Introduction

Exciting Times in the Shadow of the 'Post-Truth' Era: News, Numbers and Public Opinion in a Data-Driven World

An Nguyen, *Bournemouth University, UK*

Exciting times

Not so long ago, it would have sounded bizarre to predict that journalism – the profession that often 'gets a bad press' for its permanent hostility to and poor use of numbers – would soon engage with numbers and statistics as an essential daily routine. A browse through online and print news pages today, however, can give an immediate sense of something of an ongoing transformation: from the common most-read/most-viewed lists to the increasingly rich and diverse menu of fancy infographics, interactive data maps and so on, data are now behind many things journalists are producing and delivering to news consumers. As one enters a newsroom, that transformation becomes even more readily observable. Computer screens are jam-packed with web metrics (audience analytics), with continuous streams of such data being beamed and projected in real time onto large and prominently positioned 'performance panels' (Nguyen 2013). Editors keep 'a hawk eye' on those click and view figures behind their most-read/most-viewed/most-shared lists to make news judgement 'on the fly' around the clock (MacGregor 2007), with many having established the routine of beginning news meetings with a rundown of audience data. Here and there, an innovation mood can be felt, with newspeople across soft and hard content areas being busy experimenting with new data-driven storytelling tools. The physical space of some newsrooms has been reconfigured accordingly to accommodate new operations and teams, such as audience strategies (or audience development) and data journalism, and new workflows are created for non-news professionals – such as statisticians, data scientists, programmers and app developers – to sit and work alongside traditional newspeople.

Such are only a few sketches of the change that digital technologies are bringing to journalism's relationship with numbers and statistics. On the one hand, as seen above, the accelerating datafication of society – that is, the use of digital technologies to collect, record, store, monitor and quantify human behaviours and activities for various social, economic and political purposes – has resulted in the intensive use of audience metrics in the journalistic gatekeeping process. On the other, with the

world's move to transparency and open-access data in the digital environment and the recent introduction of simple, user-friendly data analysis and visualization software, the relatively (not entirely) new genre of 'data journalism' has enjoyed quite a surge, especially since the beginning of this decade. Both have been hailed as essentials for journalism to go into the future. Improving the use of metrics in the newsroom, for example, was seen as 'very important' by 76 per cent – and 'somewhat important' by another 21 per cent – of senior news executives and digital strategists around the world who responded to a study by Newman (2016). Another recent survey put data journalism as the third most sought additional training area by American journalists – below video shooting/editing and social media engagement but above twenty other categories (Willnat and Weaver 2014). Tim Berners-Lee, the inventor of the World Wide Web, declared in 2010 that 'data-driven journalism is the future':

> Journalists need to be data-savvy. It used to be that you would get stories by chatting to people in bars, and it still might be that you'll do it that way some times. But now it's also going to be about poring over data and equipping yourself with the tools to analyze it and picking out what's interesting. And keeping it in perspective, helping people out by really seeing where it all fits together, and what's going on in the country.[1]

Similarly, Clay Shirky (2014) points out that, unlike 'the old "story accompanied by a chart" (that) was merely data next to journalism, increasingly, the data *is* the journalism'. Hence, the first thing that Shirky (2014) advised newspaper journalists to do to save their job and their profession is to 'get good with numbers'. If journalists want to stand out in this brave new world, in his views, they will have to learn to code, take online classes in statistics, and familiarize themselves 'with finding, understanding, and presenting data'.

It is, in short, exciting times ahead.

With this potential future in mind, however, one ought to ask some hard questions. For this 'quantitative turn in journalism', as Petre (2013) calls it, to make sense and to benefit us all in the long term, it needs to be understood and explored in the historical context in which it takes place. However powerful this quantitative turn is, its socio-historical context will play a critical role in determining the extent to which its power can be unleashed and materialized for the sake of a better informed, more self-governed citizenry. Here, I would propose an examination from the perspectives of three key stakeholders in news production and communication: the public, statistical sources, and journalists as the middleperson between them.

