Comment: Technology and the Future of Statistics Education (Cobb)

I congratulate the authors of all three articles. All three are timely, all are thoughtful, all are innovative, all are potentially influential. It was a pleasure to read and learn from each, and to enjoy their provocation to think about where technology may be leading our profession.

1. A Framework

As I first read and then repeatedly re-read the three articles I found myself led to a framework for thinking not only about the articles themselves, but also about what they collectively suggest for where computers and the internet may be taking statistics education. Before offering specific comments about the articles I summarize the four-part frame that I came to find useful:

innovation and convention,
curriculum and pedagogy,
differing roles for technology, and
differing environments for teaching and learning.

Innovation and convention. All innovators face a trade-off, one that can be phrased as a question: "Do you want to move a smaller number of people a larger distance in your proposed new direction, or do you want to move a larger number of people a smaller distance in that new direction?" More briefly, "Are you a path-breaker for the few, or a path-smoother for the many?" Progress depends on having innovators along the entire continuum, but just as with Lincoln's famous comment about fooling people, there is an ineluctable negative correlation: you can perhaps move some of the people a long way, or perhaps move almost all of them a short way, but you can't move almost all of them a long way. Being radical in one direction may require being conventional in other directions. In my comments, I try to identify what I consider most radical about each contribution. When I describe other features as conventional or mainstream, I do not mean it as a criticism, but as part of the essential trade-off that challenges all innovators.

Curriculum and pedagogy: a 3x3 table. To the best of my knowledge it was David Moore (1997) who first called the attention of statistics educators to curriculum versus pedagogy as a useful model for thinking about how we use technology for what we do in class. Of course, as Moore knew, the two categories are neither mutually exclusive nor independent, but many years later his model remains useful.

For the purposes of my comments here, I extend his model to a 3x3 table. As a very crude simplification, we can classify a curriculum as retrograde, or mainstream, or innovative. Similarly, a pedagogical approach can be retrograde, mainstream, or innovative. In principle, all nine cells of the resulting 3x3 table are possible. At one extreme you or I could use cutting-edge pedagogy to present a dated curriculum; at a different extreme a talking-head wizard behind a lectern could nevertheless present truly innovative content.

Of course this 3x3 model is cross-sectional. Over time what was once innovative might become mainstream, what was once mainstream might become retrograde. (The implied Markov model suggests a possible transition to an additional, absorbing state: well-deserved oblivion. The demise of unhealthy recipe books larded with synthetic data exemplifies one such terminal transition we needn't mourn.)

The distinction between pedagogy and curriculum can be useful in a variety of contexts related to teaching and learning. In particular, as Moore suggested, it can help us in recognizing and evaluating possible uses of technology

The differing roles for technology.

Meant entirely as a high complement, I note that the three featured articles remind me of Tolstoy's unhappy families at the start of *Anna Karenina* – each uses technology in its own way. Taken together the three articles are even more valuable than the sum of their individual contributions because they illustrate a variety of uses of technology and provoke us to think about future uses as well.

It is an irony of history that the use of computers for teaching introductory statistics was becoming mainstream just at the time that John Tukey's (1977) *Exploratory Data Analysis* (EDA) was making it possible to analyze real data *without* using a computer. EDA began as a deliberate *substitute* for technology, a back-of-the-envelope approach that allowed us to teach with real data using just pencil and paper. Tukey's once-radical use of stemplots to find rank-based summaries has become mainstream, but his equally-radical break tables, which allowed transformation by hand to roots, logs, and reciprocals, were made obsolete when computers became easily available. In another irony, almost as a serendipitous afterthought, computer simulation helped Tukey's simple hand tools morph into the basis for modern robust methods (Andrews, *et al*. 1972).

Tukey's EDA, once radical, is now mainstream. The same is true of another pioneering effort, Francis Anscombe's (1981) use of interactive computing for statistical analysis. However, whereas Tukey's EDA was a conscious attempt to avoid technology, Anscombe's book was based on a deliberate and prescient vision of where computing might lead our profession. APL, the programming language Anscombe used, may have fallen by the wayside, but Anscombe's interactive approach to data analysis deserves to be regarded as a forerunner of diagnostics, graphics, and sensitivity analysis.

