

Article

Theory
&
Culture
&
Society

Statistical Practice: Putting Society on Display

Theory, Culture & Society

2016, Vol. 33(3) 51–77

© The Author(s) 2015

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0263276414559058

tcs.sagepub.com



Michael Mair

University of Liverpool

Christian Greiffenhagen

Loughborough University

W.W. Sharrock

University of Manchester

Abstract

As a contribution to current debates on the ‘social life of methods’, in this article we present an ethnomethodological study of the role of understanding within statistical practice. After reviewing the empirical turn in the methods literature and the challenges to the qualitative-quantitative divide it has given rise to, we argue such case studies are relevant because they enable us to see different ways in which ‘methods’, here quantitative methods, come to have a social life – by embodying and exhibiting understanding they ‘make the social structures of everyday activities observable’ (Garfinkel, 1967: 75), thereby putting society on display. Exhibited understandings rest on distinctive lines of practical social and cultural inquiry – ethnographic ‘forays’ into the worlds of the producers and users of statistics – which are central to good statistical work but are not themselves quantitative. In highlighting these non-statistical forms of social and cultural inquiry at work in statistical practice, our case study is an addition to understandings of statistics and usefully points to ways in which studies of the social life of methods might be further developed from here.

Keywords

ethnomethodology, knowledge, quantitative-qualitative divide, social life of methods, statistics, understanding

Introduction

Stemming from a growing interest in the ‘social life of method’ (Law et al., 2011; Savage, 2013; and for a more general overview see Mair

Corresponding author: Michael Mair. Email: michael.mair@liverpool.ac.uk

Extra material: <http://theoryculturesociety.org/>

et al., 2013), studies of the emergence, contestation, stabilization, proliferation and collapse of new methods and their diverse social, cultural and political implications have been used to challenge a series of methodological orthodoxies allegedly definitive of the social sciences since the mid-20th century (Savage and Burrows, 2007). The 'qualitative-quantitative divide' has become a principal target (see Majima and Moore, 2009; Adkins and Lury, 2012). Though it served as a significant line of demarcation in the geography of the social sciences, the qualitative-quantitative divide is now seen less as a technical distinction and more as a moral and political one, exerting a pervasive and enduring influence on the 'imaginaries' of the social sciences – something contemporary empirical work on method seeks to 'unsettle' and 'disturb' (Majima and Moore, 2009).

Those seeking to undermine a sharp distinction between the qualitative and the quantitative are not alone in voicing disquiet, joining other researchers arguing for a complete rethink of its putative bases. Current discussions about 'descriptive', 'inventive' or 'live' methods in the social sciences (e.g. Savage, 2009; Lury and Wakeford, 2012; Back and Puwar, 2012) were, in many respects, prefigured by advocates of 'mixed methods' from the 1980s onwards (e.g. Bryman, 1988; Tashakkori and Teddlie, 2010). Nonetheless, no matter how diversified the reasons for the critique, there are commonalities in the positions advanced. Those arguing for a rethink hold: (a) that allowing researchers to identify themselves in binary terms as either qualitative or quantitative, and specialize in one 'kind' of technique, is counter-productive, encouraging the subordination of phenomena of interest to the methods used to study them, thereby making the social sciences more rigid and dogmatic, less flexible and responsive (see, e.g., Majima and Moore, 2009; Tashakkori and Teddlie, 2010; Gane, 2012); (b) that describing researchers as qualitative or quantitative does not describe their actual practices – research does not neatly break into 'qualitative' and 'quantitative' work (see, e.g., Majima and Moore, 2009; Latour, 2010; Tashakkori and Teddlie, 2010; Gane, 2012); and (c) that describing researchers as qualitative or quantitative does not describe the materials they work with – the qualitative-quantitative distinction does not carve 'the empirical' at the joints, it does not neatly separate quantities and qualities, meaning that the quantitative-qualitative divide is an artificial and unhelpful dichotomy imposed upon the phenomena researchers are seeking to make sense of (see, e.g., Kuhn, 1977; Latour, 2010; Gane, 2012). Gane usefully summarizes the discontent: 'a[ny] discipline dominated by stock quantitative and qualitative methods is a discipline not only lacking in imagination, but also one that in spite of its claims can never be empirical in any meaningful sense' (2012: 159).