Numbers as a staple of modern life

Let us start with the public. The first thing is to go beyond the hype to recognize that statistics have been a critically important part of our daily life for a very long time. As

[1] *The Data Journalism Handbook*, http://datajournalismhandbook.org/1.0/en/introduction_2.html.

early as 1903, when statistics had not even established itself as an academic discipline, Herbert G. Wells had already predicted in *Mankind in the Making* that statistical skills would soon become as essential as reading and writing for efficient citizenship:

> The great body of physical science, a great deal of the essential fact of financial science, and endless social and political problems are only accessible and only thinkable to those who have had a sound training in mathematical analysis, and the time may not be very remote when it will be understood that for complete initiation as an efficient citizen of one of the new great complex world-wide states that are now developing, *it is as necessary to be able to compute, to think in averages and maxima and minima, as it is now to be able to read and write.* (Wells 1903, 189, my emphasis)

Such prophecy would have materialized over the course of the twentieth century, with statistics having now reached a point where they are 'part of the fabric of the contemporary world' (Nguyen and Lugo-Ocando 2016, 4). Numbers occupy a pervasive position in modern life because almost every key aspect of it – from the quality of the air we breathe to the national leader we choose – is numerically measured and represented in one way or another. In this 'hyper-numeric world', as Andreas and Greenhill (2010, 1) point out from a political perspective, 'if something is not measured, it does not exist' and 'will not be recognized, defined, prioritized, put on the agenda, and debated'. The age of 'big data' means numbers will exercise more power over the way we live and work – but we should not lose sight of the fundamental fact that they, as Lorenzo Fioramonti (2014) puts it, have long 'ruled the world'. In development policy-making, for example, statistics such as GDP or poverty threshold have become an indispensable tool to summarize, justify and/or rubberstamp policies across the globe:

> One extra dollar a day per capita, and the funding for a series of hospitals in India might stop on its track. A dozen more deaths of a certain disease, and the World Health Organisation declares an epidemic. For many in the world, these numbers mean the difference between life and death, poverty and prosperity, or happiness and misery. For others, who have seen Gross Domestic Product grow for decades and their lives remain as miserable as ever, these statistics mean absolutely nothing. (Lugo-Ocando and Nguyen 2017, 43)

As such, statistics have become a constant staple of our daily news diet (Blastland and Dilnot 2008; Maier 2002). A recent content analysis of British broadcasters' news output, for example, found at least one reference to statistics in more than a fifth of the sample, especially among health, politics, business and economics coverage (Cushion et al. 2016). And although there are legitimate concerns that ordinary citizens do not have the knowledge and skills to handle so much numerical information (Utts 2002, 2010), the broad assumption remains that they rely on data reported in the news to exercise some control over the social world, to make rational choices over public affairs and ultimately to maintain their influence over public institutions.

The age of dubious numbers and 'statistical bullshit'

This leads us to the second stakeholder: the people and institutions that produce and/or circulate statistics in the public sphere. The establishment in every society understands the power of numbers very well and tries its utmost to use that power to their own advantage. Today it is rare to find a resourceful political, economic or social institution that does not resort to statistical information as a crucial tool to shape public opinion. Indeed, this is a 'chicken and egg' relationship: as statistics are too central to daily life and daily news, they force such institutions to continuously measure and collect data for just everything within their operational scope to defend, legitimize and/or promote what they do or want to do (Best 2001).

One outcome of this is a deluge of bad statistics being circulated in the public sphere, thanks to both traditional media platforms and, more recently, the very helpful hand of social media and their echo chambers. Many of these numbers are deliberate attempts to win the public debate with dubious and/or misinterpreted numbers – such as the now-famous misleading £350 million per week that the Boris Johnson, Michael Gove and their Brexit followers claimed the UK had to pay for EU membership a way to persuade UK voters to vote to leave EU during the EU Referendum campaign in 2016. Some others are deliberately manipulated or even fabricated to serve some political or commercial interests. All these belong to what Tim Harford (2016) calls 'the rise statistical bullshit – the casual slinging around of numbers not because they are true or false, but to sell a message'. In the world of social media, as Harford notes, such bullshit 'spreads easily these days (because) all it takes is a click', with the bullshitter caring nothing about its truthfulness but its ability to catch attention and/or stir emotion. The bullshitter could even be more dangerous than the liar, because at least the latter cares about and defers to the truth.

That is what some have called the post-fact, post-truth era. Indeed, journalism embarks on its new quantitative journey at a very troubling time for numerical facts, a time overshadowed by the rise of fake news, the emotionally charged populist movements behind Brexit and Donald Trump's controversial election, and the increasing power of the few private companies that control big data through algorithms. Instead of helping to serve as stable reference points and to settle arguments, as sociologist William Davies (2017) puts it in the *Guardian*, statistics have lost their authority:

> Rather than diffusing controversy and polarisation, it seems as if statistics are actually stoking them. Antipathy to statistics has become one of the hallmarks of the populist right, with statisticians and economists chief among the various 'experts' that were ostensibly rejected by voters in 2016. Not only are statistics viewed by many as untrustworthy, there appears to be something almost insulting or arrogant about them. Reducing social and economic issues to numerical aggregates and averages seems to violate some people's sense of political decency.

