
Article:
Nguyen, A and Lugo-Ocando, J (2016) The state of statistics in journalism and journalism education: issues and debates. Journalism, 17 (1). pp. 3-17. ISSN 1464-8849

https://doi.org/10.1177/1464884915593234

Introduction

The state of statistics in journalism and journalism education: issues and debates

An Nguyen & Jairo Lugo-Ocando

Key words

Statistics for journalists, numbers in the news, data journalism, numeracy skills, quantitative literacy, data reporting

Authors

An Nguyen is Principal Lecturer in Journalism at Bournemouth University, UK. He has published, among other things, two books and 25 papers in several research areas, including the diffusion and social impact of online/digital journalism, news consumption and citizenship (including citizen journalism), science journalism, and news and development in a globalising world. He is in the final stage of a book on statistical newsgathering (co-authored with Jairo Lugo-Ocando). Address: Department of Journalism and Communication, Faculty of Media and Communication, Bournemouth University, Talbot Campus, BH12 5BB, England, UK. Email: anguyen@bournemouth.ac.uk.

Jairo Lugo-Ocando is Associate Professor in the School of Media and Communication at the University of Leeds. He is author of three monographs, one edited book, 18 peer-reviewed journal articles and 11 book chapters together with several public and private sector reports. He has been the Principal Investigator for a series of grants and fellowships funded by Nuffield Foundation (UK), Carnegie Trust for the Universities of Scotland (UK), the TODA Institute (Japan), AHRC-CDA (UK), John F. Kennedy Library (USA), among others. He has been a visiting research fellow at the Universidad Católica Andres Bello in Venezuela and the National University of Singapore and, since 2010, a visiting professor of the Doctoral Programme in Media & Communication at the Universidad de Malaga (Spain).

The state of statistics in journalism and journalism education: issues and debates

Statistics, without exaggeration, are part of the fabric of the contemporary world[1]. They are a pervasive feature of daily life, shaping the way we think and behave in ways that many of us no longer realise. Almost every key aspect of modern life – the quality of the air we breathe, the severity of a medical condition we have, the performance record of a soccer player we like, the potential price of the house we sell, the condition of the social care system we use, the health of the economy in which we operate, the national leader we want to vote, so on and so on – is statistically measured in one way or another. "If you live in Britain, there's no such thing as a day untouched by the ONS [Office of National Statistics]," declared the Guardian in the caption of a recent graphic illustration on July 31, 2013. Glen Watson, ONS's Director General, comments in the same article that even mundane things like the thickness of the water pipes in the house or the stock on the local supermarket shelf are shaped by statistics. As such, statistics have long been a staple of daily news – from "hard news" about politics, economics, business, finance, science and education to "softer" categories such as health, crime, sports and entertainment, community or other areas of social life. And it is likely to become even more so in the future, when the "big data" society is gradually normalised.

Yet when it comes to the quality of such news reporting, the media often "get a bad press" – to the extent that some scientists and experts have come to assume that journalists seldom get numbers right (Utts, 2010). Meanwhile, many journalists, suffering from a "blind spot" for numbers, tend dismiss statistics altogether. Confessing that they hate math at school that data make them feel dizzy, they would not hesitate to admit that they choose a journalism career to work with words, not numbers (Maier 2002; Yarnall et al 2008)[2]. For some, numbers are not only hard to swallow, but also fly in the face of what journalism is about. David Randall (2000, p. 73) observed that "quite a few" journalists see numeracy as "a kind of virus which, if caught, can damage the literary brain, leading to a permanent loss of vocabulary and shrivelling of sensitivity". In most newsrooms, according to Peter Wilby (2007), "literacy is considered

essential for reporters – or at least their subeditors – but not numeracy." Meanwhile, it is still quite hard to find statistics courses in university journalism programmes, although this varies from one country to another. In the UK, accreditation bodies such as National Council for the Training of Journalists, Broadcast Journalism Training Council or Periodicals Training Council barely mention numeracy as a professional requirement (Harrison, 2014). In the US, the provision of statistical training in journalism education is less uncommon but is by and large still not to the extent that one would expect, given that the Accreditation Council on Education in Journalism and Mass Communication has included for some time the ability to "apply basic numerical and statistical concepts" among its 13 "core values and competencies".

All of this manifests a vast misunderstanding, underestimation and ignorance of the nature and the role of statistics in daily newswork. It was time to demystify a number of crucial issues.

Statistics, not mathematics

First, it is not hard to make use of statistics for news purposes. Journalists' traditional "number phobia" is not because of statistics per se, but because their nature is either vastly misunderstood or too narrowly understood. The job of handling numbers for the news is often wrongly perceived as that of measuring, calculating and analysing things with eye-numbing formulae. Statistics and mathematics are two different things: it is not necessary to be adept at mathematics to be able to use statistics effectively. However frightening they might look, statistical analyses are about the application of valid reasoning, not calculation. Mistakes are often made in the news, but few involve getting the math wrong: most are due to flaws in the logic applied to data and their context (Moore and McCabe, 2003).