Anscombe's book is no longer well known, perhaps mainly because APL has been displaced by languages like S-plus and R. For readers unfamiliar with the book I offer a brief summary. The core of the book is based on a statistical analysis of just three data sets. The three chapters on the three data sets occupy the central 200 pages of the book – more than 60 printed pages per data set. Each analysis illustrates the interplay between two approaches, "arithmetic" and "theoretical" in Anscombe's words, "exploratory and "confirmatory" in Tukey's. Each analysis begins with exploration and description, then fits a variety of models and follows with post-fit diagnostics. In short, Anscombe's book was a pioneer in combining interactive computing, exploration, and multiple reanalyses, including robust methods and diagnostics. Computing was what enabled these anticipations of contemporary practice. In this sense, Anscombe's book is a prime example of technology *as* statistics: *technology changed practice*.

When it comes to teaching and learning, and in my discussion of the three articles, I find it useful to distinguish among four uses of technology: (1) for distribution of resources, (2) for pedagogy, (3) *for* curricular content, and (4) *as* curricular content. To illustrate the distinction between the last two, consider Anscombe's book. His book has certainly influenced the *teaching* of statistics, but I think of it as intended mainly to influence the *practice* of statistics. Thus I regard it as an instance not of technology *for* curriculum, because the curriculum could not exist or be taught without the technology; rather Anscombe's book is an instance of technology *as* curriculum. The technology is an integral part of the content itself.

Differing environments for teaching and learning. As statisticians we know that things vary. In particular, those of us who teach contend with a variety of environments, whose convex hull includes all of the following:

- Large lectures at state universities, with recitation sections and grading done by graduate teaching assistants.
- Classes of 15-35 at liberal arts colleges with teaching loads of 4-6 courses per year, perhaps with student help for marking homework or help outside of class.
- Classes of 30, plus or minus, at two-year colleges with teaching loads of 5 courses per semester, and no student assistants.
- Classes of individual distance-learners we never meet face-to-face, except perhaps virtually.

The message from this variation is clear. "One size fits all" is doomed to fail. With that conclusion as a major premise, I now turn to celebrating the variety of the articles themselves.

2. Webster West's Textbook 2.0

This article describes an on-line book with search capacity and hyperlinks to audio, video, applets, and statistical software, all integrated with course management software and social media. The curriculum is mainstream but modern, chosen to have broad appeal. Given the emphasis on technology for pedagogy and delivery, the choice of a conventional curriculum is appropriate. Moreover, there is in principle no reason why the West model could not be adapted to fit a curriculum that is farther from the mainstream. The approach is innovative in the richness of technologies used, and in the extent of their integration. Electronic distribution makes it possible to offer the on-line book to students at roughly half the cost of a printed book.

West's project, unlike those of the other two articles, is a collaborative effort with a commercial publisher. The publisher is involved with development as well as marketing, and distribution. Often, as with the other two projects, development support comes from a grant from the government or non-profit foundation. For now, working with a commercial publisher strikes me as eminently reasonable. At the same time, I worry that longer term, technology will alter the current model for commercial publishing in ways and to an extent that is hard to predict. I return to this worry later on in what follows.

Like a traditional textbook, West's package is flexible in its potential for use in a variety of learning environments, from small classes to large classes, advanced placement high school courses to lecture classes at universities, with varying degrees of student independence. West's choice of a mainstream curriculum should make it comparatively easy for teachers to adopt his approach. The main effort required of teachers will be to become comfortable with the variety of electronically-based options, and with helping their students to take full advantage.

It is easy to imagine West's project serving as a model for other, similar efforts. The flexibility, the choice of a mainstream curriculum, and the integration of a panoply of technologies, together with the advantage of marketing by a publisher all lead me to predict widespread use and commercial success, at least over the short term.

Longer term, however, I predict that we will come to see this as a needed but transitional model. In particular, for now a 50% saving over the cost of a traditional printed book may seem like a lot, but in the long run I think that options like this one are likely to face challenges and to be supplanted by other options that are far less expensive, such as the Open Learning Initiative at Carnegie Mellon University, *coursera* (Harvard, Stanford, MIT), and the one described by Cetinkaya-Rundel and colleagues.

3. Cetinkaya-Rundel, Diaz, Barr: OpenIntro Statistics (CDB)

These authors describe a textbook that is fully open-source. Their curriculum is mainstream, deliberately designed to follow the AP (College Board 2010) and GAISE (2012) guidelines. In their reliance on a traditional textbook and statistical software their pedagogy is also mainstream. Moreover, the open-source movement itself is not new. All the same, I consider the effort of these three authors to be deeply radical because of two features: their book is freely customizable and it is free. The curriculum and pedagogy, though mainstream, are mainstream for a reason – to appeal to a very broad base of users. There is nothing inherent in the model to keep it from being used with other curricula, and in principle one can imagine an open-source version of the technologically integrated model described by Webster West. What is radical is the model for distribution: the resource is available at no cost, with no restrictions on modification or use. This is potentially a model for a global revolution.

