play in so-called 'basic' or 'introductory' texts. Likewise, there are differences between the balance struck between 'theory' and 'application', or indeed the 'theory of application'.

Moore asserts that research has much to tell us about how to reform the teaching/learning process. I would agree wholeheartedly with this. However, I am less sanguine than David about the availability of that research. More particularly, even where that research does exist and is of sufficient quality, it is not always taken into account when recommendations for change are made or implemented. Shaughnessy (1992), for example, points to the gulf that exists between the findings of cognitive research and classroom research and practice.

To say that research into statistical education is guiding our progress is not strictly true at the moment. The published research is still predominantly a collection of reports of 'positive' outcomes. It does not tell us about things that did not work, and therefore about what things we should avoid. It is still rather 'development-oriented', and research into *why* a particular teaching approach is effective is relatively rare. Even when relevant research work has been carried out, statistical educationists are only just beginning to build upon existing studies. Gradually, we need to create a more synthesised body of literature, and better dissemination practices. Otherwise, we will continue to be prey to the prescriptive myths, fashions and 'fads' that are propagated by those with opinions about statistical education, and loud voices—the *really* 'radical' reformists (i.e. not me!) to whom Moore alludes. Such problems are apparently not the sole preserve of statistical education. Hart (1997), for example, highlighted the 'total lack of any research basis' that guided the introduction of, and subsequent modifications to, the UK National Curriculum in Mathematics. Ironically, of course, the greater part of the *systematic* instruction in statistics and probability for the UK is now conducted within that mathematics curriculum, and so statistical education is prey to all the generally undesirable results that Hart describes, plus a few more subject-specific ones. In particular, there is a 'tail wagging the dog' phenomenon. Once the statistical content is structured in a certain way within the curriculum, beliefs about relative difficulty levels within the subject matter, for example, quickly crystallise into 'facts about appropriate pedagogy' that are very hard to rectify subsequently.

The issue that Moore raises about statistical education producing competitiveness rather than co-operativeness is a highly pertinent one, and one that is extremely resistant to change. Current assessment practices are major stumbling blocks to reform because they encourage competitiveness. The Royal Statistical Society Centre for Statistical Education is currently conducting a two-year project entitled *Matching Education, Assessment and Employment Needs in Statistics (MEANS)*. One of the recurrent themes of the seminars/workshops and other network interactions associated with this project has been the mis-match to which Moore refers *vis-à-vis* the Motorola Company requirements of their employees.

However, what is emerging is not only a resistance on the part of the university academics to change, but also a resistance on the part of the students themselves. Even when they are available, 'team-player' skill courses may be avoided by the students in favour of more certain (less demanding? more predictable outcome?) 'theorem-proof-theorem-...-exam regurgitation-good grade' courses.

The mechanism whereby this might be changed is far from simple. It requires radical adjustments in the perceptions and practices of academics, students *and* assessors. Such changes may well require the sort of 'leaps of faith' that are rarely, or not very readily, undertaken in the corridors of established academia. There must also be *synchronous* reform, because change on the part of one group will not be sustained without reciprocal changes by the others. The Public Relations exercise that is required to encourage people so fundamentally to re-think their approaches to statistics teaching will be helped by the availability of hard evidence.

A recent sabbatical visitor to the UK who travelled relatively extensively, looking at statistical education practices in higher education, went on to report that he had found only one institution which had a real commitment to changing the pattern of statistical education in its mainstream specialist statistics courses. My own experience, particularly based on responses to the *MEANS* project, leads me to believe that this is an over-pessimistic view of UK developments. It is certainly one that underplays the effectiveness of individual 'reformers' and their influence on the statistical education experienced by their students. However, I can believe that there may be substance to the point made with respect to departments or institutions looked at as a whole. Even those departments who give strong support to the need for reforms in statistical education are not always completely ready to examine and/or reform their own teaching practices. It is all too easy for academics to fall back on the content-driven arguments that Moore describes. In fact, these become almost a defensive reflex in the economic and competitive circumstances that prevail in today's world of academia. And how much easier it can seem to ban the use of calculators in examinations rather than change our teaching and assessment methods (Garfield, 1994) to accommodate the increases in scope and power that technology offers the practice of statistics. By limiting the use of technology, though, we seriously distort the nature of the statistics that our students can and do learn. Statistics is not a subject that can be frozen in some intermediate technology period.

We have a responsibility to provide courses that allow students to experience the *real* nature of today's (and the possibilities of tomorrow's) statistics. Assessment methods must reflect, not impede, this. They cannot be allowed adversely to dictate content and pedagogy. To re-think our objectives for both non-specialist and also specialist education, and to recognise and embrace new possibilities, however, will require many of us to *forget old ideas*. This can be an uncomfortable process, but I would support Moore in his call for reforms that exploit the synergy of content, pedagogy and technology, especially if such reforms might put us in real danger of helping our students to achieve the most effective learning. Sadly, some of our current practices do not suggest that we really want to do this, nor do they always present the synergy of skills, knowledge and understanding that represent the *real and adaptable* natures of statistics.

# Discussion: Let's Stop Squandering Our Most Strategic Weapon

**Roger Hoerl, Gerry Hahn, Necip Doganaksoy**

*General Electric Company, Schenectady, NY, USA*

## 1 Introductory Remarks

We would like to express our appreciation to David Moore and the *International Statistical Review* for inviting us to participate in this discussion. We are delighted to be “part of the team” helping to improve introductory statistics education. Our comments will be from the perspective of business and industry—the employers, as engineers, scientists, managers, etc., of most students of introductory statistics courses. We were very impressed with Moore’s paper, and found little to take issue with. Our comments will therefore focus on extending the discussion in some areas particularly important to business and industry.

We view the intro course as our profession’s most strategic weapon. More and more of our colleagues, both technical and management, have had such a course in their education. Unfortunately, for the vast majority, the experience was a negative one. This causes us significant problems in promoting statistical concepts since many of our customers have already concluded that statistics is not relevant in their work. In short, we feel that we have often squandered our most strategic weapon—despite the recent efforts of Moore and others. Instead, we would like to see our future customers and associates emerge from the intro stat course with enthusiasm for the subject, and a solid appreciation of basic concepts.

A major issue hinted at by Moore, but not specifically addressed, is the low priority given to intro courses by academia in general, and Statistics Departments in particular. For example, who ever received tenure for teaching intro courses well? This course is too frequently taught by junior faculty members with no real applications experience, while more senior faculty, who might be able to provide a broader perspective, concentrate on research and advanced courses. Our entire profession must share the blame for this tremendous blunder. We in industry have not sufficiently emphasized the need for changing priorities—much less offered assistance. For example, few of us have volunteered ourselves to our local universities to help add a more practical flair to their introductory offerings. Also, ASA and ISI can play a role here—both due to its strong academic ties and by its own actions, e.g., by recognizing teaching of introductory courses through special awards or by making it one significant criterion for ASA Fellowship and ISI election to membership.