In other words, the journalist's job is not to learn and use eye-numbing formulae or calculate complicated things. For the most part, journalists deal with pre-processed data packages from other sources and what they need is a permanent determination to question data and a basic level of statistical reasoning. "You don't need to be a nerd to improve your reporting of news with numbers," says Deborah Potter (2009). "You just need to remember one basic, journalistic question: does this make sense?" Some basic knowledge of statistics is essential but what journalists need the most is not a set of skills to calculate or create their own data but one to use

logical, valid reasoning and journalistic scepticism to (a) find and acquire data, (b) explore and evaluate their real meaning in context, (c) investigate non-numerical factors shaping them, and (d) report them in a balanced, fair, accurate, accessible and engaging manner. All this does not require any special math skills. If one can add, subtract, divide and multiply, he/she can learn to handle statistics for the news, as long as he/she is willing to apply to data the same probing and enquiring mind that is essential for any other newswork.

Few journalists, for example, would need to know the scary look of the formula for a correlation co-efficient: what they need is to understand what those numbers between -1 and +1 mean in practice (what can be considered a strong, moderate or weak positive/negative association) and, more importantly, to understand that such association is not necessarily a cause-effect relationship. Even in computer-assisted reporting and data journalism, where journalists do perform software tasks to create their own data, data production skills are not necessarily the most demanding part of the process. The central task is still to find and assess existing raw datasets before using them, then to apply logics and reasoning to both software-based processes and their resulting computer output. Software skills can be acquired quickly – e.g. some free data visualisation apps on the Internet today require users no more than a few days to learn. But these skills are meaningless – and could be dangerous – if the ability and habit to question and evaluate data are absent.

It should also be noted that the stereotype of the statistically incompetent journalist is not entirely correct. SAT scores in mathematics of those starting journalism in US universities have been found to be as good as those starting other degrees (Dunwoody & Griffin, 2013). Other research shows that while it varies from one newsroom to another, journalists do have the ability to work with numbers (Maier, 2002) – and there are numerous examples of journalists turning numbers into inspiring, influential and engaging journalism (Utts, 2002), some of which will be mentioned below.

Statistical reasoning as prevention measures against "lies and damn lies"

Second, whether they like them or not, it is the journalist's duty to learn, master and use statistics competently, for at least two broad reasons. For one thing, the emergence of the "big data"

society means daily newswork itself is based more on "number crunching" than any time in the past. Many digital newsrooms today are inundated with live web metrics, erecting large data boards for reporters and editors to constantly monitor what audiences are doing with their news sites and apps, where they are from, where they go to, what they read/like/share/tweet/retweet the most and so on. These numbers behind the news are diverse and complex – e.g. some tracking software could produce several hundred variables and even pin down the advertising income that a news story can generate. If journalists have the ability to handle these metrics, they can harness them into an excellent, unique tool for more pertinent, more engaging and perhaps more viable news products. But if they do not have the statistical reasoning skills to understand and use web metrics wisely and calmly, journalists might risk falling in the trap of "the sentiment of the crowd" at the expense of their professional judgement, unless. In this new "click-thinking" culture, web metrics could easily deepen one of journalism's already severe crisis – the dumbing down of news – and bring newsroom tensions and conflicts to a new height (see Nguyen, 2013 for a review of these issues).

For another thing, which is of exclusive focus in the rest of this article, journalism is operating in an increasingly chaotic world of "lies, damn lies and statistics". Statistics, of course, do not lie. They cannot. It is those using statistics that lie, intentionally or unintentionally. Numbers are neither as neutral nor as objective as they look: they are a human invention to describe and represent the world out there and thus are subject to human reasoning and human capacity. In other words, the real problem is not numbers per se but the way people produce, use and assign meanings to them. Statistics-based lies, whenever they occur, are often because data are inappropriately produced or improperly interpreted, for all sorts of benign or malicious, objective or subjective reasons. And this is why good news reporting of statistics is badly needed. People tend to place more faith in numbers than in words, and the way they are presented in the news plays a crucial role in reinforce or challenge such faith. Without journalists' help to understand data, as Gigerenzer (2007, p. 53) said, "the public is susceptible to political and commercial manipulation of their anxieties and hopes, which undermines the goals of informed consent and shared decision making."

Such manipulation is becoming more sophisticated than ever. Today, realising the deep and powerful penetration of statistics into the way people think, believe and behave, all major social, economic and political institutions have integrated numbers as a central part of their public communication – including "news management" – strategies. Often, they mobilise complex, resourceful public-relations machines to pump into the newsroom all sorts of data that work to their advantage. One result of this is a deluge of "bad statistics" out there, i.e. those that are derived from a deliberate manipulation of the data collection and/or analysis processes. In some cases, they are from "research projects" in which the "researcher" knows the conclusion before he/she starts. In others, it is about manipulating and "massaging" data to advance some interests at the expense of their rigour. In some more shocking cases, the data arriving at the news desk are completely fabricated and do not exist. It is in part because of this chaos that the UK parliament passed in 2007 the Statistics and Registration Service Act, giving birth to the UK Statistics Authority, an independent quasi-governmental body that oversees the ONS and scrutinizes all official statistics produced in the country.

