Further, Faster, Wider

Chris J. WILD

Cobb (2015) is right on the money when he says, "For our profession, the valuable territory is the science of data," It is a territory we once had largely to ourselves but, as it expands rapidly and becomes more densely populated, we are sharing it with waves of immigrants and are in danger of becoming just one more indigenous people subsisting invisibly in the shadows of their homeland. Our competition, Cobb says, "takes place in the marketplace of ideas." It also takes place in the marketplace of credibility, in the marketplace of influence, in the marketplace of great jobs for graduates, in the marketplace of access to the best and brightest young minds, and in the marketplace of educational curriculum share. When business and organisations, as well as science, need to gain valuable insights from data, "Who you gonna call?"

As I have been reading Cobb (2015) I have also been reading the ASA Curriculum Guidelines (2014) and Diggle (2015), Peter Diggle's Royal Statistical Society Presidential Address paper. In stark contrast to the expansive vision of undergraduate statistics in Cobb (2015) and the ASA Curriculum Guidelines, Diggle says, "I would like to see less statistics in undergraduate mathematics degrees," This matters because in the UK, "most academic statistics groups now sit within departments or schools of mathematics or mathematical sciences." He did go on to talk about less statistics in undergraduate degrees being "counterbalanced by a radical expansion of postgraduate statistics teaching." But this is our old model, the conversion model whereby one obtains students to fund the discipline and replace its aging research-academics by attracting people into the fold who thought they were heading someplace else—generally into mathematics. It has been a long time since statistics graduate-programmes in western countries have been able to attract sufficient of their own nationals to be sustainable. They have been saved from withering only by an influx of students from poorer countries, mainly in Asia. Is this forever sustainable? Is this what "success" looks like?

This clash of viewpoints is a manifestation of big-tent versus small-tent statistics; see Rodriguez (2013); the "wider view" of Marquardt (1987; see also Wild, 1994, 2015) versus his narrower views; the "greater statistics" of Chambers (1993) versus his "lesser statistics," and the "wider field" of Bartholomew (1995). Big-tent statistics is increasingly favoured by the ASA as both its strategic and its service-to-society stance. It also underpins whether you see your role as educating a small number of great scientists in a fairly narrow if pure tradition, or education on a much grander scale to build a statistically capable society—as building a narrow tower or a broad-based pyramid, a select priesthood or a mass movement.

Does this really matter? It matters because statistics has an ancestry of great thinkers whose wisdom needs to be preserved and passed on. It matters because statistics has important messages for all of society, not just a few scientists. It matters because, "Those who ignore statistics are condemned to reinvent it" (attributed to Brad Efron by Friedman, 2001) and their ignorance can do real damage in the meantime. It matters because statistics is in a bad position to promulgate those messages because of the pitifully small market share it has in the educational curriculums of almost all countries. It matters because data is currently seen as exciting and valuable, and the people who know how to gain value from it are highly sought after. It matters because there has never been a better time for getting attention and market share for teaching a modern, accessible, data-centric statistics than there is right now.

Statistics education has an opportunity to help a wide cross-section of students come to a much broader appreciation of what data is and what it can do for them and society. It has an opportunity to help all students to make better sense of their world using data, to be not-easily-misled, and to prepare for a burgeoning job market. It has an opportunity to harness the power of visualization to greatly enhance the statistical understanding of a much wider spectrum of society. We can do so much more than just passive "statistical literacy," we can build significant statistical capability.

Where are most of the innovations fostering big-tent statistics education coming from? Mainly from elite U.S. liberal arts colleges like George Cobb's. They are a distinctive U.S. national treasure. With a few rare exceptions, research-universities are reactionary forces. They slow down any expansion of statistical vision because they are squeezed between the pincers of ranking systems, research assessment frameworks and research funding systems that favor traditional areas of strength and a narrow range of journals. To get promoted or funded, it is much safer to stick to your (predominantly mathematical) knitting. At this time of heightened opportunity, the needs of the wider society and the employment marketplace (more data, more accessibility to statistical capability) and the short-term health of many research-university academic units (more mathematics) seem to be pulling in opposing directions.

As with all of Cobb's writings, this paper is a joy to read both for the vividness of his imagery and his provocative messages. We have just started to settle down and get comfortable after his last onslaught when along comes George Cobb again, poking us once more with a sharpened stick. So we will pick ourselves up and lurch forward, making a little more progress.

I agree entirely with Cobb's thesis of a radical re-thinking of the curriculum from the ground up, particularly for intro and early applied statistics courses. The standard intro course has passed its use-by date. It reveals far too little of the exploding world of data and does it far too slowly. In the early stages of statistical education, I favor actually outdoing the fast-food com-

Online discussion of "Mere Renovation is Too Little Too Late: We Need to Rethink Our Undergraduate Curriculum From the Ground Up," by George Cobb, *The American Statistician*, 69. Chris J. Wild, University of Auckland (Email: c.wild@auckland.ac.nz).

petition by providing even faster food. The early stage relationship should have more in common with "courtship" than "eat-it-because-its-good-for-you" gruel. To summon up a phrase I use in talks, "Don't make students crawl over broken glass—until a desire has been aroused for what's on the other side."

