A Numbers Game: Two Case Studies in Teaching Data Journalism

Greg Treadwell[1], Tara Ross[2], Allan Lee[1], and Jeff Kelly Lowenstein[3]

Abstract

Influenced by the practices of social scientists, data journalists seek to create stories that frame social reality through quantitative data analysis. While the use of statistics by journalists is not new, exponential growth in available data and a desire for source material unmediated by political and public-relations framings have seen data journalism increasingly embraced—to varying degrees—by newsrooms, and editors increasingly seek reporters who can think in computational ways. Journalism programs keen to incorporate data journalism in curricula face a unique set of issues, including a lack of scholarship on data journalism education and how to teach it. This article reports on both the pilot of an international postgraduate collaboration in data journalism education in 2015, in which postgraduate students at two universities investigated state-run gambling in Aotearoa–New Zealand, and the introduction of an undergraduate semester-long paper in data journalism at one of the universities. A visiting Fulbright specialist supported both initiatives, helping to develop staff and student data skills, kick-start a joint investigation by students, and lay the groundwork for future international collaborations. Thanks to his visit, New Zealand educators and students were able to seek support from, a global community of journalists and journalism educators working in data journalism. Set against a literature that predicts an increasing role for computational journalism, this article explores the successes and challenges of these cases of experiential journalism education. It explores the complex but not fatal issues of data competency among both instructors and students, collaboration between geographically distinct programs, access to sensitive datasets, and publication of student work.

[1]Auckland University of Technology, New Zealand
[2]University of Canterbury, Christchurch, New Zealand
[3]Columbia College, Chicago, IL, USA

Corresponding Author:
Greg Treadwell, Auckland University of Technology, P.O. Box 92006 Wellesley St, Auckland, 1010 New Zealand.
Email: gregory.treadwell@aut.ac.nz

Introduction

The strengthening role of data journalism (DJ) in investigative practice has precipitated a growing need in newsrooms for journalists capable of producing multimedia stories through data analysis. While arguably derivative of Philip Meyer's groundbreaking theories of "precision journalism" (e.g., Meyer, 1991, 2002) and the computer-assisted reporting (CAR) of the 1990s, DJ requires a higher level of digital and data literacy to make use of the plethora of interactive Web 2.0 data applications now available and "to navigate multiple public and private databases, and represent the complex information therein in interesting, relevant and accessible ways" (Flew, Spurgeon, Daniel, & Swift, 2012, p. 161).

U.S. journalism educator Kathleen Bartzen Culver (2014) called the use of data "arguably the single most important trend in media work today." Certainly, its ability to hold power to account has been ably demonstrated in stories such as those on the United Kingdom's MP-expenses scandal in 2009, the *Wall Street Journal*'s Perfect Payday story, which exposed the manipulation of bonus-related stock options by corporate chief executives, and in its use by organizations such as WikiLeaks or ProPublica. Recently, journalists have been highlighting the behavior of some of the world's richest people after a massive leak of data that became known as the Panama Papers (https://panamapapers.icij.org/). As well, data are becoming an increasingly important research tool for investigating issues such as public spending, procurement, and public services (Felle, 2016).

Most theorists agree DJ is made up of three processes (and related skill sets)—accessing data that represents (or might represent) information in the public interest, analyzing that data using appropriate software, and developing stories on meaningful discoveries made through that analysis. While the use of statistics and basic data analysis by journalists is not new, the increased availability of large datasets along with more powerful software analysis tools (Ubaldi, 2013) have allowed journalists to report on issues in new and often more powerful ways. Journalists working in this area are reporting for wide audiences, including using multimedia presentation tools to visualize data and tell, often complex, stories in easy-to-understand ways. DJ is thriving at larger news organizations (Fink & Anderson, 2015) and producing results that would be unobtainable without journalists who can understand its processes. Not surprisingly, newsrooms increasingly demand journalists who, as Tim Berners-Lee described them,

> know their CSV from their RDF, can throw together some quick MySQL queries for a PHP or Python output . . . and discover the story lurking in datasets released by governments, local authorities, agencies, or any combination of them—even across national borders. (Arthur, 2010)