Concerns about the bipolarization of research have acquired added urgency in an age of digital data, digital 'devices' and digital methods in

their many and varied forms (Elliot et al., 2013; Rogers, 2013; Ruppert et al., 2013). In a research context characterized by rapid methodological and technological change, a context where the computationally-oriented are as likely to contribute to discussions on aesthetics (e.g. Tufte, 2006) as literary and cultural theorists are to contribute to discussions on natural language processing and data mining methodologies (Moretti, 2013; Mohr and Ghaziani, 2014), the claim that this timeworn binary scheme usefully partitions the universe of research is harder to sustain. Social, cultural or political research is no longer meaningfully mapped along clear-cut divisions, even rhetorically, with the era of big data analytics opening the way to hybridized contributions from, and collaborations with, art(i)s(ts) and humanities scholars, journalists, programmers, software developers, mathematicians, computer scientists, data analysts, and more (Savage, 2009; Savage and Burrows, 2007; Ruppert, 2013; Ruppert and Savage, 2011). These newcomers have few qualms about employing what would, until recently, have been regarded as unstable and methodologically suspect mixtures of data and techniques in order to grapple with the analytical problems at hand. The end-results of these activities are as varied as the paths by which they are initiated: software applications, computer code, probabilistic (i.e. 'statistical') and non-probabilistic (i.e. 'algorithmic', e.g., natural language processing and machine-learning) models, digital visualizations, media, crowdsourcing and citizen science platforms, searchable archives of audio, video and textual materials, project wikis, digital user-interfaces, databases, news stories, forecasts, even performances are as visible in contemporary research as the book or journal article. However, even in digital research there are fears that the qualitative-quantitative divide is being re-established just as it was losing whatever residual significance it retained. As boyd and Crawford recently put it, within big data analytics and digital research more broadly,

there remains a mistaken belief that qualitative researchers are in the business of interpreting stories and quantitative researchers are in the business of producing facts. In this way, Big Data risks re-inscribing established divisions in the long running debates about scientific method and the legitimacy of social science and humanistic inquiry. (2012: 667)

However, while many topics of interest are raised by current empirical studies of method both digital and traditional, much of the recent debate only makes sense against a background of understandings about what research involves. It is assumed that 'everyone knows' what qualitative or quantitative research consists in, an assumption that is traded upon but not itself subject to investigation. Gane's summary, much like that of boyd and Crawford's, is intelligible because it is assumed that those

reading it will *already* know what is being talked about. If, however, as critics argue, an essentialist division of research into quantitative and qualitative strands makes little sense, a position we have sympathy with, the question arises as to why those terms continue to be employed and what alternatives might be found for characterizing what researchers do.

After Garfinkel, we want to examine taken-for-granted assumptions about quantitative research in particular, treating ‘as problematic . . . the *actual* methods whereby members of a society, doing sociology, *lay or professional*, make the social structures of everyday activities observable’ (Garfinkel, 1967: 75; emphasis added). The rest of our paper, therefore, presents a case study of quantitative social scientists – statisticians – at work. We argue, based on this case study, that labelling research practices as qualitative or quantitative (or indeed ‘mixed’) may well have some uses (as badges of membership, for instance), but the labels themselves are not specifically descriptive of those practices and should not be treated as such. Knowing whether a piece of research is qualitative or quantitative, interpretive or calculative, we shall argue, is much less important for characterizing that research than understanding the specific ways in which it makes ‘the social structures of everyday activities observable’ – that is, how it puts society on display.