The situation has reached such a critical point that the highly respected University of Washington has recently started a 10-week course named 'Calling bullshit in the age

of Big Data', which they offered both face-to-face to 160 students and online to the general public (through recorded lectures/seminars and readings).

The crucial but largely forfeited role of journalism

In this increasingly chaotic world, journalists must be adept at using quantitative data to filter the true from the false and ultimately to lead the public in the right direction. Indeed, even without entering the post-truth era of fake news and filter bubbles, one would have expected that role long ago: if statistics can be worse than lies and damn lies, to use that famous old saying, then the logical outcome would be that journalism, the profession that purports to tell the truth and expose the untruth, would have long found in them a fertile land for deep-digging and world-changing investigation. The somewhat odd concurrent rise of big data and 'post-truth' can only highlight that critical importance of journalism in exposing misleading, dubious numbers that are central to public life.

On the other hand, from a more positive side, journalists need to be statistically competent because there is also a great deal of good data and statistics which, if used well as an authoritative form of evidence, could assist them greatly in searching the truth and helping the public to make sense of the complex world around them. This is by no means a new argument: as Stuart Allan says in the Foreword, it was indeed a central point that the great American thinker, Walter Lippmann, made about journalism in the early part of the twentieth century. Not long after H. G. Wells said the above, Lippmann argued ardently in *Public Opinion* for the integration of scientific rigour into journalism, calling newspeople to bring to their truth-seeking and truth-telling function the objective method of science, which is 'based on exact record, measurement, analysis and comparison' (Lippmann 1920, 138). 'It does not matter that the news is not susceptible of mathematical statement,' he said. 'In fact, just because news is complex and slippery, good reporting requires the exercise of the highest of the scientific virtues' (1922, 49). This means, as he pointed out in an earlier book, *Liberty and the News*, journalists would have to possess the ability to '(scribe) no more credibility to a statement than it warrants, a nice sense of the probabilities, and a keen understanding of the quantitative importance of particular facts' (Lippmann 1920, 82).

Flash forward a century and, as hinted at the outset, the ability to work with numbers is still not only an intellectual luxury but sometimes a cursed subject in the news industry. Many journalists simply cannot settle themselves with the idea of scrutinizing or putting statistics to good use in news production: why do they have to deal with numbers, some might ask, while they come into journalism to work with words and to avoid all those eye-numbing figures? 'Journalism is one of the few professions that not only tolerates general innumeracy, but celebrates it,' said Aron Pilhofer, executive director for the digital at the *Guardian*. 'I still hear journalists who are proud of it, even celebrating that they can't do math, even though programming is about logic. It's hard to get a journalist to open up a spreadsheet, much less open up

a command line. It is just not something that they, in general, think is held to be an important skill' (as quoted in Howard 2014, 46). Meanwhile, statistical training has been largely ignored by both universities and industry accreditation bodies – and the recent rise of big data has not changed that mindset in any substantial way (Howard 2014, Nguyen and Lugo-Ocando 2016).

In that context, it is perhaps not surprising that nearly two-thirds of 4,300 sampled references to statistics in the above study of UK broadcasters' news output were either vaguely mentioned or were clearly presented but with little context, while only 4.2 per cent of them were challenged in some way (Cushion et al. 2016). By the same token, early research into data journalism has found it to be done on a superficial level, in limited scopes and formats, with 'as much decorative as informative' impact (Knight 2015, 55). Neither is it to see that the key drivers of most data journalism projects are not traditional newspeople but those with non-journalism skills and experiences, such as programmers, developers, designers and producers (Tabary, Provost and Trottier 2016). Some influential names that have triggered the interest in data journalism, such as Nate Silver or Ezra Klein, had never worked as journalists before entering the news scene (Nguyen and Lugo-Ocando 2015).

All this begs the question: could journalists ever 'get good with numbers' – to not only make the most from the potential of 'data journalism' and audience metrics but also, and more importantly, to effectively handle, use and communicate data and statistics in general?

My personal answer is a resolute yes – and it must be a resolute yes. The traditional ignorance of statistics in journalism training and education is based on too many incorrect assumptions and prejudices about journalism's relationship with numbers. As Sisi Wei, a news app developer at *ProPublica*, puts it:

> I wish that no j-school ever reinforces or finds acceptable, actively or passively, the stereotype that journalists are bad at math. All it takes is one professor who shrugs off a math error to add to this stereotype, to have the idea pass onto one of his or her students. Let's be clear: journalists do not come with a math disability. (quoted in Howard 2014)

