EDITORIAL: THE FUTURE OF STATISTICAL LITERACY IS THE FUTURE OF STATISTICS

JIM RIDGWAY (GUEST EDITOR)
Durham University
jim.ridgway@durham.ac.uk

JAMES NICHOLSON (GUEST EDITOR)
Durham University
j.r.nicholson@durham.ac.uk

INTRODUCTION

This special edition came about as the result of a discussion at ICOTS 9 in Arizona between Maxine Pfannkuch and the guest editors, Jim Ridgway and James Nicholson. It arose from an awareness that conceptions of statistical literacy need to be updated (see, e.g., Ridgway, Nicholson, & McCusker, 2013). We have been helped greatly in the preparation of this special edition by the co-editor Manfred Borovcnik whose energy and expertise have been invaluable. There are strident calls for curriculum reform and a variety of ideas on what the nature of these reforms should be. (e.g., Cobb, 2015, and associated responses). Given the wide variety of opinions and suggestions for action, we invited essays, as well as empirical research on statistical literacy. Full papers have undergone the normal SERJ peer review process; the invited essays have not been peer reviewed and present the authors' own views. There is a set of three opinion pieces on the theme of 'Against inferential statistics: how and why current statistics teaching gets it wrong'. This has the structure of a discussion – main argument, response and authors' reply to the response. The role of inferential statistics is a current and controversial topic in statistics and science more generally, and we hope these three papers will promote and provoke further discussion.

WHAT IS STATISTICAL LITERACY?

We take a UNESCO (2004) definition of *literacy* as a starting point.

> "Literacy is the ability to identify, understand, interpret, create, communicate and compute [...] Literacy involves a continuum of learning in enabling individuals to achieve their goals, to develop their knowledge and potential, and to participate fully in their community and wider society."

The Oceans of Data Institute (ODI, 2015) describes a *data-literate individual* as one who:

> "understands, explains, and documents the utility and limitations of data by becoming a critical consumer of data, controlling his/her personal data trail, finding meaning in data, and taking action based on data. The data-literate individual can identify, collect, evaluate, analyze, interpret, present, and protect data."

We are living in interesting times. There has been a dramatic increase in the volume and quality of data available from official sources that can be used to guide decision making, new and exciting ways to present and analyse data (e.g., via data visualisation), novel sorts of data (e.g., big data), and new ambitions (e.g., measuring hard to define

Statistics Education Research Journal, 16(1), 8-14, http://iase-web.org/Publications.php?p=SERJ

constructs associated with social progress); these all present new challenges and opportunities. Nevertheless, statisticians and statistics educators are facing an existential crisis driven by two radically different forces. The first is associated with a 'post-truth' era where the phrase 'alternative facts' is a synonym for deliberate deception, where news stories are invented then circulated widely via social media, and where legitimate journalism is routinely labelled as 'fake news' by the President of the USA. The second is the explosion of machine learning techniques being applied to large data sets, and associated hubris around 'the end of theory' (e.g., Anderson, 2008).

In the invited essays particularly, many authors argue that changes in available data sources, and in ways to communicate about evidence should be associated with changes in our ideas about statistical literacy.

Robert Grant explores the relationships between statistical literacy and data science. He points out that traditional boundaries between computer science and statistics have blurred, and that defending turf is not sensible – new challenges (such as the analysis of big data) need new skill sets. Grant offers advice on how the two communities can learn from each other – the machine learning community should make more use of statistical tools such as bootstrapping and Bayesian methods, and should pay more attention to the ways that new models can be evaluated. Grant sees statistical literacy in this context as including a sensitivity to modelling assumptions (e.g., linearity and independence) and a willingness to explore data with simple tools before committing to black-box routines. Statisticians should make more use of relatively new computer-intensive methods such as *k*-fold cross-validation. Along with others (e.g., Hardin et al, 2015; Finzer, 2016), Grant argues that undergraduate statistics courses should introduce students to data science concepts. He also argues that the traditional statistics course should be turned on its head – teaching should start with families of predictive models, and work backwards to theorems.