Using examples drawn from our fieldwork, we discuss two ways in which the forms of understanding exhibited in and through statistical practice put ‘the social structures of everyday activities’ on display. Firstly, we argue that in their work, statisticians treat databases as socio-cultural not just numerical artifacts. Indeed, statisticians both investigate and reproduce the interweaving of cultural and technical reasoning in the course of working on the numbers databases are composed of. Drawing on Sacks’ notion of culture as an apparatus for both doing and seeing the doing of things (Sacks, 1992: 226), we argue statisticians must work ‘ethnographically’ as well as mathematically not only to understand their data but also to be understood when they employ that data in their analyses in turn. Secondly, we argue that statisticians’ work exhibits an understanding of those they hope will use the statistical tools they produce. Drawing on an example of a model which failed on this count, we argue such displays of understanding are most powerful where they draw least attention because they are built into the work itself. Insofar as statistical tools are relevant to others, an *acquired* understanding of those others is thus on display in the design of those tools. Where the first way of putting society on display involves ethnographic ‘forays’, as one anonymous reviewer felicitously put it, into the societies and cultures in which the data being worked on has been produced, the second involves ethnographic forays into the societies and cultures to which the products of their work are offered.

Law, Ruppert and Savage (2011) argue that methods are ‘doubly’ social: ‘methods are social because they are constituted by the social world of which they are a part’ but are also ‘social because they also help to constitute that social world’ (2011: 1–2). Our case study provides a different, less metaphysical, more concrete sense in which statistical methods have a ‘double social life’, one which consists precisely in particular ways of exhibiting understandings of social and cultural practice in terms of both their data *and* the statistical tools they develop, ways of making social structures observable central to the intelligibility of their work. Such exhibited understandings are grounded in determinate lines of practical social and cultural inquiry integral to and undertaken as an accountable, i.e. observable-reportable, part of good statistical work. By attending to these forms of inquiry we gain insight into the ways in which statistics connects with, and can be seen as relevantly speaking of and to, the societies and cultures it analyses.

Quantification and Statistical Practice

Studies of statistical practice exemplify many of the issues we find in the wider literature on methods. Probability, chance, statistics, quantification and the mathematization of natural and social phenomena have been recurrent topics within philosophy and the social sciences (Lynch, 1993) and a large number of studies have been conducted on these subjects.¹ However, as in other areas of the social sciences, despite the scale and scope of the literature on these topics, aside from historical reconstructions (e.g. Mackenzie, 1981; Stigler, 2002), exploratory studies of one kind or another (e.g. Livingston, 1987; Gephart, 1988; Bowker and Star, 1999; Desrosières, 2009; Gephart and Smith, 2009; Ruppert, 2007; McGinn, 2010; Kullenberg, 2011) and schematic ‘as if’² treatments (e.g. Latour, 1987; Law, 2009; Sandelowski et al., 2009; Verran, 2012), statistical work, what it consists in and how it is carried off, remains opaque to the general sociological public. By reading these and other studies we learn a good deal about what kinds of operations statistical work involves (e.g. classifying, scaling, calculating, modelling), what those operations effect (e.g. transformation, formalization, abstraction, simplification, reduction), how numbers acquire their credibility, their persuasive power and their utility (in the course of being harnessed for various social, political and economic ends) but acquire little sense of how any of these might be effected in practice (see Gephart and Smith, 2009). Bowker and Star’s work (1999), for instance, is instructive on the consequences of classification but much less so on the mathematical/statistical work it rests upon. Even well-regarded textbooks such as Hacking’s (2001) are largely silent on issues of statistical practice, beginning with ready-to-work-on, pre-formatted problems rather than inspection of the *in situ* work of reasoning those problems, and the worldly troubles which

accompany them, through. Asked to describe what statisticians do, we would struggle to respond armed only with the literature cited above – ‘normal statistics’, to adapt Kuhn, is a foreign land.