Which is quite true: research has shown clearly that beginning journalism students have no less ability to deal with numbers than those in other disciplines – including those in the natural sciences (Dunwoody and Griffin 2013; Maier 2002; Nguyen and Lugo-Ocando 2016). As Pilhofer (ibid.) commented: 'Journalism programs need to step up and understand that we live in a data-rich society, and math skills and basic data analysis skills are highly relevant to journalism.' All that is needed is the right statistical training approaches that (a) place a total confidence in the ability of journalism students to handle numerical and statistical information and (b) equip them with the knowledge and skills not only to source, verify and scrutinize data but also to effectively present and deliver them to the lay public. The ultimate aim of any such training is for journalists to be able to help citizens to relate themselves to the world beyond them and to vanquish the true from the false, the gold from the

bullshit. It would require not just statistical skills but also in-depth knowledge and understanding about the socio-psychological interaction between news, numbers and public opinion and its socio-political impact in today's data-driven world. It is time for media and journalism scholars, statisticians and the news profession to get together to develop a strong knowledge base for this task.

This book's intervention

It is in that spirit and towards that aim that this book comes together. While there is no shortage of practical guidelines and textbooks on statistical skills for journalists and the general public, scholarly research into the area remains relatively scarce and, at best, scattered. How do news, numbers and public opinion interact with each other? Is it really the case, as common wisdom tells us, that most citizens and journalists do not have the necessary cognitive resources to critically process and assess data and statistics? What strategies do journalists use to gather and communicate statistical data in their stories? What skills do they need to deal effectively with the daily influx of numbers in the resource-poor and deadline-driven newsroom? What sort of formats and frames do they often use to represent different types of data? What are the potential socio-psychological and political effects of such formats and frames on public information, knowledge acquisition and attitude formation? What good practices are there to promote and what bad ones to avoid – and how?

Gathering some of the most established scholars as well as emerging names in both statistics and media and journalism studies, this book offers a comprehensive, authoritative, first-of-its-kind collection of past and present research into the issues above. As represented in the word cloud below (generated from the book's manuscript

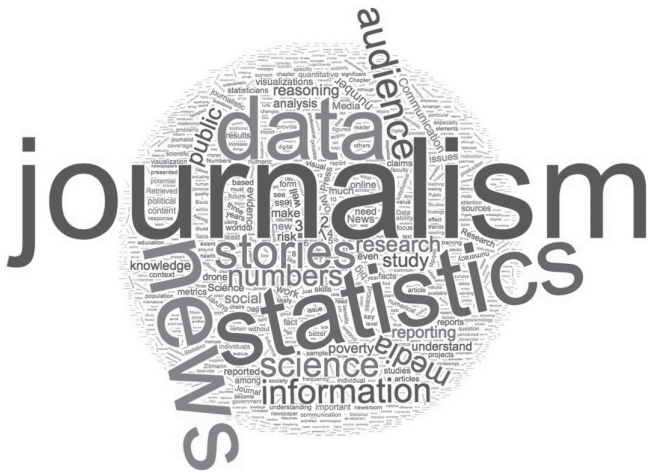

Figure 0.1 A simple word cloud featuring the key themes and topics of this book (created from its book manuscript).

with the free but powerful tools on wordclouds.com), our focus will be on how journalism makes use of data and statistics, how such numerical information is told in news stories, and how all this affects different types of audiences and publics. Towards that purpose, the book is divided into three sections, examining how numbers are received, treated and handled by at both ends of the news communication process – the journalist in the newsroom and the news consumer in the public at large – with a particular attention to their effects on public reasoning and deliberation.

Section 1: Data and statistics in news production

The first section consists of seven chapters that offer theoretical and empirical insights into how and why data and statistics are gathered, used and represented in certain ways in the news. Chapter 1 is a personally informed account of an engaged statistician with (often frustrating) experience with journalists in Switzerland. Fabienne Crettaz von Roten starts with an inventory of common statistical errors in the news before outlining three key roles that, in her view, journalists should assume but have largely failed to fulfil in the statistics – society relationship – namely verifying and scrutinizing statistics, communicating statistics in an accurate manner, and educating the public about statistics. She identifies the lack of statistical training in journalism education as part of the problem and concludes with some thoughts on potential solutions for that shortage, with particular attention to whether to teach statistical reasoning as a separate course or as embedded elements across different courses in the journalism curricula.

Chapter 2 offers some empirical insights of the overall quality of statistical news reporting, based on results from the aforementioned BBC Trust-funded review of statistics in the UK broadcast media. Stephen Cushion and Justin Lewis look deeply into the current state of everyday reporting of statistics across radio, television and online news platforms, both with a quantitative overview and two qualitative case studies. Although they found that 'statistics were routinely invoked in vague and imprecise ways, with limited context and explanation about their meaning or method behind them', some rare good-practice examples that they identified suggested that journalists would do a far better service to the public by rethinking and reinterpreting their (sometimes problematic) impartiality ethos in a way that enables and prepares them to scrutinize and challenge statistics whenever necessary.