We certainly agree with Moore’s points about the difference between mathematics and statistics. Thus, we applaud his statement that “. . . statistics is not a subfield of mathematics, and that in consequence, beginning instruction that is primarily mathematical, or even structured according to an underlying mathematical theory, is misguided”. We are, therefore, somewhat concerned that readers may misinterpret the fact that Moore talked about “mathematical sciences” in general, and then statistics in particular, rather than focusing the discussion on statistics education alone.

The casual reader may get the erroneous impression that the goals for the intro course listed in his Figure 1, which stress higher-order thinking and flexible skills, are generally agreed upon. We surely agree with these goals, but have faced vehement resistance to them from some of our statistical

colleagues (not to mention referees and editors of professional journals). Many mathematically oriented academicians consider calculating the “correct” answer to mathematical—but not necessarily realistic—problems as the “meat”, and view higher-order conceptual thinking as “fluff”. We have also found that many instructors have not given careful thought to course objectives, or consider their major goal to be that of exposing students to as many formal statistical tools as possible. This might be because a strict curriculum, perhaps built around a specific traditional text, has been pre-ordained for them by the department.

We advocate more explicit discussion of goals. Also, we would like to see consideration of “softer”—but in our judgment important—goals such as reducing “statistical anxiety”, and nurturing an appreciation that statistical thinking and methods can be helpful in solving real problems. The typical intro course today tends to foster just the opposite—our future customers, much too often, think of statistics as a long collection of chi-square tests and  $F$ -ratios. Long term, we must come to some consensus on goals if we wish to ever agree on the course design. Short term, we cannot wait for this consensus to occur, but must act now.

The format of our remaining comments will follow Moore’s major topics, in order of the volume of our comments, and not coincidentally, the degree of change which we feel is required: Content; Pedagogy; and Technology.

## 2 Content

We applaud virtually all Moore stated on this topic; for example his stand that we should give formal probability the “guillotine”. We also wish to stress the importance, from a business or industrial point of view, of tying the tools together into overall approaches to scientific inquiry. This is hinted at in various places in the paper, and we would like to extend the discussion. We have found that even “A” students do not emerge from the intro stat courses with the ability to cope with real problems that require the sequential use of several tools, including planning the investigation itself. In contrast, GE is currently applying the “Six Sigma” approach, popularized by Motorola and adopted by various other companies, throughout the corporation. This methodology for improvement is highly statistical in nature. While as professional statisticians, we might quibble with some specific elements, we must agree that, on the whole, this approach has been extremely successful. One key to its success is that there is an overall four step approach which non-statisticians can easily understand: Measure, Analyze, Improve, and Control. Specific statistical tools are then taught in class within this context, so students can see how and where each fits.

Similarly, part of Taguchi’s success must be credited to providing an overall process for robust design. The typical problems faced in business and industry are very amenable to application of statistical tools, but few, if any, can be solved through one tool used in isolation. Moore correctly notes that “students. . . are unlikely to ‘construct’ or ‘discover’ the big picture for themselves”. This is an area that merits further exploration. The current mathematically-oriented research on individual tools tends to be “an inch wide and a mile deep”, while the significant practical problem of how tools should be tied together is “a mile wide and inch deep”; this is, indeed, fertile, but unplowed, ground for further exploration.

One example of a “big picture” approach which could be used in the classroom concerns the common problem involving multiple sources of variation (i.e., lot to lot, within lot, test error, and so forth). While particularly relevant to manufacturing, this issue arises in many contexts. To address such problems, students must first understand the concept of variance components (including the ever-present time component), and the economic value of reducing variation. The assessment of variance components is often left to advanced classes because of the complexity of the usual techniques involved (components of variance models), but can actually be accomplished with simple descriptive and graphical tools, such as boxplots, run charts, and character graphs. (This might also

provide some practical motivation to learn ANOVA for those who choose to continue their statistics education). Once the major sources of variation are identified, the focus (and tools) shift towards determining root causes of variation so they can be eliminated or reduced. This typically involves design of experiments, which again can be analyzed graphically instead of by formal ANOVA and regression. Once improvement efforts have been completed, there is a need to gather additional data to verify and quantify the improvement, and compare the “before” with the “after” process with run charts, histograms, or other plots. If improvement is sufficient, we would like to leave some control system in place to maintain the gains, again, perhaps, making use of run or control charts. We believe this type of exercise would be significantly more valuable than solving textbook problems which assume a “random sample” of data, and suggest students perform a specified analysis.

We would also like to see more emphasis on teaching “statistical thinking” by helping students unlearn the “law of small numbers”. Most managers and technical people in industry, even those who have had a formal statistics course, still tend to think deterministically. They expect mass balances to match exactly, or actual financial figures to exactly equal budget. Any “variance” from budget must be explained. The costs incurred by US businesses searching for “explanations” for random variation may be in billions of dollars. This is a tragedy. Understanding that variation exists in all processes should be the cornerstone of statistical education, but most students are not coming out of the current intro course with this understanding. Why not?

Another “statistical thinking” concept that should be included in intro courses is that output measurements are the result of a process. In order to improve the outputs, we need to study and improve the process. While this is not statistical per se, it involves statistical thinking, and is very relevant. Unfortunately, intro stat courses often focus on how to quantify, test, or model output measurements, but not on how to improve them. This is a passive approach that does not directly lead to improvement, and provides little motivation for students to utilize statistical methods. For example, we can do a  $t$  test or ANOVA to prove that different branches of a bank have different average processing times for approving credit applications. So what? We would expect this—why should different branches, operating in, at least somewhat different environments and with different staffs, have identical average transaction processing times? If we understand that these output measurements ( $Y$ 's) are the result of a process, we would be motivated to try to understand the process better so we can improve the performance of the poorer branches to at least the level of the better branches.

The use of statistical methods, in conjunction with subject matter knowledge, is the best known approach for uncovering the important variables ( $X$ 's) in the process, and determining how to adjust or change them to improve the output. This is fairly obvious in engineering applications, where we can usually see the process. In business applications it is not, and business students are not generally taught process thinking; instead they are typically taught to focus on outputs. Therefore, exposing students to process thinking would make a dramatic difference in their motivation—they will see that statistics can be used to improve results, not just to “win arguments”, or to prove something which we already knew. Process thinking would also help students understand the practical issues associated with sampling from dynamic processes, rather than the theoretical “static bowl”. In real applications, of course, we virtually never sample from static populations, but from dynamic processes. More on that shortly.

Another topic which we would like to see considered in the intro course is the issue of data quality, or the importance of obtaining data which is appropriate for the problem at hand. This is typically the limiting factor in many applications. Data are too often treated like a commodity in statistics classes. For example, students may be taught how to answer the question “What is the required sample size?”, but are rarely asked to address the more fundamental questions, such as: “What information is really required to solve the problem? To what degree do the data at hand meet this need? What additional information needs to be obtained in the future and how?”

A common situation is to have large amounts of the wrong data. For example, past factory data

almost always exist in abundance (in various degrees of accessibility), but are rarely appropriate for solving real problems. The same is also true for field data. Yet, it is not uncommon to find fabricated examples, supposedly based on historical plant data, in introductory textbooks. These reinforce the completely unrealistic expectation that the needed data is there waiting for us, and all we need worry about is the optimum statistical method for analysis. A classroom discussion of data quality analogous to that in Box, Hunter & Hunter (1978, Section 14.7) would be much more useful. Similarly, we have found in working with our financial businesses that summary data on current customers is readily available, but actionable data on potential customers is not. In order to grow market share, we obviously need the latter rather than the former.