Teaching Data-Driven Reporting—The Challenges

Yet, for those tasked with teaching journalists, there is little scholarship on DJ education to turn to (Hewett, 2015), though a number of manuals and textbooks (e.g., Gray, Chamber, & Bounegru, 2012; Herzog, 2015) are emerging to support curriculum design. A study by Splendore et al. (2016) showed that "educational institutions—academic, vocational, and professional—increasingly [are now embracing] data journalism" (p. 149). But journalism schools face a number of challenges. To start with, their programs typically attract students interested in humanities with, often, unproven skills in even basic data analysis. Many students have yet to grapple with even the basics of Microsoft Excel, and as such, any thoughts of introducing Berners-Lees's "MySQL queries for a PHP or Python output" is unrealistic, especially in a crowded curriculum. Similarly, teachers are usually former journalists, not former computer scientists, and often have little more than a reporter's understanding of statistics. Yet they are now required to keep up with an explosion of new data tools that is rapidly expanding the skill base required of educators and their curricula (Krueger, 2014). What's more, the innovation required to teach a rapidly changing journalism field can be extremely difficult at academic institutions, where organizational models and processes, along with pressure on resources, get in the way (Newton, 2016).

This article reports on both a collaborative initiative in 2015 between Auckland University of Technology (AUT) in Auckland and the University of Canterbury (UC) in Christchurch, supported by a visiting Fulbright specialist and aimed at introducing DJ to postgraduate journalism students, and the introduction of a semester-long undergraduate paper in the subject at AUT. The article explores the challenges faced by both cohorts of students and suggests ways for teachers new to DJ to engage with both the topic and their students.

The Postgraduate Project

For the postgraduate project, journalism students at the two universities investigated state-run gambling in Aotearoa–New Zealand. Instructors had months earlier filed Official Information Act 1982 (OIA) requests for data on how much money the New Zealand Lotteries Commission earned and how it distributed its funds, as well as information about its marketing program and expansion. Postgraduate students enrolled in an investigative journalism paper at UC and a core, advanced-reporting paper at AUT spent the second semester of 2015 analyzing the responses to the OIA requests to map the cost of state-run gambling to different communities and track whether the money spent in each community was returned to "make good things happen," as claimed by the commission (New Zealand Lotteries Commission, 2015, p. 4). In teams, students further investigated issues around problem gambling, New Zealand Lotteries' links with gambling businesses offshore, gambling's impact on poorer communities, and New Zealand Lotteries' management.

Collaboration between the two universities came about through informal connections made across journalism programs. A Fulbright specialist in the United States, an

experienced data journalist, had led a student investigation into the Illinois state lottery, as well as working on data projects in Chile, and offered to help instigate a similar project in New Zealand. Among other things, the collaboration aimed to connect New Zealand journalism educators with a growing international network of DJ teachers and to experiment with new ways of teaching DJ within the bounds of established curricula, avoiding the usual lengthy lead-in for curriculum changes.

The experiment was collaborative, in part, because DJ is a collaborative endeavor that works best with cross-pollination between people with different skill sets (Howard, 2014). Collaboration also offered a way of quickly building hybrid skills across the two academic settings and demonstrating the networked model that is crucial to the development of DJ (Howard, 2014). Such teamwork can be found in journalism and data networks elsewhere, including Investigative Reporters and Editors (IRE) and the National Institute for Computer-Assisted Reporting (NICAR), and repositories like GitHub.

The data project also aimed to develop students' broader skill sets through a hands-on approach. New Zealand journalism education has a strong history of experiential learning (Dewey, 1938; Kolb, 2014; Lewin, 1951), and teaching DJ appeared to be a good fit with such "learning by doing" or problem-based learning. The project was designed to evolve as students accessed, cleaned, and analyzed data, and in some respects drove their own learning. Instructors and industry partners would work closely in instructional terms, but teams of students would develop the parts of an overall story package for which they were responsible in ways that they themselves would determine. By adopting an experimental mind-set, everyone was pushed to try something new, no matter their skill level; by engaging in an experiential project, students improved their capacity for the self-guided learning and motivation they need in the newsroom to keep up with new digital tools and a continually changing skill set (Howard, 2014).