Robert Gould's title is 'data literacy is statistical literacy'. Citizens need to understand and act on their relationships with data – how personal data are collected, by whom, and for what purposes, and should be aware of issues around privacy. He argues that students should learn about data provenance, data storage, data manipulation and representations, and should also be introduced to classes of models not found in traditional courses (such as new approaches to predictive modelling) and representations such as classification and regression trees (CART). He describes the use of participatory sensing in high school, mediated by mobile devices, that is part of the *Mobilize* project. Data collected in this way is messy, and Gould reports that classroom discussions address key issues such as data ownership, and sample bias, as well as being a gateway to discussing appropriate ways to analyse big data.

Sinclair Sutherland and Jim Ridgway argue that as the media are increasingly making use of interactive data visualisations (IDV), so statistical literacy requires a willingness to learn and critique novel ways to display data. IDV do more than present data in new ways; they facilitate exploration, and so effective use requires engagement and exploration – Gal's *dispositions* (Gal, 2005), as well as new skills. More radically, as the functionality of IDV increase, users will be able to explore data in increasingly sophisticated ways, further blurring distinctions between 'producers' and 'consumers' of statistics.

Chris Wild also points to the increasing capabilities of IDV to present data, and identifies the prospect of powerful analytic tools linked to large public data bases being available to everyone. He argues that educators should think about curriculum design from the viewpoint of preparing students for the current data-rich world, and for possible future worlds, and sets out some guiding principles about things likely to be empowering,

and having enduring value. Wild argues that students should be asked to consider the quality of data from different sorts of study, paying attention to bias (especially when dealing with big data), random error and confounding, and the kinds of extrapolations that are plausible from different sorts of evidence. Wild also offers useful heuristics for everyone concerned with drawing conclusions from quantitative evidence.

Theodosia Prodromou and the late Tim Dunne also explore the construct of statistical literacy in the context of the data revolution. They emphasise the need to develop students' dispositions – encouraging them to adopt a critical stance towards data, to examine metadata, the context, and to be aware of opportunities offered by new data visualisations. Prodromou and Dunne argue that statistics should be conceived as modelling, not simply the application of standard models to well-behaved data sets. They draw attention to the problems of deducing causality from observational data, especially when one might anticipate time lags between causes and effects.

Joachim Engel emphasises the importance of being able to think critically about data. He points out that the data used in political debate is unlike the artificial "small samples that dominate today's curricula". Official data sources are often used; data are multivariate; data change dramatically over time. Understanding such data requires an understanding of how variables co-vary or influence each other, and the need to think about networks of causal factors. He argues that understanding does not necessarily depend upon knowledge of advanced techniques, but can be enhanced by knowing about common misconceptions, the problems of making conclusions about causality, and phenomena such as Simpson's paradox. He points out that any evidence-informed discussion on social issues raises central questions about the scientific process – how can one operationalise constructs such as discrimination? In terms of curriculum reform, he argues for more emphasis on multivariate data displays (such as scatter-plot matrices and multivariate time series), and on non-linear relations and interactions.

On an optimistic note, Milo Schield argues that the 2016 revision of the GAISE Guidelines marks a major step forward in promoting statistical literacy via its increased emphasis on evidence appropriate for decision making – such as paying attention to study design and multivariate data and associated concepts such as confounding.

INNOVATIVE APPROACHES TO BOOST STATISTICAL LITERACY

One can distinguish between the intended curriculum, the implemented curriculum, and the attained curriculum. Much of what has been mapped out so far relates to the intended curriculum – and here, the authors setting out to specify the intended curriculum are not necessarily in a position to influence the implemented (i.e., taught) curriculum as it applies to schools or across the higher education sector. A number of authors have argued that courses in mathematics and statistics have focused on developing technique at the expense of problem solving in realistic contexts (e.g., Tishkovskaya & Lancaster, 2012) – this is a long way from the ideas about the future curriculum directions expressed here. At school level, although statistical ideas pervade much of the curriculum, there can be inconsistencies and incoherences in curriculum documents (e.g., Watson & Neal, 2012). Teachers are not necessarily confident or competent to teach statistics (e.g., de Souza, Lopes, & de Oliveira, 2014), and there can be large gaps between the intended and implemented curriculum (e.g., North, Gal, & Zewotir, 2014). Some papers in this volume address these issues.