Moore discusses the difference between observational and experimental data, but the issue of data quality goes much deeper. In particular, we urge considerable classroom attention to the topic of the quality of existing data—often obtained for purposes other than analysis. This can readily lead into the topic of the proper planning of investigations.

A partial explanation for the shortcomings in statistical education goes back to the math versus statistics issue. Perhaps because of the over-emphasis on mathematics, statisticians seem uncomfortable with statistical concepts which cannot be derived or proven mathematically. The omnipresence of variation is admitted, but often not clearly explained (Simulating a histogram on the computer does not teach business people how to interpret a financial report). Process thinking is generally not taught, hence reducing the motivation to learn and apply statistics. Data quality is ignored while data quantity is emphasized. The concept that in the real world we most frequently sample not from static populations, but rather from dynamic processes is not well understood or taught. Deming (1953) wrote on this last issue, using the terminology “analytic” versus “enumerative” studies over forty years ago, but as a whole the profession still doesn’t get it. As noted by Moore: “Those specific examples are instructive, for they point to core statistical ideas that are not mathematical in nature.” We would like to see understanding of these core concepts among the highest priority objectives of the intro course.

As our previous comments have suggested, we feel—and believe Moore agrees with us—that the intro course need also include

- Basic concepts of planning data-based investigations, including the major concepts (but not the details) of survey sampling and design of experiments, ideas that are much too important to leave to advanced courses to which most of our future customers will not be exposed. We feel this can be accomplished by following Bill Hunter’s (1978) suggestion of having students plan and conduct their own investigations—and potentially critiquing each other in small groups.
- Emphasis on graphical data analysis.
- Some introduction to user-friendly statistical software, or statistical and graphical features of popular spreadsheet programs.

With all of these proposed additions, what should we let go from the current course, other than formal probability? Certainly, calculational details that do not add to understanding, are top candidates, as Moore suggests. Statistical significance tests, of all types—from simple  $t$ -tests, to analysis of variance, to tests on the significance of regression coefficients (as opposed to confidence intervals) are next on our list. The major argument for mentioning these at all, in our opinion, is their continued predominance in the literature—rather than their practical usefulness. At a minimum, one or two tests can be reviewed to teach the process of hypothesis testing, and then others listed or referred to in an appendix. We certainly do not need to go through the calculations of 10 to 20 tests ( $t$ ,  $F$ , Chi-square, one sample, two sample, multiple sample, paired, sigma known, sigma unknown, equal sample size, unequal sample size, etc.), as is often done. It is much more important to have

students of introductory courses understand what such tests provide—and do not provide—than to have them become adept at their use. If further cuts are needed, we would emphasize the basics over the details. The specifics can then be left for subsequent courses—which, we hope, might now draw a wider audience of motivated non-statisticians, as well as future statisticians.

### 3 Pedagogy

Again, Moore's points here are well-taken, and cannot be easily disputed, e.g., "Students are not empty vessels to be filled with knowledge poured in by teachers; they inevitably construct their own knowledge ...", the summary in Figure 1, and his comment that even with interactive classrooms, leadership of the instructor is still necessary. While experiential learning is critical, and can be especially useful in imparting understanding of the planning of investigations, overuse of class participation can result in the "blind leading the blind".

A major point not made, which relates closely to our previous discussion on teaching overall approaches to scientific inquiry, is that the *sequencing* of topics within the course needs to be rethought. Most current intro courses come across to students as a miscellaneous collection of tools in a random order. Application of statistics in the real world is not miscellaneous or random, there is a natural sequence, or "flow". First, we must motivate the topic, i.e., explain why students need to understand the material (beyond meeting a degree requirement). Case studies, or exposure to real problems can be excellent ways to motivate the use of statistics. Next, we must teach "the big picture", i.e., overall approaches, reinforce this with case studies and examples which illustrate the sequential use of the tools, and have the students do *real* projects (*not t* tests on male versus female heights) that require integration of several tools. We concur with the approach recommended by Harry Roberts that case studies can be excellent vehicles to teach the theory. Similarly, the Harvard Business School has done quite well with a case study approach.

To be more specific, we recommend that at the outset of the course, or at some other early strategic point, we introduce a number of general key problems that students can readily relate to—such as evaluating the impact of smoking on health, or predicting the outcome of a forthcoming election. Other examples are listed in books such as Tanur *et al.*'s (1989) *Statistics: A Guide to the Unknown*. In fact, this book, or others like it, such as Moore's (1997) own *Statistics: Concepts and Controversies*, might make excellent supplementary reading. If students are exposed to these real problems, and taught overall approaches involving several tools, they will be more inclined to learn about the tools themselves, i.e., we will create "suction" for the rest of the course. The tools can then be taught in the context of the overall approach; hence students will understand how the tools fit together and apply to real problems. As has been noted elsewhere (Hoerl & Snee 1995), this approach is consistent with current behavioral and educational research. It also takes more of a holistic or "systems" view of both the statistical method, and the educational process.

### 4 Technology

We agree with Moore's fundamental point that technology should serve content and pedagogy. Unfortunately, we are sometimes infatuated with technology to the point where technology becomes the "what" we teach, rather than the "how". Our concern is that computer science will dominate statistics in the next century, just as mathematics has dominated in the past. The current emphasis on getting computers or Web access into the classroom has obscured the bigger issue—what should we be teaching with these computers? In our rush towards technology, we may be reinforcing the wrong thing, e.g., teaching students to calculate rather than think. The question we have heard many times: "How can we make better use of technology in the classroom?" is misguided in that it starts with a predetermined answer, and then tries to state the problem so as to support this answer. A better

question might be: “How can we better achieve our objectives for the introductory statistics course?” We believe that this statement of the question will, in fact, lead to greater use of technology, but we would like to see the horse back before the cart. An obvious obstacle to asking this question is that there seems to be little common ground within the statistics profession as to what the objectives should be. Assuming the goals recommended by Moore, which we agree with, the most obvious use of computers is to automate calculations, so that we no longer have to teach formulas, except for illustration, providing insight in the manner suggested by Moore. When formally teaching the tools, we can teach from software output, and emphasize the interpretation of the results. The Web might then become a great vehicle to provide access to a variety of real, sequential case studies.

Similar comments could be made about multimedia and graphing calculators. Graphing calculators are available today to eliminate the need to teach formulas. The real potential of multimedia is perhaps further down the road. It is a minor point, but we would not consider graphing the  $t$  distribution as “graphical data analysis”.

We value simulation, but feel that it is important to differentiate between using simulation as a teaching tool, e.g., to demonstrate the central limit theorem, or to help students differentiate between the concept of a confidence interval and a prediction interval, versus using simulation as a problem solving tool, e.g., to develop confidence intervals in complex situations. The latter may arguably be delayed to more advanced courses.