While it is beyond the scope of the present article to examine these issues in depth, we hope to partially redress the balance. Rather than come at them in the broad, we will concentrate, in the first instance, on how an operation like mathematization or simplification is accomplished and the practical work, interpretive and otherwise, that underpins it. Drawing on ethnomethodological studies of number-involving practices in a range of ordinary and specialized settings,³ the operations performed may be quantitative in character but the work involved itself resists treatment in those terms. Although we will be dealing with a form of expert practice, Bayesian statistics, our interest is in the ‘ethnomethodological foundations’ or ‘praxiological grounds’ of some of the more ordinary, and therefore more accessible, features of that work (Sharrock, 2000). Since we have dealt with more ‘technical’ aspects of these practices elsewhere (Greiffenhagen et al., forthcoming), we have chosen to focus here on matters that require little digging to gain an appreciation of.

We take up two differently leveraged but elementary examples of ‘understanding work’ drawn from a small-scale study – the process of recasting data into more computationally tractable forms and of calibrating statistical applications for use by others – precisely because they point to the more diffuse sets of background competencies upon which these specialized practices rest. These are aspects of statistical practices which are passed over in most accounts but, by focusing on them, we gain a fuller sense of how statisticians orient to and methodically go about their work *as work*.⁴ In the rush to get onto more ‘interesting’ features of statistical practice it is easy to lose sight of precisely these aspects of it, and so the ‘tangible fabric’ into which all the different activities which make up statistical practice are woven (Toulmin, 1953: 10–11; Maynard and Kardash, 2006: 1485). The examples are thus chosen to correct simplified conceptions of statistical practice in which the practices have been largely left out. To adapt Machiavelli, ‘those who [would] make maps of countries [do well to] place themselves low down in the plains to study the character of mountains and elevated lands’ (1910 [1532]).

The Praxiological Grounds of Bayesian Statistical Work

In 2010 we conducted fieldwork with a group of Bayesian statisticians in a UK university who operated across various social science domains as well as epidemiology and biostatistics. The focus of their work was very large datasets analysed using complex modelling and computational

techniques. Funded via a UK-wide, cross-disciplinary research initiative, they had an explicit remit to experiment with method.

We had a limited amount of time with the group. Although we followed the progress they made from afar for around a year all told before and afterwards (based on publications, reports, conference presentations and so on), in terms of field visits we could only see them on a handful of occasions, approximately 12 full days in total in the course of a few calendar months. As a consequence, we are in no position to discuss many aspects of their research. However, we can say something about the kind of work they do. Rather than grapple with the complexities of contemporary Bayesian modelling here, we concentrate, as discussed above, on more mundane, less technical aspects of the projects the group was involved in. While not what they regularly publish on, we believe an explication of these aspects of statistical practice is revealing, not least because it exemplifies some of the ways in which statisticians make ‘social structures’ available in and through their local activities (Garfinkel, 1967). Not only that, it enables us to be much more specific about what those local activities are.

Although we say little about the specifically Bayesian aspects of their work,⁵ we want to draw attention to some of what makes contemporary Bayesian statistics distinctive. In the wide and varied literatures on probability, literatures stretching back centuries (see, e.g., Fienberg, 2006), Bayesian statistics is normally contrasted with what came to be known in the 20th century as ‘frequentist’ statistical approaches. The frequentist approach defines probability in terms of the relative frequency of particular outcomes of a series of well-defined trials over the (indefinitely) long run. For Hacking, one of the most influential philosophical advocates of this position, this locates probability firmly *in* the world:

Statistical inference is chiefly concerned with a physical property, which may be indicated by the name long run frequency... The long run frequency of an outcome on trials of some kind is a property of the chance set-up on which the trials are conducted. (1965: iii, see also 1–26)

On Hacking’s frequentist interpretation, in studying probabilities – relative to such things as fair dies, fair coins, or fair roulette wheels, for example – the statistician is investigating the *objective* indeed *physical* properties of chance set-ups. The Bayesian approach, by contrast, is typically set out in terms of the degrees of confidence (often cast in terms of betting odds) we can have in the probability or chance of a specific outcome of any particular trial. Whereas the frequentist approach locates probability in the world, as connected to the *properties of objects*, the Bayesian approach, first formulated in the 18th century, thus locates probability, as connected to the *beliefs of subjects*, in the

processes by which we draw inferences about it (Efron, 2013). It is for this reason that Bayesians are often presented as dynamic and information-hungry figures, continuously revising their understandings in line with the latest evidence available to them.