Chapter 3 provides another insightful case study, examining the uncritical news reporting of body-count statistics during the US-led drone war in Afghanistan in the past fifteen years. Muhammad Idrees Ahmad, like Cushion and Lewis above, demonstrates how journalists, overshadowed by their 'strategic ritual of objectivity' and constrained by the difficulty to find accurate numbers about drone casualties, disregard questions about the validity of official figures from the military. In reporting drone-induced body counts to a remote passive public without knowing, for example, they included more terrorists or more civilians, journalists have greatly contributed to a long-running political and military attempt to use dubious and imprecise casualty statistics to legitimize the drone war. The consequence has been a false public image

of drones as a panacea for security problems, leading to a dramatic shift in public attitudes towards extrajudicial killing, with dangerous consequences.

Chapter 4 examines the political communication of statistics in the context of news coverage of poverty. At the heart of Jairo Lugo-Ocando and Brendan Lawson's critique is the diminishing ability of journalists to critically analyse and counter data-based hegemonic discourses, leaving too much space for political sources – such as government officials, multilateral governance bodies, NGOs and others – to spin over poverty concepts and numbers. The consistent failure of the media around the world to report and expose PR stunts around poverty statistics, according to the authors, can be attributed to not only the lack of an agreed definition of poverty between its key stakeholders but also, again, to the lack of statistical preparation among newspeople and decreasing resources in the newsroom. The result is an over-reliance on official sources, who are allowed to get away easily with their own, often flawed and ideologically driven, explanatory frameworks. With that, Lugo-Ocando and Lawson conclude the purported function of journalism as a mediator in development policy formulation and a watchdog of democratic decision-making has been ineffective.

Chapter 5 moves the discussion to another area where statistics are an inevitable part of the day: science journalism. Renata Brandao and An Nguyen start with a review of journalism's traditional deference to statistics from cultural and epistemological perspectives before citing a number of good reasons to hope that science journalism might be an exception to that relationship. However, their ensuing review of the past and present of science journalism – particularly its traditional cheer-leading approach to science, with its focus on promoting science literacy and the consequent failure to equip lay people with the ability to critically engage with and question science – suggests that that hope might be more like a dream. Employing a secondary content analysis of 1,089 science articles in four leading UK and Brazil broadsheets, they found evidence that although journalists did not deprive readers of the opportunities to understand science statistics in the social context, these statistics were presented mostly as unquestionable and undisputed facts, without the essential methodological information that readers need to critically understand, judge and, if necessary, raise their voices over science developments.

Chapter 6 explores the latest trend in the news – numbers relationship, namely the evolving shape of the new genre of data journalism. Using a content analysis of projects nominated for the Global Editors Network's annual Data Journalism Awards (DJA) between 2013 and 2015, Julius Reimer and Wiebke Loosen explore if, and how, data journalism at its finest has changed over the years in terms of, among others, data sources and types, visualization strategies, interactive features, topics, and types of nominated media outlets. Their results suggest that although data journalism is increasingly personnel-intensive and progressively spreading around the globe, the set of structural elements and presentation formats that the DJA-nominated pieces are built upon remain rather stable and restricted to relatively simple forms. In addition, although half of the sampled projects assume a watchdog role (i.e. with elements of criticism or calls for public intervention), this role is limited as journalists rely mainly

on publicly available, thus institutionally friendly, datasets. There is, however, evidence that they are increasingly looking to non-institutional data sources for their stories.

Chapter 7 deals with a very different but fast-growing use of statistics in the newsroom: the use of digital audience-tracking data in news production and distribution. Against the backdrop of the traditionally powerless position of the audience in the journalistic gatekeeping process, An Nguyen and Hong Tien Vu review the recent unprecedented penetration of web metrics into the newsroom, examining its key social and technological drivers as well as its potential impacts on news decision-making processes and journalists' work ethos and autonomy. As news moves from being exclusively 'what newspapermen make it' to also something that the crowd wants it to be, the authors demonstrate that newsroom processes and relationships are being transformed in ways that invite more misgivings and reservations than hopes and innovations. In observing the latest developments, however, the authors conclude with a more positive and optimistic note, introducing a range of promising audience data policies and measures that might help journalists to go beyond head counting and get more deeply into news engagement metrics to harness their power for a better and more sustainable future for journalism.

Section 2: Data and statistics in news consumption

The second section examines data and statistics from the news audience's perspective, with five richly insightful chapters on how statistical information in the news interacts with individuals' reasoning, knowledge acquisition and attitude/belief formation in their public and private life.