## 5 Summary

We thank David Moore for an excellent and thought-provoking discussion of ways to improve introductory statistics education. If his recommendations were broadly implemented, the intro course would be dramatically improved, and the impact of statisticians greatly enhanced. Our major points are

- We must drive for consideration of and consensus on objectives for introductory education. Our vote would be for “statistical thinking” concepts, i.e., helping students unlearn their deterministic view of the world, and viewing outputs as the result of a process, over memorizing formulas or exposure to a large number of techniques.
- Perhaps our major addition to the paper is, that we need to formally teach overall approaches to scientific inquiry utilizing several tools sequentially. This needs to be reinforced with case studies tying various tools together, and with real projects.
- We advocate teaching students about data quality, and various other non-statistical core concepts.
- We also agree with Moore’s points on pedagogy, but would add that we must completely rethink the sequence of topics.
- We feel that technology is a “how”, not a “what” that should be brought in to help us meet the course objectives.
- There is currently little tangible incentive, and in fact significant disincentive, for statisticians in academia to spend time on improving introductory courses. Could we honestly encourage young professors, seeking tenure, to spend the time required to fundamentally improve their introductory courses at the expense of publications? This is a problem with the system, and cannot be resolved easily by individuals working in isolation. However, as with the need for goals, we cannot afford to wait for this issue to be resolved. We encourage readers who have

influence over such policy to act to help change the situation now. They may even be able to evoke self-interest. After all, Lucas (1996) has noted that those statistics departments which have prioritized introductory “service courses” are among the few that are prospering in the current environment.

## Discussion

### Jon Kettenring

*Bellcore, Morristown, NJ, USA*

I believe a strong case can be made for radical changes in the way we teach statistics to beginners. In that spirit, David Moore’s article is both timely and exciting: it provides a thoughtful logical alternative plan for how such courses could be taught today.

I am especially concerned about students who end up working in industry, many of whom will not have substantial additional training in the subject. So it is best to assume that what they pick up in the first course will determine how they think statistically on the job.

The reason for this concern is my perception, after many years of experience in industry, that gross inefficiencies, major tactical and strategic errors, and expensive mismanagement of the enterprise result from a collective inability to learn from relevant data. I’m talking about data that are or should be at the fingertips of most employees—and especially managers. These missed opportunities can be found at all levels but are magnified as one moves up the corporate ladder.

If the introductory statistics course is to be a vehicle for improving the situation, then it must ensure that sound statistical instincts and thinking become part of the permanent intellectual bloodstream of the student.

Now what do business people need to know about data and statistics? They need to have an appreciation for what it means to manage by data: how to collect it, how to sample it, how to view it, how to model it, how to draw inferences, how to assess uncertainty, and how to integrate statistical analyses into the larger context of the business problem. This does not require a lot of detail or theory. However, it does require ample exposure to real problems in order to gain experience and to develop the instincts that will be needed on the job. What is possible to expect from an analysis? What questions should be asked of others who present statistical findings? How can one tell when something is fishy in the data or the analysis? In short business people need a level of statistical savvy that is probably comparable to what they have when scrutinizing a budget report or digesting a business case in a sophisticated fashion.

Moore’s approach is attractive because it downplays a lot of the tedium and cookbook aspects of statistics as well as nice but unnecessary aspects of probability and mathematics. It replaces the fire hose approach of spewing out information to students, to the extent they can take it, with the more natural and holistic one of learning through active interaction with ideas and data. Less ground will be covered with Moore’s approach but more should stick and be genuinely useful later on.

## Discussion

**Jayanta K. Ghosh**

*Indian Statistical Institute and Purdue University*

As we look back on one century and prepare for another, we cannot but be aware of great changes that are taking place. One of these may be called democratization, leading to enfranchisement of new nations and new groups within nations. The other sort of change is really a second industrial revolution, of which key ingredients have been the production, marketing and consumption of information, often in a very interactive way. If these movements retain their present momentum for a considerable period of time, the future will be very different from what we are used to and not all the transitions will be easy. Many skills will become redundant and democratization will need new levels of self-discipline and tolerance of different points of view. New values and skills will be needed to survive and prosper in a changing world. Depending on how we accept the future, preparing for the future can be a challenging, rather than a frustrating or frightening, experience.

The reason for viewing Professor Moore's paper in this more general background than the one he himself has provided is that it then becomes all the more compelling to come to grips with the emerging changes that he describes so well. I also seem to see the forces of democratization at work both in the emphasis given to an interactive mode of instruction and in the changes of content that Professor Moore explicitly describes as "democratization of mathematics". In these matters, as in the choice of right technology, the real issue is not whether such changes are absolutely good or absolutely bad, but, given these new opportunities and new constraints, what we can do for the teaching of our subject that will make it useful, relevant and interesting in a changing or changed world.

Professor Moore has argued so well for the points of view and options he recommends—after taking a broad and balanced overview of what is available and what is possible—that it is difficult to disagree with him on any major point. The following discussion therefore focuses on some important details and one broad issue: the role of a teacher in a university.

Professor Moore's recommendations are generally made in the context of "teaching statistics to beginners at the university level". At this level there does not seem to be any conflict between an interactive mode of interaction and the amount of new concepts and skills that a student is supposed to learn. The conflict becomes most serious for higher level service or theory courses. Interactions are very effective for generating interest in what is taught and absorbing it but they do seem to diminish the amount of information that can be conveyed in a course based on lectures. Multimedia can help only if students are ready to put in more time than they presently do. Given their total course load and other compulsions, is it realistic to expect more investment of time in a course? If not, there would seem to be no alternative to reducing the syllabus or the total course load. I assume the interactive mode, once it comes, will be used for most courses, not just for beginners.

Incidentally, unlike the other ingredients of the changing scenario, an interactive mode of instruction does not strike me as quite new. Undergraduate instruction in the Indian Statistical Institute has been and still is quite interactive (but we cover less material than in the U.S.). And Madame Montessori's highly successful system of teaching children is adaptive to each child, interactive and based on an appropriate technology that is an essential part of the system. What I find especially impressive about the Montessori method is that such a highly individualized method can be taught to would-be teachers so that the method has survived for generations after the death of Madame Montessori.

Such has not been the case with other attempts at interactive teaching, as in Tagore's Santiniketan (The Abode of Peace) and many innovative and initially successful experimental schools elsewhere. Success in these experimental schools seemed to depend on the skill and enthusiasm of a few rare gifted individuals.

As we opt for interaction, we must explore the possibilities and make a correct choice as in the case of technology.

I turn now to contents. Historically, our discipline evolved by combining collection of data and data based studies with probability based inference in the face of uncertainty. In this blending the concept of uncertainty in probability theory corresponds to the notion of variability in statistical data and to de-emphasize this would take away an essential flavor of the subject. Perhaps Professor Moore only wants to de-emphasize intricate mathematical calculations of probability in exotic problems. That would be natural and right in a basic course for beginners. But take away probability completely from a basic course, and it would cease to be completely basic. One of Professor C.R. Rao's favorite stories about his teaching B. Stat at the Indian Statistical Institute in the sixties is that he made the students go to hospitals and collect data about the sex of newborns. The data was a sequence of females and males. They then compared this with the data generated by random numbers or coin tossing and understood the conceptual unity underlying the uncertainty and regularity in these apparently very different situations. A two way path from an experience like this to elementary probability theory should be a natural component of any basic statistical course.