The visiting Fulbright specialist was instrumental in establishing the parameters of the project, visiting both campuses to introduce the essentials of data analysis and visualization techniques to students, as well as staff at AUT. This was an arguably successful response to the issue of data competence among teachers. The specialist ran workshops more than a week at AUT (during a study break; nevertheless those students who turned up went on to take lead roles in the project) and then spent a week at the UC, working with students. It is unlikely the New Zealand instructors would have been as ambitious with this project had they not had that mentoring and oversight. His role also included beginning to integrate students and teachers into a global network of journalists, teachers, and others working in the DJ area, connecting people with others at IRE and the NICAR, for instance. He helped build support, too, from non-journalists working with data. For instance, when students asked for help in accessing and cleaning data that a local business would not provide, he connected them with a sympathetic data enthusiast in Australia who scraped and cleaned the data for them. At another point, a Ministry of Health data analyst (tracked down through Open Data networks) helped one of the UC teams locate and interpret alternative datasets on problem gambling. This kind of support was integral to the levels of success the

project enjoyed and reinforced for instructors and students alike the availability of support from those within the data field.

Getting the Data

The difficulty we faced sourcing data for this project clearly suggested a need for greater transparency in the running of New Zealand's state-gambling agency and the distribution of gambling funds. We sought data on both the operation and the income of the agency, and the distribution of its profits to the community. Despite initiating a request for data more than 4 months before the beginning of the postgraduate project (the authors were in contact via Skype meetings and email to map out and plan the project for at least 6 months before the teaching semester started, and the AUT and UC instructors put in many hours to craft and fight for the OIA data requests), some data—arguably the most sensitive and potentially newsworthy—were never released. A lengthy dispute over the students' rights to the data under the OIA has ensued and is, at the time of writing, subject to an investigation by the Office of the Ombudsman. The risk of such a situation arising must be considered when planning DJ projects as part of journalism courses. On reflection, we may not have done anything differently, but the question that needs asking is "What is the right data to seek, given the need to balance the likelihood of access and the potential impact of its analysis, and subsequent news stories." The possibility of a high-impact story seemed key to keeping students strongly motivated. If innocuous data with little potential for strong news stories are sourced, student application—and consequently student learning in an experiential environment—is likely to be reduced. The success of former projects the Fulbright specialist had helped run gave the New Zealand project organizers confidence in the impact of the gambling story—a quick check of freely available data (e.g., annual reports) suggested similar trends were occurring in the New Zealand state-gambling sector. However, the finely granulated data that would show exactly how much the poorest citizens were contributing to the funding of New Zealand's arts, elite sport, and filmmaking industries was the data the agencies stubbornly refused to release, citing, at different times, reasons of commercial sensitivity, unavailability of the data, and the broadness of the request. As well, the agency refused to release 5 years of the minutes of its board for the same reasons (though, at the 11th hour, it did release heavily censored copies of one year's minutes). These refusals, though frustrating, provided powerful teaching and learning moments. Students were able to follow, and ultimately report on, the struggle for data, problem-solved about where to get alternate data, and along the way learned more about their rights as seekers of information and the appeal processes available when data are withheld. The agencies' lack of cooperation strengthened the students' resolve to seek sources elsewhere and they wrote convincing stories on the growth of gambling and its impact in poorer communities.

As a side note, the fight for data on this project also allowed the two journalism schools to take a proactive role in advocating for greater data transparency and accountability from key public agencies. In New Zealand, the Lottery Grants Board is charged with distributing gambling revenue back to the community through its national

and regional funding committees and three separate statutory bodies: Creative NZ, the NZ Film Commission, and Sport New Zealand. Students quickly discovered that when analyzing the datasets that were released, that data were collected inconsistently across the different agencies, which barred meaningful comparisons between funders and finer-grained analysis of aggregated lotteries and grants data. Although we were forced to abandon our attempts to measure and compare some aspects of gambling finances, we were able to use this as an opportunity to advocate for transparent and accessible government data. In New Zealand, as elsewhere, there are growing concerns about access to public information (Drinnan, 2013) and it was our view that journalism schools should take a lead in this regard. As such, we pushed for a commitment from one agency to re-examine its failure to record funding by region, and we continue to lobby other agencies through appeals to the Ombudsman's office for improved access to data, which we hope to analyze and publish.