The fact that anyone is allowed to edit or customize the book raises questions about how to manage content. If I modify the book, will I be allowed to post my version as another open-source document? If I and others can do that, versions will proliferate. Should these multiple versions all be equally available *caveat emptor*, or should there be versions that are recognized by a committee designated for that purpose? The model of CAUSE (2013) and its website might be useful. Alternatively, professional organizations (NCTM, AMATYC, MAA, ASA, IASE) might post approved versions on their own websites. Then such sanctioned versions would gain both credibility and publicity. There are other questions of a similar sort, but it seems reasonable to think that experience will lead to acceptable answers and workable mechanisms.

A second set of questions related to zero cost strikes me as more important but harder to answer. This is where the greatest potential for revolutionary impact lies, perhaps as revolutionary as the impact of the printing press. We had books for millennia before we had the printing press, but Gutenberg radically changed the balance between fixed and marginal costs of producing books. The fixed cost of machinery and typesetting was high, but compared to the cost of copying by hand, the marginal cost of an additional printed book was negligible. Plummeting marginal costs led to an epidemic spread of books, a rise in literacy rates, and a challenge to the existing order, a challenge we now know as the Reformation, made possible because technology had undermined a monopoly on reading the Bible. Hierarchy loosened.

In much the same spirit, we had printed books for half a millennium before we had the world-wide web, and now, once again, technology is radically changing the fixed and marginal costs of producing and distributing books. For a second time, a drastic reduction in marginal costs will challenge an existing order. It seems plausible, perhaps even likely, that once again hierarchy will loosen. Currently, book publishers are gatekeepers due to high fixed costs and marketing costs. Colleges also are gatekeepers due to high marginal costs of providing education to an additional student. Technology threatens to open both gates. Already books can be produced on a lap-top and offered for free via the web. Already anyone with internet access can take free courses such as those offered by Harvard, MIT, and Stanford through *coursera*, or through Carnegie Mellon's Open Learning Initiative.

As a thought experiment, imagine that you are in charge of hiring someone for a technical position: How much extra will your company pay for an applicant with US citizenship and a traditional degree from a residential college, as opposed to a non-degreed non-US-citizen with equal technical skills acquired on-line? How much value comes from attending face-to-face classes led by a teacher who knows you personally and cares about your learning? What is the value of time spent at a residential college? What will be the consequences for US high school students making choices about college? What will be the consequences for US institutions of higher learning? Longer-term and more broadly, what will be the consequences for wages of mid-level technical workers? … for income inequality? … for our political system?

4. Zieffler, Isaak, and Garfield: The Course and the Textbook (ZIG)

This article is unlike either of the other two. It is the only one to propose a radical curriculum, one based on substituting randomization-based methods for traditional methods that rely on asymptotic theory and the normal approximation. Its curriculum is also the most radical of the three in terms of its tightly knit reliance on cognitive science and a research-based understanding of how students learn the logic of inference (and why, sadly, they sometimes fail to learn).

The pedagogy is also radical, in that it is based at root level on model-eliciting activities. These are open-ended activities, designed and refined through research, and intended to spur exploration of an applied problem as a way to engage students in constructing their own understanding how statistics works.

Although both the curriculum and pedagogy are radical, they pose different challenges for the teacher, and they have different relationships to technology.

> (1) Technology. The randomization-based curriculum depends heavily on technology. Although the simulation-based methods in this project may be introduced with coins and cards, that tactile beginning is just a concrete lead-in to an essential technological follow-up. In theory, you can do enough simulations with coins and cards, but in practice, you need a computer. In contrast to curricular dependence on technology, the radical pedagogy via model-eliciting activities (MEAs) is much more independent of technology in that you can use MEAs in many situations where you don't need computers.
>
> (2) Challenges to the teacher. Learning to teach a randomization-based curriculum is mainly an intellectual challenge, one that requires learning a set of ideas that is very different from those in the usual course. The traditional, normal-centric curriculum is heavy on formulas and a large, diverse collection of exceptions, rules and adjustments that try to ensure that approximations work the way they are hoped to work. The randomization-based alternative has far fewer formulas but a critically important small set of general ideas. Making the transition from one curriculum to the other can present a major challenge, especially for those without a background in statistics.
>
> Learning to teach a course based on MEAs poses a challenge that is not so much intellectual as personal. Even though it make sense in theory to give up the control that lecturing brings in order to allow learning be more spontaneous, I can testify first-hand that Yogi Berra had it right: In theory, there may be no difference between theory and practice; in practice, there is.