On the other hand, I agree that we should only emphasize formulas that tell us something important or useful about the quantity being computed. For example, I too have stopped telling my class formulas that we learned as aids to our calculation. For example, I don't discuss  $var(X) = E(X^2) - \{E(X)\}^2$  but once in a while discuss  $E(X^2) = \{E(X)\}^2 + var(X)$  to stress that expectations don't commute with non-linear operations.

I finally come to technology. I have no doubt at all that Professor Moore's choices, like his rejections, are right. The only thing that bothers me is that introduction of technology, though inevitable, will further increase the difference between students from developing and developed countries. Is there a way for developing countries?

There is a remaining issue that disturbs everyone—the role of the teacher in the new environment. The position taken by Professor Moore is balanced and positive but many others have painted a darker picture where the teacher is redundant, or nearly so. I myself tend to agree with Professor Moore, but often feel disturbed by discussions which seem to hold the teacher, or his research interests, responsible for whatever problem a student faces as she copes with the transition from a school to the more demanding environment of a university. A good university will continue to be a place where good men and women wish to come to engage in their scholarly pursuits and communicate to the next generation the values, knowledge, and skill that have served them well.

# Discussion

**Richard L. Scheaffer**

*Department of Statistics, University of Florida, Gainesville, FL 32611, USA*

## **The Setting**

Allow me to set the stage for my remarks by providing a little personal background for the perspective that I bring to this discussion. As a teacher, administrator, writer and participant in many committees, I have been active in statistical education for over thirty years at the college level and for nearly twenty years at the school (K-12) level. Mostly, these activities have been organized through the American Statistical Association (ASA), the International Association for Statistical Education (IASE), a section of ISI, the National Council of Teachers of Mathematics (NCTM), or the Mathematical Association of America (MAA). It is important to note that the first two organizations have statistics as their purview while the second two have interests ranging across a broad spectrum of the mathematical sciences. I believe that statistics education must be a unified effort of the statistical and mathematical sciences so that sound teaching of modern statistics gets woven into the fabric of education in quantitative thinking and practice. It is true that good statistics courses can be found in other disciplines, but it is the statistics/mathematical sciences disciplines that must take the lead on rigorous and forward-looking development of instructional programs in introductory statistics, ideas from which should be supported and extended in other quantitative courses. No other discipline is interested in developing and supporting a well-rounded and integrated program of education in statistical concepts and practice. I agree that statistics has its own substance, content and modes of reasoning, but these must become part of the education of mathematical scientists, whether they intend to teach at the school or college level. There are too few statisticians and statistics departments to do otherwise.

Practice leads to the consideration of the other key component of statistical education, business and industry (or, what we might refer to as those in charge of the workplace). Schools, colleges and universities must work closely with leaders from the workplace not so much so that our graduates (at whatever level) are assured of getting jobs but simply because both sides have much to offer in terms of content, pedagogy and technology, the big three of Professor Moore's paper.

Before moving to more specific comments I must say that much of the insight I have gained in statistics education has come from the books, papers and talks of David Moore, so I have no substantial disagreement on the main points that he elucidates in his paper. My comments may be taken more in the spirit of amplification or fine tuning around the edges. In particular, I agree with the synergies (at least the potential synergies) among content, pedagogy and technology even though I will now separate my remarks into those three categories for simplicity.

## **Content**

With regard to the content of an introductory statistics course, statisticians are in closer agreement today than at any previous time in my career. As an illustration consider a vignette from the history of the development of the Advanced Placement (AP) statistics course in the USA. (An AP course is a college-level course taught in the high schools.) Efforts to establish such a course became serious in

the 1970's but when statisticians were invited to meetings to discuss such a course, the statisticians spent all their time arguing among themselves as to the proposed content. By the late 1980's a task force of statisticians and teachers of statistics considering the feasibility of such a course had little disagreement on the main threads of content. Thus, constructive work could begin on how to make the course a reality.

The content of the AP course parallels that commonly seen for introductory statistics courses taught in statistics and mathematical sciences departments, and reinforces Professor Moore's assertion that content is moving toward the concrete, emphasizing data exploration, data production and inferential reasoning, but not the memorization of a list of formulas. On what might be thought of as a continuum from data to inference, I think there needs to be stronger emphasis at both ends. In the generally positive emphasis on data that now permeates introductory statistics (both in the schools and in colleges), teachers and students grab data wherever they can find it and rush to plot it on whatever plot they may have learned most recently. (Most recent is always best!) What is lost is a thorough discussion of how the data originated, what the numbers might mean, if anything, and what plot or numerical summary might be appropriate. So, categorical data gets put on stem pots and averaged; age gets subtracted from heart rate. In short, more emphasis must be placed on number sense.

At the other end of the spectrum, concepts of statistical inference often get short changed by the amount of time and energy placed on data exploration. Plots are fun and easy these days, but statistics is more than finding a clever way to display data. Decisions must be made and this requires that students understand the concepts of inference, as well as the fact that in some cases formal inference is neither possible nor desirable. Some introductory statistics courses either have become pure data exploration or have remained exercises in formula manipulation with the formulas now residing on a piece of technological equipment. Some instructors view a modern course in the subject as a mixture of the two. All three outcomes are undesirable. The key is to find ways of teaching inference that are in keeping with the notion of constructivism (naive, of course) but still allow closure on a few critical ideas.

I agree that only the probability essential for understanding statistics should be taught in an introductory statistics course but, in light of the above, I opt for a little more in the essential area. The main thrust is certainly relative frequency notions of basic probability and emphasis on distributions rather than calculation of single probabilities. It is important, however, that students grasp the notion of conditional probability, if only from a two-way table of data. The concept of association between two variables is a theme that carries through much of statistical inference, from looking at categorical data on a two-way table to looking at a scatterplot, and this concept is best understood and expressed in terms of conditional distributions. Familiarity with probabilistic reasoning is also an essential part of the understanding of randomization, and the key role it plays in designing statistical studies.

Seeing the synergy between data analysis and the core mathematics of primary school is an astute observation that should be amplified. It begins even before whole number counting, as children are arranging buttons into piles of different colors (categorical, unordered) or separating the books on the bookshelf into small, medium and large sizes (categorical, ordered). I have seen elementary school teachers get absolutely excited about the possibility of using a stem plot to teach place value. The commonly heard statement that elementary teachers are scared by statistics because of their non-quantitative backgrounds is categorically false. They are put off by irrelevant presentations of statistical ideas. This is the place to begin teaching statistics and, cast in the right spirit, elementary teachers will make great use of it.

## **Pedagogy**

That the learning of statistics must move from passive to active seems to be accepted by a large number of teachers these days. At least in the United States, today's students seem to have little

capacity to sit for a long period of time and absorb a lecture. There is a fast-paced world. The struggle in the teaching is to find an appropriate balance between the “telling” of the past and the “discovery” of the present. Both have strengths and weaknesses for a subject in which there are some specific ideas that we want students to learn, as Professor Moore correctly points out.