And it is not just serious data that constantly seek the limelight: seemingly silly or bizarre things do, too. Consider, for instance, the following topics that Kevin Peachey (2010), a BBC consumer-affairs reporter, compiled from press releases:

- An average British woman walks the same distance as that between London and Hull as a result of shopping every year, according to research by one of the UK's largest retail chains.
- Liverpudlian women have the largest breasts in the UK, the same retailer announces on another occasion.
- Batman is the superhero boss that most UK employees would like to have, says a recent campaign to raise awareness of employees' hidden potential.
- Dry cleaning experienced a "dramatic" increase as a result of the credit crunch, according to a dry cleaning company.
- Holidaying at least once a year is more important than owning a home or having children for a "massive" 95% of the British population, a travel search site claims.
- The ideal annual salary for UK workers is £38,000, a poll for a recruitment company shows.

- More and more people are fabricating ill and dying animal stories to "pull sickies and get out of work", a pet insurance firm announces.

Peachey observes that "many, arguably the majority" of these would go unreported, but they still "prove to be appropriate fodder for stories". At least, they hold a strong appeal to some of the tabloid press.

The opposite of the urge to release data also happens. Some sources – especially those that have legal obligation to release available data for public scrutiny (e.g. government agencies, publicly traded firms) – often make efforts to hide data that might work against their interest and reputation. In some cases, they use a range of tactics to "bury bad news" or distract journalists from sensitive data – e.g. choosing to publish data at a time when journalists cannot treat them with sufficient attention and due diligence, or releasing a huge amount of data all at once to make it extremely hard for journalists to spot controversial data within a short time span. In other cases, they find all possible measures to make it difficult, expensive, intimidating, or even hopeless for journalists to obtain data – e.g. it took the Daily Telegraph five years of legal fighting with the UK House of Commons before they were granted access to Members of Parliament's expense data, which gave rise to a ground-breaking investigation into how these expenses were abused by politicians (more on this case below).

In that context, it does not make sense for journalists to say that they are not here to work with numbers. People badly need the media to find and provide them with high-quality and trustworthy data so that they can function effectively, either as citizens or consumers, in their increasingly data-driven daily environment. As such, journalists' "laissez-faire" attitudes to data – on the naïve and convenient basis that numbers speak for themselves – cannot be afforded. Without a watchful eye on numbers and the basic skills to handle them, journalists would not be able to obtain an adequate understanding of many news events and issues. They would not be able to detect and expose misleading or cheating data that are of paramount importance to the public. They might be easily lured away from vital figures that sources want to conceal and the public needs to know. Sometimes, such uncritical reporting leads to sadly hilarious claims. More common and more serious is that journalists would serve as unwitting mouthpieces for powerful

and resourceful sources. Very often, without the confidence to scrutinise and challenge data, journalists opt for the easier route: they adopt the superficial but safe "he said, she said" formula and highlight to audiences what is already highlighted by sources. As Wilby (2007) pointed out:

> Look carefully at the statistics-based stories in any newspaper and you will see that few are the result of original journalistic research. The press challenges official figures only after think-tanks, pressure groups or MPs have done the work.

As such, Battersby (2010, p. 59) observes that numbers that play a powerful role in public debate and government-decision making "are often given a credence and significance they simply do not deserve". American public policy scholar, Kathleen Geier (2012), made similar observations on how "crap social science" damages the body politic. It is "bad enough," she noted, to have ideologues to "deploy dubious studies and statistics directly, in their own writings, speeches and interviews" without being held accountable. It is worse, however, "when shoddy research and dubious factoids get respectful attention from mainstream reporters and pundits". This, as Geier went on to argue, happens all too often and

> when these "facts" get disseminated, poisonous "conventional wisdom" about a subject can develop, and the general public is bamboozled. Much reporting about social science is, unfortunately, quite terrible. Sensation-seeking journalists will promote the research that reports the most dramatic, headline-grabbing results; research that reaffirms pre-existing prejudices tends to get a disproportionate amount of attention… All too often, the most basic questions about a piece of research are not asked.

