New Pedagogy and New Content: The Case of Statistics

David S. Moore
Department of Statistics, Purdue University, West Lafayette, IN 47907 USA

International Statistical Review, 65 (1997), 123–165.

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 and Ahlgren (1988), Kapadia and 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 poured in by teachers; they inevitably construct their own knowledge by combining their present experiences with their existing conceptions.

Figure 1 near here

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 and 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 expouse 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?

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.

Figure 2 near here

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 and 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 and 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 and 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, and Lichtenstein (1982) and Zeckhauser and 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.

Figure 3 near here

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 $\Longleftrightarrow$ Hands-on work

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

Content $\Longleftrightarrow$ *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?"

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 $\Longleftrightarrow$ *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 $\Longleftrightarrow$ *Content*: New content in statistics reflects the computing-intensive nature of statistical practice. Exploratory data analysis is characterized (Velleman and 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 telescopy 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, and Swift (1986); Landwehr and Watkins (1986); Landwehr, Watkins, and Swift (1987), and Newman, Obremski, and Schaeffer (1986).)

Figure 4 near here

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 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 $\overline{x}$ 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 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 and 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 near here

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

1. Biehler, R. (1993). Software tools and mathematics education: The case of statistics. In *Learning from Computers: Mathematics Education and Technology*, Ed. C. Keitel and K. Ruthven, pp. 68–100. Berlin: Springer.

2. 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.

3. Bullock, J. O. (1994). Literacy in the language of mathematics. *American Mathematical Monthly* **101**, pp. 735–743.

4. 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.

5. Cobb, P., Yackel, E., and Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education* **23**, pp. 2–33.

6. Garfield, J. (1994). Beyond testing and grading: Using assessment to improve student learning. *Journal of Statistics Education* (electronic journal) **2**, no. 1.

7. Garfield, J. (1995). How students learn statistics. *International Statistical Review* **63**, pp. 25–34.

8. Garfield, J. and Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education* **19**, pp. 44–63.

9. von Glasersfeld, E. (1990). An exposition of constructivism: why some like it radical. In *Constructivist Views on the Teaching and Learning of Mathematics*, Ed. R. B. Davis, C. A. Maher, and N. Noddings. Reston, Va.: National Council of Teachers of Mathematics.

10. Gnanadesikan, M., Schaeffer, R. and Swift, J. (1986). *The Art and Techniques of Simulation.* Palo Alto, Calif.: Dale Seymour.

11. Kapadia, R. and Borovcnik, M. (Eds.) (1991). *Chance Encounters: Probability in Education.* Dordrecht: Kluwer.

12. Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society* **B21**, pp. 272–319.

13. Krantz, S. G. (1993). *How to Teach Mathematics: A Personal Perspective.* Providence, R.I.: American Mathematical Society.

14. Landwehr, J. and Watkins, A. (1986). *Exploring Data.* Palo Alto, Calif.: Dale Seymour.

15. Landwehr, J., Watkins, A. and Swift, J. (1987). *Exploring Surveys and Information From Samples.* Palo Alto, Calif.: Dale Seymour.

16. Laridon, P. (1996). Graphics calculators in South Africa. Paper presented at the Eighth International Conference on Mathematics Education, Seville, Spain.

17. Moore, D. S. (1992). Teaching statistics as a respectable subject. In *Statistics For the Twenty-First Century*, Ed. F. S. Gordon and S. P. Gordon, pp. 14–25. Washington, D.C.: Mathematical Association of America.

18. Moore, D. S. (1993). The place of video in new styles of teaching and learning statistics. *American Statistician* **47**, pp. 172–176.

19. Moore, D. S. (1997). *Statistics: Concepts and Controversies*, 4th ed. New York: Freeman.

20. National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards For School Mathematics.* Reston, Va.: NCTM.

21. National Research Council (1990). *Reshaping School Mathematics: A Philosophy and Framework for Curriculum.* Washington, D. C.: National Academy Press.

22. National Research Council (1991). *Moving Beyond Myths: Revitalizing Undergraduate Mathematics.* Washington, D. C.: National Academy Press.

23. National Research Council (1993). *Measuring What Counts: A Conceptual Guide for Mathematical Assessment*. Washington, D. C.: National Academy Press.