All these features make the ZIG project very different from the other two. Alone among the three it takes advantage of technology to base a course on statistical content that could not be taught without technology, specifically the use of randomization-based alternatives to large-sample approximations based on the normal curve.[1] This makes it an instance of technology *for* content, and an example of what I regard as perhaps the most overlooked

---

[1] Although Webster West's applets make it easier for students to implement normal-based methods, their use does not change the statistical content from what we began to teach 30 years ago when data-driven courses became mainstream.

and most underexploited opportunity that computers have opened up to us as teachers of statistics. This kind of reliance on technology also makes the authors path breakers rather than path smoothers. The down side is that what they propose may prove to be too hard a challenge to appeal to more than a small number of teachers in the short run. The up side is that the challenge they pose is one that can provoke or inspire all of us: even those who consider their approach too hard to try out may still imagine what it could be like, and so to have their thinking and teaching changed in good ways.

5. Looking ahead

In conclusion, I reiterate my applause for what these authors have accomplished for our profession. At the same time, as I look to the future in the context of the three innovative efforts, my sense is also one of imbalance and opportunity: *Our use of technology for conveyance is far ahead of our use of technology for content*. By "conveyance" I mean both pedagogy (conveyance of ideas) and distribution (conveyance of resources). By "content" I mean curriculum, but more specifically, I have in mind an ambitious agenda for the future of statistics education. As I see it, despite our use of real data, of interactive graphics, and of diagnostics, on balance technology has *widened* the gap between statistical practice and what we teach in the first course, whether "first" refers to the first-year introductory course or the upper division course in mathematical statistics.

Technology has advanced statistical practice in many areas, including these five:

(1) Randomization and the bootstrap as alternatives to asymptotic methods based on a normal approximation;
(2) Multiple regression, multi-way ANOVA, and other models with multiple "predictors";
(3) Data mining, micro-arrays, and other methods for large data sets;
(4) Generalized linear models, and logistic regression in particular;
(5) Bayesian posteriors generally and hierarchical models in particular.

In my concluding paragraphs, I offer a few thoughts about my hope that *over the next decade technology will help integrate these topics into our introductory curricula*. I suggest that the technology is already there. What is required is to recognize and acknowledge an important fact: Our teaching of statistics is still largely shaped by a conviction that mathematics provides the only path to understanding. In fact, despite the immense value of a mathematical understanding, practitioners in our client disciplines are demonstrating that it is quite possible through technology to understand enough to do meaningful applied work even without much mathematical sophistication. *We need to find better ways to use technology to rethink the content of what we teach*. As I look to the future I am encouraged by a number of efforts to do just that:

Randomization and the bootstrap. Until recently, randomization-based methods were not taught at all in the elementary first course. Now, thanks to computer simulation, that is changing. See Hesterberg *et al*. (2010) for the current version of a pioneering singlechapter foray. For book-length treatments, existing examples include not only ZIG, but also Lock *et al*. (2012) and Tintle *et al*. (2013).

Multi-predictor models. Despite the leading example of Mosteller *et al*. (1983), until recently it was unusual for an elementary beginning course to do more than touch on multiple regression or analysis of variance (ANOVA). Now, thanks to technology, we can hope for change. As an important example, Daniel Kaplan (2009) at Macalester College teaches a first course that is largely multivariate. As an additional example, for several decades Mount Holyoke College has taught a course on design of experiments and multi-factor ANOVA, a course with neither mathematics nor statistics prerequisites (Cobb 1998).

Data mining, and other methods for large data sets. Until recently, these methods were considered beyond the reach of an elementary course, but this is now another area where computers are changing what we teach. For several years Richard De Veaux (2009) has been teaching data mining as a second statistics course at Williams College. More recently at Amherst College Amy Wagaman (2013) created and taught a first-year introductory course whose topic was applied multivariate methods, and which included methods for data mining. The seminal paper by Nolan and Temple-Lang (2010) addresses directly the role of computing in the statistics curriculum.

Logistic regression and other generalized linear models. These models and related methods have become a central part of statistical practice, but they are rarely taught in a first course, mainly because our exposition is still based largely on a mathematical understanding. Computers offer an alternative path: The computer does the work, and the context provides the meaning. Moore and McCabe (1998) broke new ground by including an optional chapter on logistic regression. Lisa Dierker (2012) and her colleagues at Wesleyan College have developed an elementary first course that makes this approach of computer-and-context (as an alternative to deductive derivation) the basis for an entire course.