There are long-standing concerns about the absence of quantitative skills in social science graduates (e.g., ESRC, 1987). In the UK, a large-scale initiative – *Q-step* – was launched to address this problem. Two approaches are described here. Jackie Carter and

her colleagues use an approach to teaching statistical methods which includes paid placements where undergraduates work in oganisations which use social research, undertaking research projects of value to the host organisation. They provide three case studies (from over 100 students who have had placements) that illustrate benefits both to students and to the host organisations. The taught part of the course focuses on contexts of direct relevance to social science students; students search for data sets relevant to their own interests. This leads naturally to key questions about data provenance, representativeness, metadata and the like, that can be missing if the focus is exclusively on statistical technique. Student placements have a long history and some obvious disadvantages, such as funding, administrative time, assessment, supervision; Carter et al. describe how these were addressed.

Julie Scott Jones and John Goldring address the problem of poor statistical literacy in social science classes. They argue that there are a number of barriers to student learning, and offer some pedagogic strategies. In particular, they argue that both teachers and students need to be empowered to use statistics, and that this can involve removing barriers to learning such as mathematics anxiety, intellectual resistance and 'troublesome knowledge'. They describe a new course designed to put students at the centre of learning. There is some diagnosis and remediation of poor numeracy skills, illustrations of how simple numbers changed the Victorian world for the better, and students are required to tell sociological stories based on data. Students can see road maps about course content, and there is a system of student self-checking, and students can sign off achievements; there is also evidence that the new course works.

Ayse Biglin and her colleagues describe *Opening Real Science,* a project that brings together subject specialists, teacher educators and educational designers to create online learning modules for pre-service teaching programmes – including modules on statistical reasoning. They begin by emphasising the role that statistics plays in people's lives using resources that include Hans Rosling's *Gapminder* videos. They asked students to critique newspaper accounts of data from PISA and TIMMS; students conducted investigations with a statistical theme, and presented this work; and engaged in a critical evaluation of their peers' work.

Stephanie Budgett and Drusilla Rose describe their investigations designed to inculcate key ideas about margins of error in final school year pupils, using media reports. They offer a hypothetical learning trajectory for statistical literacy, and some insights into conceptual difficulties experienced by students.

Tamires Queiroz and her colleagues explore the importance of context and affect, and document the high proportion of emotional responses, feelings and beliefs that students exhibit when interpreting statistical information on important social topics. They argue that interpretation of statistical information is a complex activity that draws upon affect, as well as on statistical and mathematical knowledge.

Rosemary Callingham and Jane Watson replicate the hierarchy of statistical literacy described in Watson and Callingham (2005) using new data (of 2007-09) and a revised task set that includes more items on probability (the paper presents tasks that can be used both for teaching and for assessment). Descriptors of the levels of the hierarchy reflect increasing cognitive complexity (from *idiosyncratic* to *critical mathematical*). They report on increases in attainment in each of five school years associated with a programme based on increasing the pedagogical content knowledge of teachers. However, they are cautious about the extent to which higher levels of the hierarchy could be reached by younger children, pointing to work of developmental psychologists such as Piaget and Inhelder (e.g., Inhelder & Piaget, 1958) who locate the emergence of formal operations – such as the ability to abstract and generalise – around 12 years and beyond.

Steve MacFeely and his colleagues consider the key success factors for statistical literacy poster competitions, based on the International Statistical Poster Competition. Posters are significant because they encourage students to use data to tell stories, not simply as the object of an exercise to demonstrate statistical technique. Posters call on the traditional skills of analysing and presenting data, but also on accessing data and communicating findings – skills often missing in traditional curricula, where 'communicating findings' is often no more than stating the outcome of a hypothesis test or interpreting a single graphical representation. The authors point to national statistics offices and statistical societies which promote statistical literacy via competitions of various sorts, including posters and other outreach activities. Key success factors include: co-operation between institutions (e.g., schools, statistics offices, and government departments); celebrating achievement; support for teachers (e.g., by providing pedagogical materials on-line); support from statistics offices, universities and statistical societies; and technology. We return to the theme of groups collaborating to promote statistical literacy, later.