In order to have confidence enough to move from a course scripted by a textbook to one that is free for discovery, the teacher must clearly understand the goals of the course and must have a multi-faceted knowledge of the content in order to bring the various “discoveries” that might be demonstrated into a coherent focus. This implies, among other things, that teachers of statistics must have much more knowledge and experience than many now possess, knowledge of what statistical inference really is and experience in handling data. Here is another place where schools and colleges would find the cooperation of business and industry invaluable. A good industrial program in quality management education could supply teachers with a wealth of information on both content and pedagogy.

It is true that students learn to value what is assessed. Thus, multiple assessments (tests, quizzes, short reports, term projects, computer lab assignments) should be used and the big picture of statistics should be emphasized over the concern about small details. I would give a high grade for a project that talked correctly about randomization in a comparative experiment and about transforming the data because the distribution looked skewed, but had a mistake in counting degrees of freedom. Again, teachers need to understand what the big ideas are.

## **Technology**

In synergy with content and pedagogy, technology does indeed play a key role in modern statistical education. We could not attempt modern data analysis without it. In fact, I think the advent of technology and the exploratory techniques that technology helped advance are the main reasons we have achieved some success in getting statistics into the secondary school curriculum in the US. Students in high school or college can now do the same statistical analysis that a quality engineer may be doing on the job.

In the content section I mentioned that some introductory statistics courses are going in the direction of too much data exploration and too little inference. That is because the data exploration is now fun and quick, and produces exciting pictures that can sometimes be changed dynamically. Inference is still a black box, whether done by hand or on the computer. In fact, black-box inference can be even more of a problem now that a computer or graphing calculator can automatically fit a variety of models to data in almost no time. What is needed to balance this out is to develop ways to have students discover concepts of inference by using technology. Straightforward simulations, as in the case of the central limit theorem, help. But, other concepts such as bias, randomization, confidence, significance, and power are not so easily shown. If that is attempted, the student tends to get lost in a maze of computer commands. Concept development software, as opposed to data analysis software, is beginning to appear; this could be a major improvement in the synergy between technology, concepts and pedagogy. The beat goes on!

## Discussion: Leading Horses to Water

**Judith M. Tanur**

*Distinguished Teaching Professor, State University of New York at Stony Brook, USA*

I am in full agreement with Moore's general thrust; indeed, I have been a minor participant in the movement to reform the teaching of statistics, having served as a member of the ASA/MAA Joint Committee that developed the recommendations shown in Moore's Figure 2 and as an instructor for workshops originated by that committee. For years I have expressed my teaching philosophy in such catch phrases as "statistics is not spectator sport" and "passive learning is an oxymoron" and have implemented that philosophy in the classroom by employing many of the variety of activities that Moore mentions and keeping the content of the course on a practical—rather than a mathematical level. Indeed, such a level is mandatory, for I teach the required introductory statistics course in the sociology department and sociology is a major that is often selected by students who prefer to have as little contact with mathematics as possible.

Although, as a sociologist, I believe that some of the things we take as "facts"—for example, race—are socially constructed, I agree with Moore that there are hard and fast facts of mathematics. In that sense I am neither a radical constructivist nor even a social constructivist. But we must remember that statistics is not mathematics, and when Moore requires that his students in their problem solutions "state a conclusion in the context from which the data come" (as I do, too) we are (quite properly, I believe) bringing in opinion, subjectivity, and social construction. Indeed, one of the goals of teaching statistics according to the recommendations of the ASA/MAA Joint Committee on the Curriculum is to teach students that there is rarely a single "right" answer to a statistical problem.

I have been teaching this course for some 30 years, through the student rebellions of the 1960's, through the movement from slide rules to mechanical calculators, to electronic calculators, to batch programming, to DOS-based PC programs, to menu-driven present-day statistical software. Starting out in the classroom, like so many college and university teachers, without any training in pedagogy, I have learned by doing. By trial and error and at enormous cost to generations of students, I have learned to teach. I even teach our department's teaching practicum on a regular basis. (In the practicum I regularly echo Moore's point that it is a snare and a delusion to teach in the way that "worked for us," because faculty—and the graduate students in the teaching practicum—are a self-selected bunch and not at all representative of the learning styles of undergraduates in general.)

I present my credentials in this rather long-winded and certainly immodest way in order to make a point. Recently something has gone very badly wrong with my teaching and/or with my students' learning.

As I write this I am slowly recovering from what I call "the semester from hell". Let me make clear what I was doing. I was, as usual, teaching introductory statistics, a required course for sociology majors. I was using David Moore's fine book *The Basic Practice of Statistics* (1995) and the companion workbook for Minitab, the statistical package that is implemented in our statistical laboratory. The class met twice a week for an hour in the lab and then for an hour and twenty

minutes in a classroom. I decided at the outset not to insist that students work in groups because I have found that procedure very difficult to implement at Stony Brook with unselected students. But I encouraged working together, and several of the brighter students took advantage of the option. There was ample time in the laboratory to explore the capabilities of the software and to use it to carry out data descriptions, simulations, and analyses. The students had access to the lab outside of class time and received regular assignments and frequent feedback. I followed the curriculum in Moore's book rather closely. That implies that very little stress was placed on probability and a good deal of emphasis fell on ideas of data generation, randomness, and exploratory data analysis. I put somewhat more stress than does Moore on cross-tabulations, their understanding and interpretation and their use in uncovering lurking variables, for this is the bread and butter of sociological analysis.

I am accustomed to students being frightened of statistics and for many of them to find it difficult. I have had classes that objected to what they saw as the rigorous requirements of the course. (Recall that these are students who are extremely leery of "math". As an example of the application of the concrete in the classroom, as Moore urges, when I ask these students to percentage a table, many are at a loss about how to calculate a percentage. I remind them that they know what they save when they go to purchase an article at a 10% discount. When I ask them introspect on the process they used to give me that answer, they usually respond that 10% is an "easy" number—they're mechanically moving a decimal point. So I give them a "harder" number—20%. They come up with an answer, and on introspection can come up with the process they used to get that answer, and thus realize that they know the algorithm. Then we tackle the more difficult reverse case—if you saved \$12 what percent discount did you get? The point is that these students are not innumerate—they are able to use these calculations in their everyday lives, but by and large they are not facile with arithmetic, let alone algebra or statistics and they suffer symbol shock easily.)

But what I was faced with this last semester is something different from lack of math background or discomfort with calculation. It was massive disengagement. I have never taught a class in which at a normal class meeting some 50% of the registered students were absent (usually terror—I fondly believe of the material, not of me—keeps attendance high). And I have never taught a class in which whenever I asked a question, I was confronted with a roomful of blank looks or, worse, of students who refused to meet my eyes. Nor have I ever taught a class in which few of the students handed in the homework assignments (although they counted heavily towards their course grade) and fewer still handed them in on time.