Data Competence

The level of "data competence" among students was an obstacle to overcome. A handful had some data analysis, Excel and coding skills, but none were advanced in this area and most students had, at best, minimal knowledge of Microsoft Excel. On the Canterbury program, students were given preliminary classes in Excel, data analysis, and data visualization to give the class a common grounding in data principles and techniques but arguably needed more to be proficient and confident enough to work alone. We would recommend more collaboration with data experts in the preparation of the DJ curriculum to scaffold students into future projects. In the postgraduate project, coding was considered too far a step. It is true that theorists are increasingly calling for some level of coding among journalism graduates. Anderson, Bell, and Shirky (2015) wrote,

> Journalists should learn to code. It's true that to be fluent and useful in many programming languages requires very highly developed skills; not every journalist will be able to do this, and not every journalist should do this. But every journalist needs to understand at a basic literacy level what code is, what it can do, and how to communicate with those who are more proficient. (p. 31)

It was noted by industry experts whose advice was sought for this project that the real power of DJ is best released through bespoke programming and visualizations. However, to make this project manageable within its time constraints and student skill sets, both universities kept mostly to analysis of data through spreadsheet-based software and visualization of trends through relatively user-friendly online services, including Google Fusion Tables and Datawrapper. The Fulbright specialist reinforced the usefulness of basic analysis techniques (e.g., simple pivot tables allowed students to slice and dice datasets until they saw story angles) which helped students work out the best ways for collating, cleaning, analyzing, and reporting on the data, thereby introducing them to the fundamentals of DJ. Varying levels of data

competence within both the instructor group and the student cohort was not—at least in this case study—a fatal issue. Similarly, faced with the realities of building data-sets, students learned to experiment, to assess their own strengths and weaknesses, and how to find and ask for help.

A key lesson learned during the postgraduate project was the need to monitor the quality of the data throughout. Keen as they are to find their story, students need to be taught to manage the data trail patiently in an organized manner, and to invest adequate attention to cleaning and fact-checking the data. Although it should be student-driven, help with project management is vital as problems missed in early stages potentially risk the validity of stories further down the track. The project highlighted a need to keep students focused on the data to keep questioning and affirming what their core story was—what is your data, what are the apparent trends, what more do you need, and what questions does it answer/pose—and to have data tests at multiple staging points to audit the accuracy of students' data analysis.

Changes for Next Time

The experience in the postgraduate pilot project has underscored the need for data requests to be lodged perhaps even a year ahead if the data might be in any way sensitive (which they perhaps should be). A curriculum change would see an introduction to coding included, if only to convince students of its importance to DJ. As well, the instructors agree they could make better use of communication interfaces such as Google hangouts to join collaborating classes more tightly in their work. In terms of publication outcomes, it would seem a good idea to team up with an industry-based news organization as commissioning publisher. The investigation into state-run gambling did have initial support from an editorial manager at a weekend newspaper who was keen to run the story package. However, his involvement with the project ended thanks to industrial restructuring, and publication would now seem likely in school media, from where it may well be picked up by that same newspaper. It would be fruitful to have a partner publication not only keen to publish student work but also to devote some of its data specialists' and editors' time to mentoring students through the project.

The Undergraduate Paper

Meanwhile, AUT university's new DJ course, Jour703, was introduced in the second semester of 2015, running as a final-year elective within the university's Bachelor of Communication Studies program. The cohort was 14 students, of whom 11 were journalism majors, two were public-relations majors, and one majored in television. To gain entry to the course, students had to sit a basic maths and statistics test. At least five students who had provisionally enrolled in Jour703 pulled out when they learned they had to sit a maths test.

The teacher for Jour703 had previously taught a "maths and statistics for reporters" unit within the journalism program but had no prior experience of DJ. Therefore, while planning the course, he collaborated with several professionals with relevant expertise

in DJ and education: one was the Fulbright specialist, who ran workshops for AUT staff and who shared his own DJ curriculum and provided feedback on course drafts; a second was the data editor for Auckland's daily newspaper, the *New Zealand Herald*, who provided ideas on course content; a third was an AUT statistics lecturer. Another influence on the content of Jour703 was the scope of the textbook selected for the course: *Data Literacy: A User's Guide*, written by David Herzog (2015), an associate professor at the Missouri School of Journalism in Columbia.