COMBATTING THE EXISTENTIAL CRISIS

The Oxford Dictionaries word of 2016 was post-truth, which is defined as "relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion or personal belief." This has been commonly used in the phrase post-truth politics. A climate of post-truth can make statistical literacy an irrelevance.

Queiroz et al. (this volume) demonstrated the extent to which emotional responses colour interpretation. Contempt for evidence is socially corrosive. It violates the core values of the statistical community. Components of statistical literacy must include the ability to evaluate the quality of evidence and argumentation that is offered; at least as important as a core of technical competences is a disposition to engage actively with evidence.

An Ipsos MORI survey (2016) of 1019 British adults used the prompt "Now I will read you a list of different types of people. For each would you tell me if you generally trust them to tell the truth, or not?" Of the 24 professions listed, politicians generally came bottom (15%), then government ministers (20%) then journalists (24%). Scientists had an 80% endorsement; nurses were top, at 93%. This can be viewed positively; the results suggest a critical stance to some information sources. However, if the results reflect a nihilistic view that things are unknowable, advocates of post-truth are winning. If the lack of trust means that people are looking for more reliable sources of information, then there is scope for optimism.

In the run-up to the UK referendum on Brexit, a leading politician (Gove) announced that 'Britain has had enough of experts'; both sides of the campaign made claims that could not be justified. A charitable interpretation of Gove's position would be that he was claiming that Brexit would take the UK into uncharted waters, where there are no experts, or that political decisions should be informed by subject expertise, but not driven by it. In the 2016 US presidential election, there was a great deal of fabricated evidence.

Wild (this volume) poses a critical question: *"What can we come to know and how can we come to know it?"* There are dangers in believing that all questions of substance can be answered by recourse to data. The mantra of 'evidence-based decision making' is naïve. Decisions do not happen in a vacuum; data is not always fit for purpose, and at best evidence can *inform* political decisions. Following Wild, we need to look carefully at our implicit acceptance of an 'evidence-informed' world view, and to acknowledge that

scientific communities sometimes claim to know more than they can know. Over-claims are dangerous because they reinforce Gove's dismissal of expertise.

STRENGTHS AND LIMITATIONS OF STATISTICAL METHODS

White and Gorard (this volume) argue that social scientists are using the wrong methods to study social phenomena; Nicholson and Ridgway agree with this criticism, further pointing to failures to replicate results, and to unprincipled data dredging. A number of authors in this volume argue that students should be aware of the strengths and limitations of different data gathering techniques, and that core statistical ideas such as confounding and causality should receive more attention. Statistics should be thought of as a modelling exercise, and we should take note of a comment by Sir David Cox "Most real life statistical problems have one or more nonstandard features. There are no routine statistical questions; only questionable statistical routines" (Cox, quoted in Chatfield, 1991, p. 240). These are all things that can be done within the academic community. However, 'post-truth' is an urgent and pervasive problem that needs a co-ordinated assault from as many agencies as possible.

COLLABORATIONS TO IMPROVE STATISTICAL LITERACY

MacFeely et al. (this volume) point to the many collaborations between different organisations associated with the International Statistical Poster Competition. The paper by Phoebe Arnold (this volume) describes the roles of Full Fact, an influential fact-checking organisation in the UK. She describes a range of initiatives aimed at honest reporting of evidence in public places. These include: fact checking; offering help to journalists and politicians who are trying to represent things accurately; describing key statistical ideas in simple terms, often using analogies; creating short summaries of debates; preparing briefing documents; and producing a catalogue of mistakes that are commonly made. All of these promote statistical literacy, and all would be worthwhile curriculum activities. At least as important is Full Fact's achievement in assembling a large team of statisticians from a variety of backgrounds (including statistics offices) paid by their employers to take part in an intense period of fact-checking around election time.