Bayes. I've saved the biggest but perhaps most important challenge for last. Applied practice has become increasingly Bayesian over the last two decades, thanks mainly to Markov Chain Monte Carlo. Our teaching lags far behind. Why? Here, also, our exposition is held back by a conviction that a mathematical path to understanding is the only acceptable route. Specifically, this conviction makes an obstacle out of the marginal probability in the denominator of Bayes Theorem, which is typically a multiple integral. Computers can free us from the tyranny of the denominator: (1) The posterior is proportional to likelihood (Laplace's 1774 principle, see Hald, 1998, p. 160); (2) we get posteriors by simulation, and (3) the denominator is simply the number of replications. *Anyone who understands division can be a Bayesian*. Our first courses need to recognize and take advantage of this fact. We can (and probably should, and no doubt will) continue to argue about when and whether a Bayesian model is useful, but computers have mooted the argument that Bayes is too hard for a first course.

In conclusion, I offer a final salute to the authors, together with a hope that they and others who read their articles will continue to explore and exploit a truly vast set of exciting possibilities.

References

Andrews, D. F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H., and Tukey, J.W. (1972). *Robust Estimates of Location: Survey and Advances*. Princeton University Press.

Anscombe, F.J. (1981). *Computing in Statistical Science through APL*. New York: Springer-Verlag.

CAUSE (2013). *Consortium for the Advancement of Undergraduate Statistics Education.* Retrieved October 25, 2013 from https://www.CAUSEweb.org

Chihara, Laura and Hesterberg, Tim (2011). *Mathematical Statistics with Resampling and R*, New York: John Wiley and Sons.

Cobb, George W. (1998). *Introduction to Design and Analysis of Experiments.* New York: Springer-Verlag.

Cobb, George W. (2007). "The Introductory Statistics Course: A Ptolemaic Curriculum" *Technology Innovations in Statistics Education*, v.1, no. 1.

College Board (2010). *Statistics Course Description*. Retrieved October 27, 2013 from *http://apcentral.collegeboard.com/apc/public/repository/ap08_statistics_coursedesc.pdf*

De Veaux, Richard (2009). "Successful Data Mining in Practice," Retrieved October 27, 2013 from *http://www.jmp.com/about/events/explorers/pdf/deveaux_boston09.pdf.*

Dierker, L. Kaparakis, E., Rose, J., Selya, A. and Beveridge, D. (2012). "Strength in numbers: A multidisciplinary, project-based approach to introductory statistics education," *Journal of Effective Teaching, 12(2)*, 4-14.

GAISE (2012). *Guidelines for Assessment and Instruction in Statistics Education: College Report*. Retrieved October 27, 2013 from *http://www.amstat.org/education/gaise/GaiseCollege_Full.pdf.*

Hald, Anders (1998). *A History of Mathematical Statistics from 1750 to 1930*, New York: Wiley, p. 160.

Hesterberg, T., Moore, D.S., Monaghan, S., Clipson, A., Epstein,R., Craig, B.A., and McCabe, G.P. (2010), "Bootstrap Methods and Permutation Tests," Chapter 16 for *Introduction to the Practice of Statistics*, 7th edition, by David S. Moore, George P. McCabe and Bruce A. Craig, New York: W. H. Freeman.

Kaplan, Daniel T. (2009). *Statistical Modeling: A Fresh Approach*.

Lock, R., Lock., P.F., Lock Morgan, K., Lock, E.F., and Lock, D.F. (2012). *Statistics: Unlocking the Power of Data.* New York: John Wiley and Sons.

Moore, David S. (1997). "New Pedagogy and New Content: The Case of Statistics" (with discussion), *International Statistical Review*, 65, 123-165.

Moore, David S. and McCabe, George (1998). *Introduction to the Practice of Statistics: Nonparametric Tests and Logistic Regression: Chapters 14 and 15 Print Supplement.* New York: W H Freeman & Co.

Mosteller, Frederick, Fienberg, Stephen, and Rourke, Robert (1983). *Beginning Statistics with Data Analysis*, Reading, MA: Addison-Wesley.

Nolan, D. and Temple-Lang, D. (2010). "Computing in the Statistics Curricula," *The American Statistician*, 64, 97-107.

Tintle NL, Chance B, Cobb C, Rossman A, Roy S, Swanson T and VanderStoep J. Introduction to Statistical Investigations" Preliminary version. Fall 2013. Additional references via *http://math.hope.edu/isi.*

Tukey, John W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

Wagaman, A.S. (2013). "Meeting Student Needs for Multivariate Data Analysis: A Case Study in Teaching a Multivariate Data Analysis Course with No Pre-requisites." Retrieved October 29, 2013 from *http://arxiv.org/abs/1310.7141*.