Scott Maier starts with a thought-provoking chapter that challenges our common beliefs about the power of numbers in the news. Statistics are to trigger rational action but their impact on human emotion could be big and undesired. Maier proves it through the case of news about humanitarian crises, where experimental psychology research shows that statistical information can actually diminish empathy and discourage humanitarian response. This 'psychophysical numbing' effect (lives are valued less as their numbers increase) was tested and confirmed in his research, including a year-long examination of reader responses to the use of personified numbers by *New York Times* columnist Nicholas Kristof and an experimental study on how numbers-based stories differ from other storytelling forms in their influences on reader emotions. The cautionary lesson he offers: the news media might select some statistics to raise awareness but should not rely solely on them to drive responses to calamity and mass injustice. Embedding personal stories in numerical news reporting might be the way, he concludes.

That, however, comes with potential issues that Chapters 9 and 10 deal with. In Chapter 9, Charles Berger offers an authoritative review of research into the effects of a regular use of anecdotal cases and quantitative data to depict threatening trends (e.g. traffic facilities, crime and health problems) in the news. He begins with a discussion of how news consumers, in the absence of relevant statistical data, use mental shortcuts and heuristics – such as extraordinary cases and non-representative

anecdotes that are regularly provided in the news to illustrate hazards and threats – to make wrong inferences about the general trends/risks that such anecdotes depict. Even when journalists present quantitative data in lieu of anecdotes, however, there persist many problems, especially with respect to their widespread, advertent or inadvertent, use of mere frequencies rather than base-rate data in depicting threatening trends. Through an analysis of some recent examples in the US media and a thorough review of experimental research, Berger shows how such statistical depiction may encourage news consumers to form distorted estimates of their risk of exposure to hazards and threats. This, as he reviews, leads to many individual and social consequences in a public with generally low numeracy, such as excessive anxiety (constituting a public health problem), attenuated achievement expectations and increased risk aversion.

In Chapter 10, Rhonda Gibson and Coy Callison delve even more deeply into the rational and emotional responses of news consumers to numeric information and exemplar cases, with attention to how these responses vary across different quantitative literacy levels and different numerical presentation formats. Following a brief introduction to exemplification theory, which predicts how base-rate data and exemplars in news reports affect audience perception of the reported issues, the authors review research that measures the effectiveness of various data formats – such as percentages, fractions, probabilities, and primary versus secondary ratios. The second part of the chapter addresses more recent research into (a) the effects of an individual's quantitative literacy level on his/her willingness and ability to process numeric information in news reports, (b) the rational impressions and affective reactions that news users with different numeric ability form after exposure to statistical information and/or exemplifying cases about risk-related topics, (c) the subsequent effects on those individuals' levels of empathy with victims, their personal risk assessment as well as their assessment of risk to others. The chapter ends with useful recommendations for journalists and others who create statistical messages designed to inform the public.

If Chapters 9 and 10 raise issues about the potential undesired effects of statistical information in the news, Chapter 11 seeks an explanation for the so-called number paradox: the fact that the media use a lot of numbers, even though it is known that news consumers do not usually recall, understand or appreciate such numbers in news articles. Willem Koetsenruijter first reviews the relevant research literature to demonstrate that the impact of numbers on the way news consumers recall and perceive news facts might not be as strong as expected. Assuming that reporters are reasonable actors with purposes behind everything they do in news making processes, the author hypothesizes that journalists use numbers as a rhetorical device to appear credible, that is, they use the authoritative status of statistics as objective and scientific facts to boost their ethos in the mind of news consumers. Koetsenruijter presents a series of experiments designed to test this hypothesis in which he found enough evidence to conclude that the use of numerical information did increase the perceived credibility of the news both in newspapers and on television.

Chapter 12 reports an innovative, comprehensive mixed-method study on how news consumers use and evaluate infographics, the major distinctive form of the new data journalism genre. Yael de Haan, Sanne Kruikemeier, Sophie Lecheler, Gerard Smit

and Renee van der Nat first conduct an eye-tracking study to explore the extent to which news consumers pay attention to and use data visualizations in three different news modalities (a print newspaper, an e-newspaper on a tablet and a news site in the Netherlands). Based on the results of this part, they then use the same material to conduct focus groups and an online survey to investigate whether news consumers appreciate and value data visualizations. The results show that news consumers do read news visualizations, regardless of the platform on which the visual is published. In addition, visualization elements are appreciated, but only if they are coherently integrated into a news story and fulfil a function that can be easily understood. Given the scarcity of research into this area, the chapter will provide the first comprehensive picture of the usefulness of infographics in the news and contribute to a growing literature on alternative ways of storytelling in journalism today. Its innovative mixed-method approach also suggests some fresh ideas for future research into the use and effect of new storytelling formats in journalism and beyond.