CONCLUSION

Statistical literacy is a pre-requisite for an informed democracy. Increasing statistical literacy is a key element in warding off the existential crisis we face. Revising current curricula in school and at university to ensure that there is an adequate focus on using evidence to make decisions in realistic contexts is an essential starting point. At least as important is for statistics educators to take a broader view of their task, and to engage directly with the illiteracies encountered in broadcast and social media – for example by direct critique, or by promoting statistical literacy directly. There is a need for disparate elements of the statistics community to come together; cultivating statistical literacy across the whole of society should be a goal that brings like-minded people together with a common cause.

REFERENCES

Anderson, C. (2008). *The end of theory: the data deluge makes the scientific method obsolete.* [Online: archive.wired.com/science/discoveries/magazine/16-07/pb_theory]

Chatfield, C. (1991). Avoiding statistical pitfalls (with comments by D. F. Andrews, R. A. Bailey, M. K. Clayton & E. V. Nordheim, N. Glick, C. L. Mallows & D. Pregibon, and D. A. Zahn, and a rejoinder by the author). *Statistical Science, 6*(3), 240–268.

Cobb, G. (2015). Mere renovation is too little too late: We need to rethink our undergraduate curriculum from the ground up. *The American Statistician*, *69*(4), 266–282.

de Souza, A. C., Lopes, C. E., & de Oliveira, D. (2014). Stochastic education in childhood: Examining the learning of teachers and students. *Statistics Education Research Journal*, *13*(2), 58–71.
[Online: iase-web.org/Publications.php?p=SERJ_issues]

Finzer, W. (2016). *Common online data analysis platform (CODAP)*. Emeryville, CA: The Concord Consortium. [Online: concord.org/codap]

Gal, I. (2005). Towards "probability literacy" for all citizens: Building blocks and instructional dilemmas. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 39–63). New York: Kluwer Academic Publishers.

Hardin, J., Hoerl, R., Horton, N. J., Nolan, D., Baumer, B., Hall-Holt, O., Murrell, P., Peng, R., Roback, P., Temple Lang, D., & Ward M. D. (2015). Data science in statistics curricula: preparing students to "think with data". *The American Statistician, 69*(4), 343–353.

Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. London: Routledge and Kegan Paul.

Ipsos MORI (2016). Enough of Experts? Ipsos MORI Veracity Index 2016.
[Online: www.ipsos-mori.com/researchpublications/publications/1896/Enough-of-Experts-Ipsos-MORI-Veracity-Index-2016.aspx#gallery[m]/0/]

North, D., Gal, I., & Zewotir, T. (2014). Building capacity for developing statistical literacy in a developing country: Lessons learned from an intervention. *Statistics Education Research Journal, 13*(2), 15–27.
[Online: iase-web.org/Publications.php?p=SERJ_issues]

Oceans of Data Institute (2015). *Building global interest in data literacy: a dialogue*. Waltham, MA: Educational Development Center. [Online: oceansofdata.org/our-work/building-global-interest-data-literacy-dialogue-workshop-report].

Oxford Dictionaries (2016). Word of the year 2016 is … .
[Online: en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016]

Ridgway, J., Nicholson, J. & McCusker, S. (2013). 'Open data' and the semantic web require a rethink on statistics teaching. *Technology Innovations in Statistics Education*, *7*(2), 1–12.
[Online: escholarship.org/uc/uclastat_cts_tise?volume=7;issue=2]

Tishkovskaya, S. & Lancaster, G. A. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education, 20*(2).
[Online: www.amstat.org/publications/jse/v20n2/tishkovskaya.pdf]

UNESCO (2004). *The plurality of literacy and its implications for policies and programs.* UNESCO Education Sector Position Paper: 13. Paris: UN Educational and Cultural Organization. [Online: unesdoc.unesco.org/images/0013/001362/136246e.pdf]

Watson, J. & Neal, D. (2012). Preparing students for decision-making in the 21st century – statistics and probability in the Australian curriculum: Mathematics. In B. Atweh, M. Goos, R. Jorgensen, & D. Siemon (Eds.), *Engaging the Australian curriculum mathematics – Perspectives from the field* (pp. 89–113). Wahroonga: Mathematics Education Research Group of Australasia. [Online: www.merga.net.au/onlinebooks]