The problem can be worse when journalists force-fit objective facts and figures into their own personal perception and/or professional framing of the world. As Michael Blastland – a founder of BBC Radio 4's More or Less – told us in an interview:

> Journalists tend to use statistics to reinforce their own views and pre-conceptions of reality. They take data that can fit in their own narrow scope of what the story should say. They have to fit the format of news stories that have a beginning and an end. In so doing, journalists are not reporting accurately the events and phenomena that they are trying to describe or analyse.[3]

However, sifting through recent cases and examples scrutinised by pioneer number-checking websites – such as Fact Check of Channel 4 News, More or Less of BBC Radio 4, FullFact.org in the UK, or the Pulitzer-winning PolitiFact of the Tampa Bay Times and The Numbers of the Wall Street Journal in the US – suggests that "lies and damn lies" can still get through the news media with relative ease because journalists do not have a determination to apply their traditional watchful and sceptical eyes to data. The ultimate consequence is on the public's ability to judge what is going on in the world. The consequence of improper reporting of statistics can be far-reaching and hard-hitting. Poor crime data in the news, for example, could promote an excessive fear of crime, create a state of moral panics, and push the public and elected leaders into unsound decisions (Schlesinger and Tumber, 1994). Medical news based on questionable data – especially those about health risks and "miracle cures" – can generate unnecessary worries, false hopes, meaningless lifestyle changes, wasteful spending on medicine, undue resistance to doctors' advice and so on (Reyna et al 2009). Inaccurate or insensitive reporting on stock market data could cost the entire life savings of many "mum and dad" investors (Hope, 2011; Kleinnijenhuis et al., 2013). The excessive reliance of business journalists on the book values of companies (i.e. profit and loss figures, rather than real assets and market values) has been attributed to some extent to recent financial scandals, such as the 2008 credit crunch that gave rise to the recent global economic crisis, or the 2001 collapse of the energy giant Enron in the US. In some cases, poorly reported numbers might fuel tension, conflict or bloodshed, with the post-911 public sentiment against the Muslim world being one contemporary example (Esposito & Mogahed, 2007).

Does this mean journalists can check every statistic that comes their way? Realistically, we doubt that they can, given the dynamics of today's newsrooms. Under the intense pressure of deadlines and editorial politics, newspeople often have little time to work on stories in general, let alone number-crunching ones. Insufficient newsroom resources add insults to injuries. Decades of declining audiences and advertising revenues have seen continuous savage cuts to news budgets, with many news outlets now working with less than half of the personnel they had in the 1970s and depending heavily on press releases. Deep-digging investigative journalism initiatives have become either a luxury or a rare species in danger of extinction, being replaced with trivial, dumbed-down content and celebrity gossip. Meanwhile, the unprecedented intrusion

of digital technologies has placed more pressures on journalists, requiring them to multi-task for multi-platform news deliveries within the same, if not less, amount of time as they used to have.

It is, however, precisely these problems of contemporary journalism that justify the need for every journalist to possess a basic level of statistical competence, or at least a habit to question statistics in the same way as they do to any other kind of raw information. In the increasingly resource-scarce and demanding newsroom, that habitual scrutiny of data might be seen as a most effective prevention measure for journalists to avoid or minimise the risk of falling victim to bad statistics and those behind them, or a "vaccine" against the many data-based myths and untruths that can be dismissed with a basic check. Journalism, after all, survives and thrives because of its mission as a verification discipline (Kovach & Rosenstiel, 2001). If journalists keep merely attributing data and data-based claims to the powerful and the elite or letting dubious figures go off the radar the day they are published, journalism risks becoming even less worthy in the long term. Their already at-risk profession might further fail to fulfil its very public service, recede into a deeper crisis of public trust and confidence, and even lose its raison d'être in a world where data are such powerful weapons for so many interest groups. As Livingston and Voakes (2005, p. 9) argue, "if we value independence as a cornerstone of journalism ethics, part of that independence must be the ability to assess numerical information without relying on the source".

Statistics for inspiring, world-changing journalism

From a positive perspective, it should be reminded that there is a great deal of good statistics out there. They can be seen as one of the most valuable assets that modern journalism might have to proactively approach events and issues of the day. A basic competence in statistical reasoning would enable and empower journalists to harness the power of good data – namely their capacity to summarise patterns, to depict and predict trends and to identify possible causes and effects – for news purposes. Used wisely and properly, good data allow journalists, among other things, to overcome the sometimes tyrannical rule of intuition, to link anecdotal evidence to the big picture, to bust myths, to challenge prejudices, to connect seemingly distance incidents/events, and to turn ostensibly boring, soulless numbers into vivid, meaningful representation of the world.

And if they manage to go beyond basic statistical reasoning to be able to generate their own data, journalists will have wide-open opportunities to produce deep-digging, ground-breaking and world-shaking journalism. The recent history of journalism is not short of such reporting. As early as 1967, when riots were taking place in Detroit, a journalist named Philip Meyer partnered with two scholars at the University of Michigan, using an IBM 360 mainframe computer to analyse survey data on the social profile of rioters. What he found and reported in his stories was striking: college-educated people were as likely as high school dropouts to participate in these riots. In a time when low education was widely believed to be behind riots and computers were still something of a cumbersome novelty, Meyer's work won him a Pulitzer Prize and, along with other achievements, led him to later become an international authority in the area of "precision journalism" – journalism based on scientific evidence.