Section 3: Agenda for the future

The third section offers some forward-looking perspectives on what can be done to improve the status quo of the journalism, statistics and society relationship and how we might theoretically and empirically approach to improve this relationship in the age of big data.

In Chapter 13, Kevin McConway, an academic adviser to BBC Radio 4's More or Less show and the UK Science Media Centre, offers a useful mix of history, theory and personal experience from working with journalists to discuss how the interaction between statisticians and journalists can be more fruitful. Although they do similar things at the macro level, he argues, fundamental differences in the way each profession works exist and need to be mutually understood and acknowledged if they are to work effectively with each other. Reflecting on the strange case of news about potential links between mobile phone use and the risk of brain tumours, McConway 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 personalize stories from numbers. His 'take-home message': statisticians need journalists as they do not have the strengths of journalists – namely telling stories to the right audience in a short space. Therefore, they 'should not simply blame journalists for getting things wrong' but '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'.

Chapter 14 looks to the future with a perspective from inside the newsroom, examining the potential role of science journalism as a possible bridge between data and journalism. Holger Wormer charts the recent rise of data journalism – from the tendency to tell stories in 'charticles' in the late 2000s to the growing use of colourful and interactive graphics in online media – as an extension of the media's traditional 'love affairs' with numbers. At the same time, he argues, there is not any real evidence that journalists in general have significantly improved their knowledge on data and

statistics in recent years. Neither is it clear whether data journalism projects are more about a truthful appearance than substance, that is, whether they can offer something really meaningful from a statistician's point of view and be better than the simplistic 'he-said-she-said' journalism formula. In exploring these issues, Wormer argues that one possible bridge to connect both sides could be the traditional cadre of systematically educated science journalists who are familiar with journalistic needs and skills (between news selection, classical investigation and storytelling) but also with basic principles of statistics and scientific methods. From this perspective, science journalism education strategies may be not only a key to improve the strange relationship between journalists and statistics but also predestined for a 'second generation data journalism'.

Chapter 15 focuses on the current place of statistical reasoning in journalism education to propose some strategic directions for the teaching of these skills in the future. Robert Griffin and Sharon Dunwoody showed from findings of twin surveys with US journalism chairs in 1997 and 2008 that a clear majority valued statistical reasoning skills as a competitive advantage for their students in the job market. The level of statistical reasoning instruction, however, remained relatively low and saw only a small improvement between the two waves of the study. Although these surveys date from, respectively, one and two decades ago, the authors yield some useful and generalizable points that bear fresh implications for today and the future. For journalism education to be at the helm in the 'next big thing' of data journalism and to equip journalists with evaluative capacities to align truth claims with evidence to suppress fake news, universities will need to address the many factors that the surveys found to be behind the past and present low-key profile of statistical training in journalism curricula – namely the perceived unwillingness and inability of students to acquire statistical reasoning, the shortage of faculty with relevant expertise, the tightness of existing journalism curricula and other structural factors, and the lack of reward for faculty entrepreneurship (i.e. attempts to bring statistical reasoning into their classes). 'But getting there', they warn from their decadal surveys, 'will be challenging' because of the very slow pace of changes in academic institutions and requires much entrepreneurship on part of journalism chairs themselves.

The last chapter shifts the focus to journalism research, with Oscar Westlund and Seth Lewis offering a set of conceptual and theoretical toolkits to approach the increasing datafication of society and its potentially transformative consequences on how news is perceived and practised. Their chapter presents four conceptual lenses on big data and journalism: *epistemology* (the legitimization of new journalistic claims about knowledge and truth based on big data); *expertise* (the negotiation of occupational status, authority, and skill sets as new specializations are developed and deployed in the newsroom); *economics* (the potential for and challenges with new efficiencies, resources, innovations, value creations and revenue opportunities that big data might induce for journalism); and *ethics* (ethical issues raised for the norms and values that guide human decision-making and technological systems design). While each of these addresses relevant questions for news media and helps to guide future research, Westlund and Lewis choose to devote the rest of the chapter to the current

research literature on the epistemology of big data and journalism and future directions for such research. Their investigation can never be more timely in this post-truth moment. As the authors argue themselves, exploring the epistemology of journalism and big data is in itself 'an opportunity to understand how information is legitimated as "real" or "fake" to make sense of journalism's co-production of knowledge with publics, to identify dynamics of information production and circulation that lead to various interpretations on the part of audiences'.