What did I do? First I tried all the usual tricks—when I asked a question of the class I resisted the temptation to answer it myself, and let the silence drag on seemingly interminably. Sometimes I got an answer, if only a partial one. I put more energy into my own interactions with the class, hoping to make up for the students' lack of energy and to model more active behavior—to the point that I began to feel like a high school cheerleader going through routines. Suspecting that the problem was lack of before-class preparation, I announced that I would give pop quizzes to check that the students were doing the reading on time. I gave the quizzes intermittently and they showed clearly that the students were not doing the reading, or at least not reading carefully. (And fear of the quizzes did nothing to improve either the preparedness of the students or their attitude towards the course and towards me.) I gave the students more time in the laboratory to work out problems for themselves before pooling class solutions. And finally, when it was clear that nobody in the class was following what was going on after the umpteenth simulation of a sampling distribution, I went back three chapters and we went over all the material yet again. Nothing worked!

What went wrong? The short answer is that I don't really know. Student evaluations were, as is so often the case, equivocal. Some liked the variety of examples while others wanted more concentration on sociology. Some felt the class went too quickly while others found it too slow. Performance on the final exam suggests that while a few students did achieve a mastery of the material, the vast majority did not.

Students rarely came to my office hours, but they did come with some frequency to those of my teaching assistant, and their interaction with her offers what I consider a clue. She reports that the students basically wanted to know “which formula to apply when”. I was reminded of a story told by my colleague in the Applied Mathematics and Statistics Department at Stony Brook, Alan Tucker, who has been active in Calculus Reform. He reports that one student was overheard confiding in another “I liked the old book better—you could get the right answer without having to understand what was going on”. My students also had no desire to understand what was going on—they just wanted to know how to get the right answer. I have no doubt that active learning is the only way to understand statistics, and probably any other subject. The question I raise is how do we motivate our students to learn actively when they consider it easier to be spoon-fed. (Note that in that last sentence I have generalized the individual problem I have been describing to one that I assume others face. Surely it would be more modest to assert that the problem is a failure of my teaching alone, and perhaps it is. But conversations with other professors in the social sciences at Stony Brook suggest that it is widespread in quantitatively oriented courses in those departments. Perhaps the problem is broader than those departments or than our campus and I shall address it below as if it were.)

My guess is that the phenomenon I’ve been describing is, in part at least, a consequence of the democratization of higher education that Moore describes. Many of the students at a state university like Stony Brook have to work (and many of them have to work full time) in order to pay tuition and support themselves. And in order to receive low cost student loans, they have to maintain a full course load. Thus employment pressures, together with the more traditional diversions of college life, leave them little time or energy for the business of the education itself. Surely students have had to work their way through college for many generations. But at least the popular image visualizes such students in earlier generations as being highly motivated. Moore tells us that over 60% of high school graduates in the United States now go on to some kind of post-secondary education, and we know that a Bachelor’s degree has become an entry-level credential for practically all white collar—and even pink collar—jobs. So while our students are motivated to get the degrees we offer them, they do not all seem to be strongly motivated to learn the content we attempt to teach them. (The “gentleman’s C” is surely a long-standing institution. What seems to me to be new is the widespread lethargy I have been describing.) And we must remember that lethargy is contagious—it is easy for a class to develop a norm of non-participation and socially ostracize those who violate that norm. Such social pressures in the classroom are subtle but powerful, and very difficult to eradicate once established.

What is to be done? I am sorely tempted to pass the buck and argue that it is our methods of primary and secondary education that need to be changed. If students came to college more accustomed than they currently are to active learning and critical thinking, they would not have developed a taste for spoon-feeding, and our motivational problems at the level of higher education would be mitigated. But we need to work with the students we are receiving at the university level, so I am rather at a loss.

My only speculation goes as follows: in general I agree with Moore that lectures are necessary for offering the students the “big picture”, but perhaps my teaching suffers from a tendency to outline that picture before students feel the need for such synthesis. Perhaps I must let them—indeed force them—to generate more pieces of the puzzle before I show them how to fit the pieces together. In practice, I think this may mean that a course centered around projects that students carry out from start to finish would provide the necessary motivation. Students who had their own data to make sense out of ought to be motivated to seek tools for that purpose, and even earlier, students faced with the problem of gathering their own data would inevitably bump into the real problems of data collection. While I have taught courses in which students did projects, I have never yet based an entire course on such an activity. I intend to try it the next time I teach introductory statistics.

And keep my fingers crossed!

# Response

**David S. Moore**

It is a pleasure to read the comments of so distinguished a group. The discussants' remarks offer a thoughtful set of expansions, emendations, and mild alternatives to my suggestions. Taken together, they mark out general boundaries for contemporary teaching of statistics to beginners. Rather than respond to individual comments, I want to conclude by emphasizing three wider issues suggested by the discussion.

*We (mostly) agree!*

First, *there is a striking degree of general agreement* on the overall nature of beginning instruction that we find desirable. Dick Scheaffer, whom I am willing to take as an authority, says that "With regard to the content of an introductory statistics course, statisticians are in closer agreement today than at any previous time in my career". *This general agreement is, I think, the most important message that readers of the article and the discussion should carry away.* Teachers of statistics should note that we have reached as clear a consensus as could be hoped on the guiding principles for teaching our discipline. The discussants are hardly an idiosyncratic group—they represent several nations, include both academics and non-academics, and among the academics list educational psychology and sociology as well as statistics as disciplinary affiliations. That their comments can be described as variations on common themes is revelatory.

I am often reminded, particularly when speaking with mathematicians, how unusual our degree of consensus is. For example, the U.S. National Council of Teachers of Mathematics has asked professional societies in the mathematical sciences to comment on its *Standards* (NCTM, 1989) in preparation for their revision. The NCTM standards generally apply to school mathematics the kind of thinking that undergirds our consensus about statistics instruction. Comments from the American Mathematical Society reflect the deep division among mathematicians concerning the "reform" approach to teaching. The AMS group could not even agree that students should not be taught to extract square roots by hand!

*Our agreement may have limited impact*

The relative consensus among statisticians who teach statistics, and the lack of similar consensus elsewhere, leads to my second summary comment: *To achieve widespread improvement in the teaching of beginning statistics, statisticians must find a way to cooperate more effectively with teachers who are not statisticians.* Dick Scheaffer has pointed to the need for a unified effort across the mathematical sciences. The need is broader, and applied scientists may be more cooperative than mathematicians.

Statistics may be unique among academic disciplines in that it is most often taught by faculty who lack advance training in the discipline. Behavioral and social scientists, mathematicians, and professors of business or engineering bring a variety of assets to the teaching of statistics, but are nonetheless a collective exception to the usual academic standard that instruction should at least be

designed and directed by scholars specializing in the discipline taught. Statistics is also a staple of in-house training by employers, a sector that in the United States is comparable in number of students taught and dollars expended to the entire higher education system. Here again the training may not be designed and directed by statisticians.

I am willing to assert (with, alas, only extensive anecdotal evidence) that non-statisticians are more likely to retain traditional content that focuses almost exclusively on formal inference and to present this content in a quite formal way that does not show beginners how statistics can shed light on data and help data shed light on reality. Mathematicians are schooled to think this way. Other teachers, in my experience, are often in the grip of the professional's fallacy—they teach statistics as a research tool for their field, though few of their students will become researchers. The common course title “research methods” gives the game away. The changes in the discipline of statistics that inform our consensus about beginning instruction percolate only gradually into the consciousness of teachers who read primarily literature in business or engineering or mathematics.