Similar successes have become more common since then. Daniel Golden of The Wall Street Journal, for instance, won a Pulitzer Prize in 2004 for using simple descriptive admission statistics at elite US universities to show how children of rich and powerful alumni and donors were given preferences at the expense of many smart and talented students. Nate Silver's almost perfect prediction of the 2008 and 2012 US election results on his political blog, FiveThirtyEight, earned him a place in Time's 100 most influential figures in the world in 2009, two Webby Awards in 2012 and 2013, and a best-selling book, The Signal and the Noise. In the UK, the aforementioned exposure of expense abuses by MPs in 2009 is another excellent example. When the expense data were obtained, through a long legal battle and a leak, and reported in the Daily Telegraph, the public did not seem to get enough of the story. It shocked the nation, cost the career of prominent politicians, including the then Speaker of the House of Commons, and eventually led to the imprisonment of several MPs and lords.

"Blockbusters" like these might not be the norm in journalism yet, but they can serve as inspiring or textbook-like examples for current and future journalists to see how far data could take them. Although in the day-to-day handling of statistical news, journalists do not often enjoy the high level of resources and skills that such investigations require, these success stories show that statistics are neither cold nor boring, but an endless source of inspiration and information for much excellent, far-reaching and career-changing journalism. With the world's continuing move

to transparency and open-access data in the digital environment, opportunities for such journalism are becoming more open. Indeed, the open data movement, along with the emergence of many user-friendly statistical analysis and presentation tools in recent years, has led some big thinkers of our time (including Tim Berners-Lee, the inventor of the world-wide-web) to come to see data-driven journalism as the future. The first of the three things that Clay Shirky (2014) suggested newspaper journalists should do to save their job and their profession is to "get good with numbers":

> The old "story accompanied by a chart" was merely data next to journalism; increasingly, the data is the journalism. … Learning to code is the gold standard, but even taking an online class in statistics and getting good at Google spreadsheets will help. Anything you can do to make yourself more familiar with finding, understanding, and presenting data will set you apart from people you'll be competing with, whether to keep your current job or get a new one.

This special issue

It is in that context of the vital role of statistics in journalism that we have long been puzzled by the fact that this critical and rather urgent area has received very little attention in journalism scholarship. For one thing, as said from the outset, it is quite inexplicable that journalism education in universities rarely incorporates statistical skills in its curricula, despite repeated calls from experts, scientists and policy makers and despite being better positioned than any news organisation or professional training body to do so. For another, although there is no shortage of useful textbooks or discussion by journalists and commentators, scholarly research in this area remains relatively scarce and, at best, scattered. The emergence of data journalism (DJ) in recent years – and the development of computer-assisted reporting in previous decades – has generated some useful and important research, but much of this is more or less technologically driven, focussing on what computer technologies can help journalists to produce data. What is much needed is research into issues related to the "average" journalist's daily statistics-based work and the many factors that shapes such work. How do they perceive the role of quantitative literacy in societies and the function of statistical skills in newswork? How do they handle the daily influx of data into the newsroom? What sort of frames do they often use to represent

numbers in the news – and what are the potential socio-political impacts of such framing? What good practices are there to promote and what bad practices to avoid? What are the social and institutional conditions that might facilitate or impede such practices? What skills and resources do journalists need to deal effectively with numbers in a resource-poor and deadline-driven environment? The list can go on. Given that quantification plays a crucial role in societies and that statistics have been part of the news agenda for several hundred years, the scarcity of such scholarly enquiry is not less inexplicable than the lack of statistical education for journalists.

This special issue, in that sense, is a long overdue effort to bring together eight articles that can shed more light on some of the issues we raised earlier. Muhammad Idrees Ahmad inaugurates with an insightful case study of body counts in the US-led drone war in Pakistan, in which he demonstrates how and why flawed and dubious statistics can infiltrate the news media and circulate the public sphere with ease. Although less than 4% of the thousands of people killed in drone strikes could be identified as members of Al-Qaeda and 2% as "high value targets", the drone war receives consistently high public support in the US and the UK, where its sanguinary human toll causes "little discomfort among decision makers". Ahmad attributes this to the false "image of a surgical war with little collateral damage" – "a cure-all for terrorism" in Barack Obama's words – that a "docile press" creates and sustains over the years. This, in his account, is due not only to the media's downplaying or ignorance of sufferings in distant places but also, perhaps more importantly, to journalists' statistical credulity. This credulity manifests in (a) an uncritical subscription to the notion of objectivity that privileges official military and political sources and (b) the fetishization of statistics as hard facts that leads to an undue disregard for how the underlying data are collected and compiled. The very duty of the journalist – verifying facts, questioning evidence and dispelling myths – seems to have become a "luxury" in this particular case. Ahmad argues that the responsibility should not lie only with individual journalists. Under continuous deadlines in a fast moving news cycle, "they cannot all be expected to find multiple sources for events happening in remote regions," he notes. "But editors can ensure accuracy by systematically adding caveats to all official claims until they have been corroborated by independent sources."