All in all, it is our ambition that this book will make a fresh contribution and invite further research to an enquiry area that should have received much more scholarly attention in the literature. In addition, we hope that the book offers the news profession some systematic perspectives and principles to build a pragmatic framework for a more effective, more fruitful interaction between journalism, statistics and society. We have never needed such theoretical and practical work as badly and urgently as we do today.

References

Andreas, P. and Greenhill, K. M. (2010), *Sex, Drugs, and Body Counts: The Politics of Numbers in Global Crime and Conflict*. Ithaca, NY: Cornell University Press.

Best, J. (2001), *Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists*. London: University of California Press.

Blastland, M. and Dilnot, A. (2008), *The Tiger That Isn't: Seeing through a World of Numbers*. London: Profile Books.

Cushion, S., Lewis, J., Sambrook, R. and Ghallagan, R. (2016), Impartiality reporting of BBC reporting of statistics: A content analysis. Retrieved 30 September 2016 from http://www.bbc.co.uk/corporate2/bbctrust/our_work/editorial_standards/impartiality/statistics.

Davies, W. (2017), 'How statistics lost their power'. *The Guardian*, 19 January. Retrieved 19 January 2017 from https://www.theguardian.com/politics/2017/jan/19/crisis-of-statistics-big-data-democracy.

Dunwoody, S. and Griffin, R. J. (2013), 'Statistical reasoning in journalism education.' *Science Communication*, 35(4), 528–38.

Fioramonti, L. (2014), *How Numbers Rule the World: The Use and Abuse of Statistics in Global Politics*. London: Zed Books.

Harford, T. (2016), 'How politicians poisoned statistics', http://timharford.com/2016/04/how-politicians-poisoned-statistics/ (accessed 15 May 2016).

Howard, A. (2014), *The Art and Science of Data-Driven Journalism*. Tow Centre for Digital Journalism. Retrieved 12 February 2015 from http://tinyurl.com/jmxskoz.

Knight, M. (2015), 'Data journalism in the UK: A preliminary analysis of form and content.' *Journal of Media Practice*, 16(1), 55–72.

Lippmann, W. (1920), *Liberty and the News*. New York: Harcourt, Brace and Howe.

Lippmann, W. (1922), *Public Opinion*. New York: Pearson Education.

Lugo-Ocando, J. and Nguyen, A. (2017), *Developing News: Global Journalism and the Coverage of 'Third World' Development*. London: Routledge.

MacGregor, Phil (2007), 'Tracking the online audience.' *Journalism Studies*, 8(2), 280–98.

Maier, S. R. (2002), 'Numbers in the news: A mathematics audit of a daily newspaper.' *Journalism Studies*, 3(4), 507–19.

Newman, N. (2016), Journalism, Media and Technology Predictions 2016. Oxford: Reuters Institute for the Study of Journalism. Retrieved 30 April 2016 from http://tinyurl.com/h86ak9t.

Nguyen, A. (2013), 'Online news audiences: The challenges of web metrics.' In Karen Fowler-Watt and 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.

Nguyen, A. and Lugo-Ocando, J. (2015), 'A vaccine against that anti-data journalism brain.' *Data-Driven Journalism*, 4 November. Retrieved 4 November 2015 from http://datadrivenjournalism.net/news_and_analysis/a_vaccine_against_that_anti_data_journalism_brain.

Nguyen, A. and Lugo-Ocando, J. (2016), 'The state of data and statistics in journalism and journalism education: Issues and debates.' *Journalism*, 17(1), 3–17.

Petre, Caitlin (2013), 'A quantitative turn in journalism?' Retrieved 1 March 2015 from http://towcenter.org/a-quantitative-turn-in-journalism/.

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

Tabary, C., Provost, A. M. and Trottier, A. (2016), 'Data journalism's actors, practices and skills: A case study from Quebec.' *Journalism: Theory, Practice, and Criticism*, 17(1), 66–84.

Utts, J. (2002), 'News and numbers: A guide to reporting statistical claims and controversies in health and related fields.' *The American Statistician*, 56(4), 330–31.

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*. Retrieved 30 January 2015 from http://iase-web.org/documents/papers/icots8/ICOTS8_1G2_UTTS.pdf.

Wells, H. G. (1903), *Mankind in the Making*. London: The Echo Library. Retrieved 20 April 2014 from http://www.gutenberg.org/ebooks/7058.

Willnat, L. and Weaver, D. (2014), *The American Journalist in the Digital Age: Key Findings*. Bloomington, IN: School of Journalism, Indiana University. Retrieved 20 April 2016 from http://news.indiana.edu/releases/iu/2014/05/2013-american-journalist-key-findings.pdf.