The topics covered in Jour703 included evaluating examples of best practice in DJ across worldwide media; statistics principles and practice linked to DJ; sources of data, open government, social media, and finding news angles in data; cleaning, analyzing, summarizing, and filtering data; and visualization. The main tools students used included Microsoft Excel, Google Sheets, Google Fusion Tables, Datawrapper, Open Refine, Piktocharts, and Zamzar. The weekly 3-hr Jour703 class was usually split into three parts—a lecture or guest speaker, student presentations on published DJ projects, then a hands-on workshop working with data.

There were three assessments in Jour703. The first was a presentation to the class (worth 20%) where students worked in pairs to lead a class discussion on a significant DJ project. The second was a proposal for a data story (30%) that outlined the scope of the story, identified the databases that would be used and four sources to interview, and included a plan to analyze and visualize the data. The third was a 1,500-word data story (50%) complete with infographics and tables, a case study profile, a data analysis report, plus supplementary spreadsheets and interview transcripts. Students' data stories tended to focus on trends in social issues with data in many cases drawn from the latest (2013) national census. Story topics included the following: women having children later in life, trends in teenagers taking up driving, changing baby names reflecting greater population diversity, and changes in regional vaccination rates. In addition to their assignments, students were also expected to contribute to a class blog that shared opinions on a series of academic journal readings on DJ.

Student feedback on Jour703 was very positive. The first graph (Figure 1) below shows data from an online survey of student satisfaction conducted in the final week of the course. The second graph (Figure 2) shows satisfaction levels with Jour703 compared with other papers taught across the Communication Studies School, the wider faculty, and the university.

Despite the positive feedback, Jour703 presented several challenges for the teacher. The first challenge was determining the scope and boundaries of the course. Among the questions arising were what skills to cover, how much maths and statistics to include, what software tools to focus on, and whether or not to include coding/programming languages like Python, Ruby, PHP, or Perl? Programming is used in many newsrooms, and the data editor at the *Herald* argued for its inclusion. Some journalism schools in the United States include programming in their DJ curricula, while others limit themselves to spreadsheet programs such as Excel. The teacher, who had no personal experience with programming, had to weigh up what was achievable in an introductory course. In *Data Literacy: A User's Guide*, Herzog (2015) covers identifying, obtaining,   evaluating,   cleaning,   analyzing,   and   visualizing   data   but
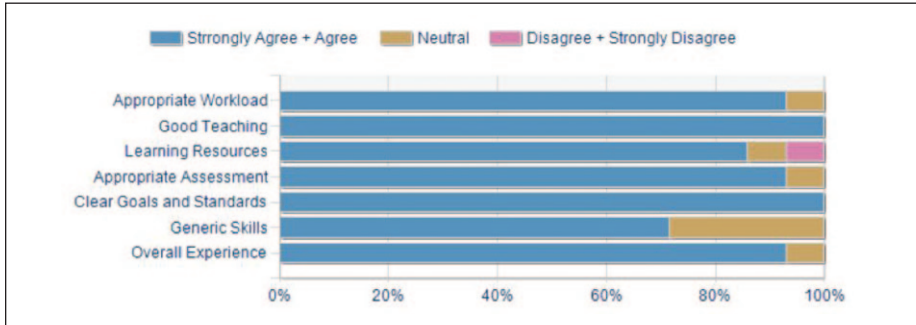
Figure 1. Paper rating—overview.
Source. AUT University Student Paper Experience Questionnaire report on Jour703 DJ, Semester 2, 2015.
Note. AUT = Auckland University of Technology; DJ = data journalism.
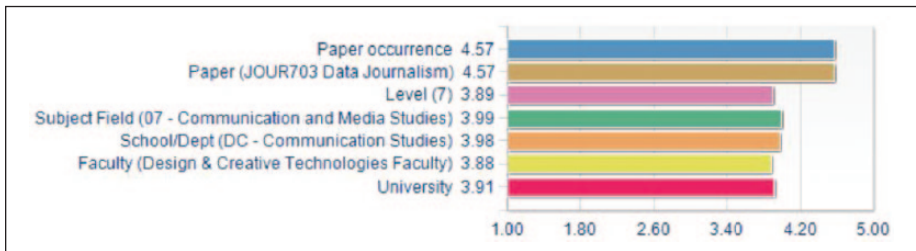
Figure 2. Mean overall experience for Jour703 versus university averages.
Source. AUT University Student Paper Experience Questionnaire report on Jour703 DJ, Semester 2, 2015.
Note. AUT = Auckland University of Technology; DJ = data journalism.

mentions programming languages only in passing, seeing them as beyond the scope of an introductory text. Jour703 had no practical programming component, but the *Herald* data editor did a 2-hr class in which he showed students examples of coding and explained how it worked in principle.