The quality of statistics in the news depends on the ability – or otherwise – of journalists not only to verify and question data but also to communicate them properly to the audience. Idoia Portilla touches this second aspect in the next article, investigating how five leading Spanish newspapers included essential methodological details in reporting electoral poll results during the Catalonian Parliament election in 2012. Survey/poll data are ubiquitous in daily news, with substantial influences on public opinion. As such, various professional polling associations recommended journalists to include essential methodological information – such as sampling methods, sample sizes, response rates, fieldwork dates, sampling errors, names of research and commissioning organisations, and so on – so that the public can judge the quality and value of the reported data. Even if they assume audiences will not understand everything, Portilla argues, the media should still provide such details and "note at least if the survey conforms to standards" so that the audience can distinguish professional from unprofessional polls. In Spain, such information is even required by the law. But her review of the relevant literature in Spain and other advanced media systems in Europe and America shows that such information remains sparse in the news – and where it is provided, mistakes are not uncommon. Her content analysis of the Spanish newspapers' coverage of poll-based news during the Catalonian election campaign shows a general improvement in the inclusion of such details in news stories, especially when the reported polls are sponsored by the newspapers themselves. But by and large, much remains to be done on both sides of pollsters and the media to ensure that news audiences have adequate and accurate information to judge the reliability and validity of poll data for themselves.

Portilla's conclusion suggests another important determinant of the quality of statistical news: the interaction between journalists and statistical experts (and the institutions they represent). In the next article, Kevin McConway, an academic advisor to BBC Radio 4's More or Less and the UK Science Media Centre, delves into this relationship from a statistician's perspective, offering a useful mix of theoretical perspectives and his personal reflection on working with radio and newspaper journalists. Starting with a brief historical review, McConway attributes the prominence of numbers in the news to "the special status of numerical facts in our society" where "numbers are not construed in the same way as other, equally factual and trustworthy, information". Although statisticians and journalists do similar things at the macro level, he argues, fundamental differences in the way each works exist and need to be mutually understood

and acknowledged if they are to work effectively with each other. Through the "strange" case of news about potential links between mobile phone use and the risk of brain tumours, McConway illustrates how statisticians have "very little idea of how stories actually get into the papers or the broadcast news". He goes on to outline the key differences between the two worlds, calling on statisticians to understand journalism in at least three respects: its timescales (deadline pressures), its agenda (news values and editorial policy), and its pressure to personalise stories from numbers. His "take-home message": statisticians do not have the strengths of journalists – namely telling stories to the right audience in a short space – and "should not simply blame journalists for getting things wrong". Instead, statisticians "must help them to get things right" by being "proactive in making known to journalists what we do, and why and how we do it." While McConway addresses statisticians, his critical reflection – and some practical advice that emerges out of this reflection – bears direct relevance and usefulness for journalists and journalism scholars.

The lack of statistical skills seems to have led journalists to a new challenge in the emerging area of DJ: they have to recede to a minor, auxiliary role in the production process, leaving the lead to those with no traditional journalistic talent. This is among the findings of a study by Constance Tabary, Anne-Marie Provost and Alexandre Trottier in the next article. In examining 178 DJ projects at six major media outlets in Quebec, they found only 40% of the credited authors were journalists, and most of these were involved in background research tasks that require minimal computer and statistical skills. Meanwhile, five key actors with extensive interdisciplinary skills in graphics, programming and statistics – only one of whom has a journalism background – produced and supervised the production of 55% of the studied projects. Thus, instead of encouraging journalists to learn new statistical skills, DJ seems to see the arrival of "a new kind of professionals" with atypical professional profiles in the newsroom. Tabary et al's content analysis also found that most of the studied DJ projects were of little depth, presenting relatively unsophisticated data visualizations and lacking essential elements such as navigation facilities and information exchange with readers. Finally, due to a heavy dependency on pre-processed – often "friendly" – public datasets made available by public institutions, these projects do not stand up as original independent analyses. "In order to develop more meaningful and deep-

digging DJ for the future" and to avoid merely relaying information from powerful institutions, "journalists must control data collection," they conclude.

Despite these caveats, data journalists hold a positive and sanguine view on the potential of their new trade for the media's fourth estate function, according to Tom Felle in the fifth paper. Seeing DJ as a new form of accountability journalism, the 26 data journalists who responded to Felle's interviews strongly identified themselves, through numerous examples, as agents of democratic accountability who employ a novel form of data-based investigation to hold powers to account. In addition, while regarding digital DJ as a continued evolution of CAR, these journalists pointed to a fundamental difference: CAR is merely an investigative tool while DJ is both an investigative tool and a new method of story telling in its own right. Its unique capacity for visualisation and interactivity allows reporters and audiences to interpret, contextualize, examine and analyse news in quite different ways. Some contended that audience engagement tools – such as crowdsourcing, geo-coding and social media – allow for a potentially limitless number of data stories to be told in personalised and localised manners. In response to the criticism that DJ tends to consolidate the position of an already technologically literate elite audience, the interviewed journalists unanimously objected, arguing that DJ indeed engages wider audiences than they traditionally can reach. Felle, however, observes that the "data class" problem might still persist because DJ mainly attracts investment from broadsheet-style news outlets and therefore does not reach far beyond their traditional ABC1 demographic groups.