I am all for the ambitious experimentation Cobb praises and advocates to push the limits of what we can do without (or with very few) pre-requisites. Where students will take many of these courses it may, fittingly in this big-data age, enable educational timespans to be shortened by massive parallelization using divide and recombine (see Cleveland and Hafen, 2014)!

But the end aim should not be a smorgasbord of choose-one intro courses each covering a small number of "advanced" topics at an intro level. The data world is exploding in scope and potential and we need to efficiently convey as much as possible of the scale and excitement of this in intro courses as well as providing personal empowerment—to create a sense of possibility and potential "for me in my life," not just something for some high-powered PhDs somewhere. We should target "What I can do with data and what data can do for me" to build a desire to learn more. Beginning experiences of data analysis should feel like driving a shiny sports car at breakneck speed along the Riviera, sliding around hairpin bends overlooking thrilling vistas. We spend too much time in working in windowless workshops with our heads stuck under the hood. This is not fanciful reverie. I have done enough prototyping with my iNZight (*http://www.stat.auckland.ac.nz/~wild/iNZight/*) and VIT (*http://www.stat.auckland.ac.nz/~wild/VIT/*) software projects and "Data to Insight" MOOC (*http://www.stat.auckland.ac.nz/~wild/d2i/4StatEducators/*) to believe that this is within our grasp. If struggling to get the right stuff out of software chews up a significant proportion of intro students' time then its the wrong software.

And let's stand some common practices on their heads. Let's excite students about what can go right before moderating that with "keep yourself safe" messages about what can go wrong. Let's distinguish between the fundamental statistical messages and enabling skills. Where time is short, let's concentrate on fundamentals and strive for the fastest ways to convey them. I fully favor the expanded curriculum that empowers students to speak the enabling languages used by statisticians (computing/algorithmic, graphics/visualisation, and mathematics). But they are not the fundamentals. They are great enabling skills for helping people on their way once they know where they are going. Visualisation, stands out from the others because it can provide a fast track to understanding fundamentals.

I applaud the more "data-science" agenda for those making a serious commitment to statistics. Integrity (Section 2.2) also demands that undergraduate statistics programmes provide good employment skills for the majority who will not go on to further study. There are more of these in the "data science" aspects of statistics than in most of what we have traditionally stressed. But we have to consider the striptease of what to reveal and when. Not everything we newly think is important has to be revealed straight away.

I worry about starting too early with wrangling messy data. Yes we need to teach statistics majors to deal with messy data. But extracting jewels from gloop is not something most people do because they love messing around in gloop. They want the jewels. But first they have to know (i) that jewels exist, and (ii) they might be in there. So lets first have them discover jewels in places where they are easier to find. Also coming into vogue for intro courses is working in "reproducible-research" modes. Yes statisticians should work like this, and yes, there is a stage in a program where it should become the standard way of operating. But all these things slow down what you can see and how fast you can see it. There should be a sniff test. Is this an enticing element of courtship? Or do I feel the skin-pricks of glass shards? So should we save it for after marriage? Or at least till after moving in?

Can statistics secure a central position in the new data world? Is there a will to find a way? Sadly I think statisticians in the best liberal arts colleges and Bob Hoggs BIG (Business, Industry, and Government) offer more hope than most research universities. The former want it. As to the latter, Im not so sure.

References

ASA Curriculum Guidelines (2014), "American Statistical Association Undergraduate Guidelines Workgroup," *2014 Curriculum Guidelines for Undergraduate Programs in Statistical Science*, Alexandria, VA: American Statistical Association, *http://www.amstat.org/education/curriculumguidelines.cfm*.

Bartholomew, D. (1995), "What is Statistics?" *Journal of the Royal Statistical Society*, Series A, 158, 1–20.

Chambers, J.M. (1993), "Greater or Lesser Statistics: A Choice for Future Research," *Statistics and Computing*, 3, 182–184.

Cleveland, W. S., and Hafen, R. (2014), "Divide and Recombine (D&R): Data Science for Large Complex Data," Statistical Analysis and Data Mining, 7, 425–433.

Cobb, G. W. (2015), "Mere Renovation is Too Little Too Late: We Need to Rethink Our Undergraduate Curriculum from the Ground Up," *The American Statistician*, 69.

Diggle, P. J. (2015), "Statistics: A Data Science for the 21st Century," *Journal of the Royal Statistical Society*, Series A, 178, 1–18.

Friedman, J. H. (2001), "The Role of Statistics in the Data Revolution," *International Statistical Review*, 69, 5–10.

Marquardt, D. W. (1987), "The Importance of Statisticians," *Journal of the American Statistical Association*, 82, 1–7.

Rodriguez, R. N. (2013), "Building the Big Tent for Statistics," *Journal of the American Statistical Association*, 108, 1– 6.

Wild, C.J. (1994), "On Embracing the 'Wider View' of Statistics," *The American Statistician*, 48, 163–171.

——— (2015), "On Locating Statistics in the World of Finding Out," *http://www.arxiv.org/abs/1507.05982*.