A second challenge was assessment: Should assessments narrowly test students' technical skills in data analysis, visualization, and so on; or should students also be expected to interview sources, take photographs, and present their data work as part of broader pieces of journalism with a text-based narrative holding the story together. The teacher felt it was important that the data work was not seen as an end in itself, but that it served journalism. A third challenge was estimating students' prior knowledge of basic maths and statistics concepts, and whether they would bring any prior experience of working with spreadsheets and visualization tools. It turned out most students were not particularly strong in maths and statistics, and most had no prior experience with the software used.

For the second iteration of Jour703, the teacher plans to add in a couple of practical sessions on programming. In their detailed feedback, several students said they would

have appreciated some hands-on experience with coding. Beyond that the teacher plans to add some new visualization apps to the toolbox and make some minor changes to the assessments.

Conclusion

DJ is increasingly being introduced to the journalism curriculum as editors begin to seek graduates who can bridge the worlds of reporting and computer science. Instructors introducing projects and courses in DJ face, among other issues, relatively low levels of data competence among students, and perhaps faculty too. However, international support is available for those wanting to establish a foothold in DJ education. The postgraduate project demonstrated the value of collaborative teaching to help instructors tackle new curriculum content, including both data analysis skills and innovative approaches to their delivery. Tackling new skills through experiential learning helped students adopt an experimental mind-set—they learned a lot, despite the fact final stories did not come together as planned—and answered calls for a teaching hospital model (e.g., Newton et al., 2012). As it evolved, the project ran less like a classroom and more like a newsroom—students defined and created the journalism, professionals mentored them to improve the quality and impact of that journalism, specialists brought in expert knowledge, and everyone worked together on a story with real-world impact.

Although the postgraduate project faced the usual issues such project-based approaches face, including extra demands on instructors' time and restrictive time-tabling that made it difficult to build the investigation around students' other commitments (work placements and other significant projects), it demonstrated the value of attempting curricula change iteratively—and collaboratively. The partner programs now have real-world practice datasets with which to train subsequent journalists. They also have data on which to keep building the lotteries investigation (i.e., the second attempt will be stronger for the groundwork done), and a real-world and close-to-home investigation to share with future students. By attempting a large project through international collaboration, instructors were able to not only tackle new curricula and innovation but also gain valuable teaching strengths from one another. The value of regular debriefs between project partners in this regard cannot be stressed enough.

The undergraduate paper, meanwhile, demonstrated that even with a cohort with limited knowledge of statistics, a successful Level 7 introduction to DJ can be achieved, complete with high student satisfaction and real-world outputs as a result of problem-based teaching. The point, it would seem, is that the obstacles to teaching DJ should not prevent us from attempting to embed it in our programs. These are early days in DJ education in New Zealand, but these two projects have set early templates for development.

Authors' Note

Declaration of Conflicting Interests

Funding

References

Anderson, C. (2013). Towards a sociology of computational and algorithmic journalism. *New Media & Society*, *5*, 1005-1021.

Anderson, C., Bell, E., & Shirky, C. (2015). Post-industrial journalism: Adapting to the present. *Geopolitics, History, and International Relations*, *7*, 32-123.

Arthur, C. (2010, November 22). Analysing data is the future for journalists, says Tim Berners-Lee. *The Guardian*. Retrieved from http://www.guardian.co.uk/media/2010/nov/22/data-analysis-tim-berners-lee

Culver, K. (2014, July 7). Special series: Teaching data journalism. *MediaShift*. Retrieved from http://www.pbs.org/mediashift/2014/07/special-series-teaching-data-journalism

Dewey, J. (1938). *Experience and education*. New York, NY: Simon & Schuster.