Shifting the focus from statistics in journalism to statistics in journalism education, the last three articles examine current developments and potential reasons for the presence/absence and successes/failures of teaching statistics in higher education. In a thought-provoking study, Robert Griffin and Sharon Dunwoody found from twin surveys with US journalism chairs that a clear majority valued statistical reasoning skills as a competitive advantage for their students in the job market. Pedagogically, the chairs preferred integrating such skills across the journalism curriculum (rather than offering stand-alone statistics courses). The level of statistical reasoning instruction, however, remained relatively low and saw only a small improvement between the two waves of the study. Possible reasons included a perceived unwillingness and inability of students to acquire statistical reasoning, the shortage of faculty with relevant expertise to teach it,

and the tightness of existing journalism curricula. In addition, both surveys showed that less than three in ten chairs would reward entrepreneurship (i.e. faculty attempts to bring statistical reasoning into their classes), although those who perceived benefits were more likely to bestow rewards, especially if it is consistent with university goals. The authors found some evidence to suggest that such rewards might encourage faculty entrepreneurship. Placing these findings in the context of DJ, Griffin and Dunwoody offered a list of practical take-away advices for journalism educators and conclude: "How ready journalism students are for working in, and leading, the new professional world of data journalism depends on how ready journalism undergraduate programs are to prepare them to do so. That, in turn, may require some programs to change hiring preferences to favor new faculty who can integrate some statistical reasoning instruction into their journalism courses, or reward current faculty for doing the same. SAT data show that journalism students, on the average, are not math dummies. It might be best to avoid treating them as such."

In the next article, Jonathan Hewett offers an in-depth case study of specific pedagogical and institutional obstacles to innovation in DJ education. Drawing on a balanced mix of the literature, a review of stakeholders and course documents and personal reflection, Hewett highlights a whole range of complexities in developing a DJ module for the MA in Interactive Journalism that he directs at City University London. Issues lie not only in the interdisciplinary nature and the particular new demands of DJ per se (which is at odds with journalism educators' well-documented resistance to change and their already heavy teaching and administrative workload), but also many socio-structural factors in our increasingly market-oriented higher education sector. The inclusion of DJ at City did not escape issues such as its market relevance, its appeal to students with primarily arts and humanities background, job/career outcomes, university reputation, student satisfaction, programme management, curriculum coherence, supporting resources (e.g. computer labs and staff time), the uneasy relationship between financial costs and less tangible benefits, and so forth. A notable point from Hewett's study is that students' socialisation with practising data journalists (through guest seminars/lectures, work placements, networking events, freelancing and social media interaction) is not incompatible with educational innovation – i.e. it does not necessarily reinforce old patterns or marginalize new practices as it has been widely criticised. This is because, he argues, DJ is still evolving with more fluid norms

and practices with which journalism educators are unfamiliar. “Rather than holding it back, more specialized socialization could assist journalism education to innovate effectively,” Hewett suggests.

Moving from a single programme to the macro picture, Sergio Splendore, Philip Di Salvo, Tobias Eberwein, Harmen Groenhart, Michal Kus and Colin Porlezza close this special issue with an international comparison of DJ education strategies and approaches across six European journalism cultures (Germany, Switzerland, the Netherlands, Italy, Poland, and the UK). Following Hallin and Mancini’s seminal classification of media systems and combining desk research with interviews with instructors, they found that the peculiarities of national media systems (e.g. professionalism level, educational system, government openness, market forces) affect the extent to which DJ education in each country is developed, the types of its prevailing providers (academic, professional, vocational, or civic organisations), its diversity and quality of courses, its instructors’ academic/professional background, the extent to which expertise from different disciplines can be brought together, and so forth. More professionalized countries, for example, tend to have more structured and interdisciplinary DJ courses by teachers with an educational background in journalism and communication. In all six countries, however, DJ education is still in an experimental stage, lacking financially sustainable models and qualitative standards and neglecting fundamental journalistic topics such as transparency, accountability, ethics and responsiveness. And since the media still see DJ primarily as “a cost rather than an investment” and lack a genuine “open journalism” philosophy, there persists a disincentive for further embracement of DJ education initiatives.

All in all, we hope that this special issue will “set the scene” and serve as an invitation to much more research into this increasingly critical area of journalism and public life.

Acknowledgements

We are grateful to the HE Academy and the Royal Statistical Society for the funding and institutional support that made possible the research symposium from which this special issue arose. A snapshot of the symposium and its presentations can be found at

https://www.heacademy.ac.uk/teaching-how-use-and-communicate-statistics-journalism-students-statistics-journalism-practice-and.