24. Newman, C., Obremski, T. and Schaeffer, R. (1986). *Exploring Probability*. Palo Alto, Calif.: Dale Seymour.

25. Pearl, D. K. and Stasny, E. A. (1992). *Experiments in Statistical Concepts*. Dubuque, Iowa: Kendall/Hunt.

26. Rossman, A. J. (1996). *Workshop Statistics: Discovery With Data*. New York: Springer.

27. Scheaffer R. L., Gnanadesikan, M., Watkins, A., and Witmer, J. (1996). *Activity-Based Statistics: Student Guide*. New York: Springer.

28. 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.

29. Steen, L. A. (Ed.) (1988). *Calculus for A New Century*. Washington, D. C.: Mathematical Association of America.

30. Slovic, P., Fischhoff, B., and Lichtenstein, S. (1982). Facts vs. fears: Understanding perceived risk. In *Judgment under Uncertainty: Heuristics and Biases*, Ed. D. Kahneman, P. Slovic, and A. Tversky, pp. 463–489. New York: Cambridge University Press.

31. Spurrier, J. D., Edwards, D. G. and Thombs, L. A. (1995). *Elementary Statistics Laboratory Manual*. Belmont, Mass.: Wadsworth.

32. Tanner, M. A. (1990). *Investigations for a Course in Statistics*. New York: Macmillan.

33. Tversky, A. and Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin* **76**, pp. 105–110.

34. Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* **90**, pp. 293–315.

35. University of North Carolina (1997a). *Teach-Stat Activities: Statistics Investigations for Grades 1–3*. Palo Alto, Calif.: Dale Seymour.

36. University of North Carolina (1997b). *Teach-Stat Activities: Statistics Investigations for Grades 3–6*. Palo Alto, Calif.: Dale Seymour.

37. Velleman, P. F. and Hoaglin, D. C. (1992). Data analysis. In *Perspectives on Contemporary Statistics*, Ed. D. C. Hoaglin and D. S. Moore, pp. 19–39. Washington, D.C.: Mathematical Association of America.

38. Velleman, P. F. and Moore, D. S. (1996). Multimedia for teaching statistics: promises and pitfalls. *American Statistician* **50**, to appear.

39. Vere-Jones, D. (1995). The coming of age of statistical education. *International Statistical Review* **63**, pp. 3–23.

40. Zeckhauser, R. J. and Viscusi, W. K. (1990). Risk within reason. *Science* **248**, pp. 559–564.

Figure 1: The Reform of Pedagogy

- **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 2: Recommendations of the ASA/MAA Joint Curriculum Committee

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 3: Synergy in Statistical Education

- **Content $\Longleftrightarrow$ Pedagogy**

  Data analysis $\Longleftrightarrow$ Hands-on work
  Statistics in practice $\Longleftrightarrow$ Communicate, cooperate
  More concepts $\Longleftrightarrow$ Less proof

- **Pedagogy $\Longleftrightarrow$ Technology**

  Visualization $\Longleftrightarrow$ Automate graphics
  Problem-solving $\Longleftrightarrow$ Automate calculations
  Active learning $\Longleftrightarrow$ Multimedia

- **Technology $\Longleftrightarrow$ Content**

  Computing $\Longleftrightarrow$ Data analysis, diagnostics, bootstrap, ...
  Automation $\Longleftrightarrow$ More concepts
  Simulation $\Longleftrightarrow$ Less proof

Figure 4: Primary School Synergy

- **Whole-number counting $\Longleftrightarrow$ Bar graph**

  "How often do you wear a hat?"
  Bars of stick figures count responses: Often, Sometimes, Seldom, Never.

- **Number line $\Longleftrightarrow$ Dot plot**

  "What is the date on your coin?"
  Interesting left-skewed distribution.

- **Place notation $\Longleftrightarrow$ Stem-and-leaf plot**

  "How many pages does your favorite book have?"

- **Betweenness and grouping $\Longleftrightarrow$ Histogram**

  "How long does it take you to get to school?"

Figure 5: Multimedia is . . .

Text
Sound
Still images
Full-motion video
Cartoon-style animation
Dynamic computer graphics
Computing for calculation and graphics

IN ONE SYSTEM

Learner interacts with keyboard and mouse