Drinnan, J. (2013, August 23). Media freedom under fire. *New Zealand Herald*. Retrieved from http://www.nzherald.co.nz/business/news/article.cfm?c_id=3&objectid=11112909

Felle, T. (2016). Digital watchdogs? Data reporting and the news media's traditional "fourth estate" function. *Journalism*, *17*, 85-96.

Fink, K., & Anderson, C. (2015). Data journalism in the United States. *Journalism Studies*, *16*, 467-481.

Flew, T., Spurgeon, C., Daniel, A., & Swift, A. (2012). The promise of computational journalism. *Journalism Practice*, *5*, 157-171.

Gray, J., Chambers, L., & Bounegru, L. (2012). *The data journalism handbook*. Sebastopol, CA: O'Reilly Media.

Herzog, D. (2015). *Data literacy: A user's guide*. Thousand Oaks, CA: SAGE.

Hewett, J. (2015). Learning to teach data journalism: Innovation, influence and constraints. *Journalism*, *17*, 119-137. doi:10.1177/1464884915612681

Howard, A. (2014). *The art and science of data-driven journalism*. Retrieved from http://towcenter.org/research/the-art-and-science-of-data-driven-journalism/#v.-pathways-to-the-profession

Kolb, D. (2014). *Experiential learning: Experience as the source of learning and development*. Upper Saddle River, NJ: Pearson Education.

Krueger, K. (2014, July 7). How J-schools are tackling the demand for data journalism skills. *MediaShift*. Retrieved from http://mediashift.org/2014/07/how-j-schools-are-tackling-the-demand-for-data-journalism-skills/

Lewin, K. (1951). *Field theory in social sciences*. New York, NY: Harper & Row.

Meyer, P. (1991). *The new precision journalism*. Bloomington: Indiana University Press.

Meyer, P. (2002). *Precision journalism: A reporter's introduction to social science methods*. Lanham, MD: Rowman & Littlefield.

Newton, E. (2016, January 28). Seeing innovation in a new, better way for journalists, educators. *MediaShift*. Retrieved from http://mediashift.org/2016/01/seeing-innovation-in-a-new-better-way-for-journalists-educators/

Newton, E., Bell, C., Ross, R., Philipps, M., Shoemaker, L., & Haas, D. (2012). *An open letter to America's university presidents*. Retrieved from http://knightfoundation.org/press-room/other/open-letter-americas-university-presidents/

New Zealand Lotteries Commission. (2015). *Annual report 2014/2015*. Retrieved from https://mylotto.co.nz/assets/footer-pages/about-us/reports-and-statement-of-intent/Annual-Reports/Lotto-NZ-Annual-Report-FY15.pdf

Splendore, S., Di Salvo, P., Eberwein, T., Groenhart, H., Kus, M., & Porlezza, C. (2016). Educational strategies in data journalism: A comparative study of six European countries. *Journalism*, *17*, 138-152.

Ubaldi, B. (2013). *Open government data: Towards empirical analysis of open government data initiatives* (OECD working papers on public governance, No. 22). doi:10.1787/5k46bj4f03s7-en

Author Biographies

Greg Treadwell, a former newspaper reporter, photographer and editor, is a senior lecturer in journalism at Auckland University of Technology in Aotearoa-New Zealand. He teaches investigative journalism, news reporting and photojournalism and has research interests in freedom of information and journalism education.

Tara Ross is head of the journalism programme at the University of Canterbury. She is a former award-winning senior journalist for Fairfax Media newspapers *The Press* and the *Sunday Star-Times* and is completing a PhD on Pasifika news media in Aotearoa-New Zealand.

Allan Lee is a senior lecturer in journalism at Auckland University of Technology. He teaches news and feature writing, data journalism and news editing and design. His research includes production journalism and reporting on the economy. Allan has worked as a business journalist for newspapers and magazines in the UK, Australia and Aotearoa-New Zealand.

Jeff Kelly Lowenstein is an investigative journalist, writer and assistant professor at Grand Valley State University. His work has been published by *The New Yorker* and the Center for Public Integrity, and has earned local, national and international recognition. A Fulbright teacher, scholar and specialist, he has written three books.