References

Battersby, M. (2010). Is That a Fact? Peterborough, Ontario, Canada: Broadview Press.

Dunwoody, S., & Griffin, R. J. (2013). "Statistical reasoning in journalism education." Science Communication, 35(4), pp. 528-538.

Esposito, J. L., & Mogahed, D. (2007). Who Speaks for Islam? What a Billion Muslims Really Think. New York: Gallup Press.

Geier, Kathleen (2012). "On the importance of statistical literacy." Washington Monthly, May 12. Online: http://www.washingtonmonthly.com/political-animal-a/2012_05/on_the_importance_of_statistic037307.php (accessed on March 10, 2013).

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). "Helping doctors and patients make sense of health statistics." Psychological Science in the Public Interest, 8 (2), pp. 53-96.

Harrison, S. (2014). "History of numeracy education and training for print journalists in England." Numeracy, 7 (2). Online: http://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1154&context=numeracy (accessed on March 15, 2015).

Hope, W. (2011). "Global Financial Crisis| Time, Communication and Financial Collapse." International Journal of Communication, 4 (2). Online: http://ijoc.org/index.php/ijoc/article/view/799 (accessed on March 10, 2013).

Kleinnijenhuis, J., Schultz, F., Oegema, D., & van Atteveldt, W. (2013). "Financial news and market panics in the age of high-frequency sentiment trading algorithms." Journalism: Theory, Practice & Criticism. 14 (2), pp. 271-291.

Kovach, B. & Rosenstiel, T. (2001). The Elements of Journalism. London: Guardian Books.

Livingston, C. and Voakes, P. S. (2005). Working with Numbers and Statistics: A Handbook for Journalists. New Jearsey: Lawrence Earlbaum Associates, Publishers.

Maier, S, R. (2002). "Numbers in the news: a mathematics audit of a daily newspaper." Journalism Studies. Vol. 3, No. 4, Pages 507-519.

Moore, G. and McCabe, D. (2003). Introduction to the Practice of Statistics. London: W H Freeman & Co.

Nguyen, A. (2013). "Online news audiences: the challenges of web metrics." In Karen Fowler-Watt & Stuart Allan (eds.), Journalism: New Challenges. Bournemouth University Centre for Journalism and Communication Research. Online: http://eprints.bournemouth.ac.uk/20929/1/The%20challenges%20of%20web%20metrics.pdf.

Peachey, K. (2010). “It's official ... Do surveys reflect our lifestyle?” BBC News Online. http://news.bbc.co.uk/1/hi/business/8462375.stm (accessed on March 1, 2015).

Potter, D. (2009). “News with numbers.” Advancing the Story. Online: http://www.advancingthestory.com/2010/01/21/news-with-numbers/ (accessed on March 1, 2015)

Randall, D. (2000). The Universal Journalist. London: Pluto Press.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). “How numeracy influences risk comprehension and medical decision making.” Psychological Bulletin, 135, pp. 943–973.

Schlesinger, P. & Tumber, H. (1994). Reporting Crime: The Media Politics of Criminal Justice. Oxford: Clarendon Press.

Shirky, Clay (2014). Last call: the end of the printed newspaper. Medium. Online: https://medium.com/@cshirky/last-call-c682f6471c70 (accessed on March 1, 2015).

Utts, J. (2010). “Unintentional lies in the media: Don’t blame journalists for what we don’t teach.” Invited paper for the International Conference on Teaching Statistics, Slovenia. Online: http://iase-web.org/documents/papers/icots8/ICOTS8_1G2_UTTS.pdf (accessed on January 30, 2015).

Utts, J. (2002). “News and numbers: A guide to reporting statistical claims and controversies in health and related fields.” The American Statistician, 56(4), 330-331.

Wilby, P. (2007). “Damn journalists and statistics.” The Guardian. Online: http://www.theguardian.com/media/2007/nov/05/mondaymediasection.pressandpublishing (accessed on February 12, 2015).

Yarnall, L., Johnson, J. T., Rinne, L., & Ranney, M. A. (2008). “How post-secondary journalism educators teach advanced car data analysis skills in the digital age.” Journalism & Mass Communication Educator, 63(2), 146-164.

[1] The term “statistics” has two broad meanings. First, as a non-countable noun, it refers to a discipline – the science and practice of developing human knowledge through the use of numerical empirical data, based on statistical theory. It encompasses all techniques and procedures for analysing, interpreting and displaying data as well as making decisions based on data. Second, “statistics” is – as used in this article – the plural form of “statistic”, a piece of number-based information such as mean, median, mode, frequency, correlation and so on.

[2] This “number phobia” is, to be fair, not unique to journalists: it is common among other professionals in the liberal and creative arts and even some scholars in social sciences and humanities. For the news profession, this fear – or at least a lack of self-confidence – is exacerbated by the fact that some experts have come to assume that journalists can never get numbers right.

[3] Interview with the authors on May 11, 2010 in Edinburgh.