It is easy to say what is needed in the academic arena: a breaking down of the barriers that separate disciplines, so that team teaching and other “obvious” cooperative approaches can flourish. This will happen slowly at best. It is also easy to see that a field that is widely viewed as one variety of “research methods” has an uninteresting and probably short future. Propagating a broader view of the science of data is essential to our survival as a discipline. Having no short remedies to offer, let me urge statisticians to look around them and (to paraphrase Tom Lehrer), “Don't shield your eyes—evangelize”.

### *We are still too narrow*

Grumbling about stereotyping statistics as “research methods” leads naturally to my final comment: *Our reformed introductory instruction is much improved, but remains too narrow*. Our consensus on content can be briefly stated as “more data, more design, more concepts”. Probability-based inference finds its place in a broader framework, with a degree of emphasis that varies with the audience. I suggested that we consider such additional topics as data ethics and national statistics. The non-academic discussants, Jon Kettenring and the General Electric group, have pointed to more general and more important ways in which we might broaden the content of our first courses. Anne Hawkins' development of the statistical literacy theme points in the same direction. I hope that academic teachers of statistics will not be drawn by the professional's fallacy into quick rejection of the “big picture” ideas emphasized by the non-academic discussants. I accept them as suggestions that I should have made.

In fact, industrial statisticians have for some time been urging that the role of even a professional statistician must extend considerably beyond that of an expert on research methods. Hunter (1981) and Marquardt (1987) are two classics in this hortatory tradition. We academics have responded—that “one-and-only” statistics course looks quite different now than it did 15 years ago. It is fair to ask Roger Hoerl, Gerry Hahn, and Necip Doganaksoy if their technical and managerial colleagues gained so negative an impression of statistics from a course taught by statisticians. But we must also, in the spirit of continuous improvement, constantly reconsider the needs of our students and client disciplines. For almost all students, this is indeed their “one and only” statistics course. Just as we should steel ourselves to teach no probability not required to understand the statistical content, so we should refuse to teach statistical topics needed only as gateways to further courses in statistics. We should also continue to reconsider what happens in our classrooms, perhaps pulled along by Joan Garfield's comments.

### *What of the future*

Our relative consensus is of course not static, if only because our discipline continues to evolve. The discussants hint at good reasons for some pessimism. Judy Tanur observes deterioration in students' motivation, study habits, and attention spans, no doubt consequences of the degenerate youth culture that we Americans are busily exporting to the rest of the world. Jayanta Ghosh reminds us that many skills will become redundant, and Hoerl *et al.* wonder if statistics will become a branch of computer science. Statistics education certainly faces many challenges. Can research into teaching and learning help us interact with Judy's new class of student customers? Will we content experts listen to such advice? Can we broaden the vision that informs our beginning instruction enough to change the prevailing perception of the value of our discipline? Can we cooperate effectively with other disciplines and with the non-academic employers of our graduates?

In the end, the future of statistics education is tied to that of statistics as a field of intellectual endeavor. The compelling reason to continue to broaden our beginning instruction is that our field must broaden or perish in our rapidly-changing, globally-integrated, information-driven (choose your adjectives) environment. Rapidly changing environments bring about maladaptation and some combination of extinction and evolution. Change may sweep "statistics" away in a wave of "creative destruction", that metaphor of capitalism triumphant. That is, there may be no "statisticians" a generation from now, but more Information Services experts than there are statisticians now. Look at the courses taught in a top engineering school in 1920 to see what wholesale destruction has occurred, yet there are more engineers than ever. Is "statistics" more analogous to "electric trolley engineering" than to "electrical engineering?" What and how we teach will help determine the answer.

### **Additional references**

- Box, G.E.P., Hunter, W.G. & Hunter, J.S. (1978). *Statistics For Experimenters*, John Wiley & Sons, New York.
- Deming, W.E. (1953). On the Distinction Between Enumerative and Analytic Surveys. *Journal of the American Statistical Association*, **48**, 244–255.
- Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. D. Reidel Publishing Co. Dordrecht. Tr. C.A. Sheppard.
- Garfield, J. (1995). Respondent: "How should we be teaching statistics?" *The American Statistician*, **49** (1), 18–20.
- Hart, K. (1997, to appear). What responsibility do researchers have to mathematics teachers and children? *Proceedings, ICME-8, July 1996*.
- Hawkins, A. (1996a). *Common Variance*. Presented to the East Midlands Local Group Meeting of the Royal Statistical Society, March 1996.
- Hawkins, A. (1996b). *Can a Mathematically-Educated Person Be Statistically Illiterate?* Presented at the Nuffield Conference 'Mathematics for the New Millennium: What Needs Changing and Why?' October 1996.
- Hoerl, R.W. & Snee, R.D. (1995). Redesigning the Introductory Statistics Course. *Technical Report No. 130*, Center for Quality and Productivity Improvement, Madison, Wisconsin.
- Hogg, R. (1992). Report of Workshop on Statistics Education, in Heeding the Call for Change, *MAA Notes*, 34–43.
- Hunter, W.G. (1977). Some Ideas about Teaching Design of Experiments with 2\*\*5 Examples of Experiments Conducted by Students. *The American Statistician*, **31**, 12–17.
- Hunter, W.G. (1981). The practice of statistics: the real world is an idea whose time has come. *American Statistician*, **35**, 72–76.
- Lucas, J.M. (1996). *System Change and Improvement: Guidelines for Action When the System Resists*, presented at the 40th Fall Technical Conference, October 24th, Scottsdale, Arizona.
- Marquardt, D.W. (1987). The importance of statisticians. *Journal of the American Statistical Association*, **82**, 1–7.
- Moore, D.S. (1995). *The Basic Practice of Statistics*, New York: Freeman.
- Moore, P.G. (1990). The Skills Challenge of the Nineties. *Journal of the Royal Statistical Society, Series A* **153**(3): 265–85.
- Piaget, J. & Inhelder B. (1951) (trans 1975 Lowell Leake Jr., Paul Burrell, Harold D. Fischbein) *La genèse de l'idée de hasard chez l'enfant (The origin of the idea of chance in children)*. Routledge & Kegan Paul.
- Scheaffer, R., Watkins, A. & Landwehr, J. (1997). What every high school graduate should know about statistics. To appear in Lajoie, S.P. (Ed.). *Reflections on Statistics: Agendas for Learning, Teaching, and Assessment in K-12*. Hillsdale, NJ: Erlbaum.
- Snell, J.L. (1996). The Internet: A New Dimension to Teaching Statistics. *The American Statistician*.
- Tanur, J.M., Mosteller, F., Kruskal, W.H., Pieters, R. S. & Rising, G.R., eds. (1989). *Statistics: A Guide to the Unknown*, 3rd edition, Wadsworth, Pacific Grove, California.

- Tucker, A. & Leitzel, J. (eds.) (1995). *Assessing Calculus Reform Efforts*. Washington, D.C.: Mathematical Association of America.
- Wackerly, D. (1996). *ExplorStat Software*, University of Florida. (NSF-funded project).

[Received September 1996, accepted February 1997]