While the frequentist approach dominated 20th-century statistics, the Bayesian approach has enjoyed incredible growth over the last 30 years (Fienberg, 2006; Efron, 2013). A field that was a peripheral and derided fringe interest is, more and more, being treated as a legitimate area of statistical research and practice, and can be said to have greater appeal in areas such as epidemiology, biostatistics and high energy physics than its frequentist 'rivals'. In order to understand this recent shift in the fortunes of Bayesian statistics, it is helpful to draw a distinction between two of its many strands: a distinction between philosophical and methodological Bayesianism. Philosophical Bayesians are committed to a particular philosophy of statistics with a take on probability and inference quite distinct from that of frequentists. Methodological Bayesians, on the other hand, are committed to using the methodological approaches – forms of model building and so on – that have been developed within the Bayesian framework to address particular kinds of statistical problems. A statistician can be a methodological Bayesian without being a philosophical Bayesian, and, indeed, those we worked with in our study were of that cast of mind. They felt the 'tools' of Bayesian statistics were useful for answering the kinds of statistical problem they were most interested in and had a pragmatic, instrumental orientation to their work, showing little interest in discussions about the 'ultimate' nature of probability and inference – discussions we tried and failed to engage them in. Bayesian approaches are spreading, not because philosophical Bayesianism has triumphed over philosophical frequentism (as in all areas of metatheoretical and indeed metaphysical debate, a source of endless, unresolvable dispute), but for reasons more associated with methodological Bayesianism: computing power has improved so much that it is now possible to undertake the calculations needed to do Bayesian modelling.

The key breakthrough has not been a metaphysical but a practical one – the rise of, in particular, Monte Carlo Markov Chain (MCMC) methods, the availability of open-source Bayesian statistical analysis packages, such as the 'BUGS' (Bayesian inference Using Gibbs Sampling) project, and powerful enough computers to process MCMC computations in those packages to determine the distributions from which their models proceed (Diaconis, 2009; Efron, 2013). Bayesian statistics is much more prominent today, in other words, because it has become a viable way of doing statistics, whereas in the past it was not.⁶ Furthermore, Bayesian statistics is not as radical a departure from frequentism as it can sometimes appear to be. Bayes' theorem, the engine of Bayesian inference, can be unproblematically derived from the probability axioms accepted by statisticians of all hues and in many other respects too

converges with frequentist statistical approaches (Gelman and Shalizi, 2013: 26). As a result,

we may see contemporary statistics as a place of happy eclecticism: the wealth of computational ability allows for the application of countless methods with little hand-wringing about foundations. Contemporary practitioners may work blissfully free of the old frequentist-Bayesian controversies; younger statisticians... seem only distantly aware of them. (Mayo, 2011: 79)

Debates about mathematical foundations, any more than debates about philosophical foundations, are not where the relevant differences could or do lie. The most important work-relevant distinctions between a Bayesian and a non-Bayesian are instead to do with such things as the modelling techniques they are most familiar with, the problem-solving strategies they have been trained in, the computer programmes they turn to in order to do the work of modelling, calculation and computation, and the literatures and practical examples they regard as useful guides in their own work. Within this, what is unusual about many areas of contemporary Bayesian statistics, as opposed to frequentist ones, is that they employ different kinds of data and modelling techniques for handling them. This heterogeneity of material, methodology and modelling techniques is the result of an emphasis on making use of as many different types of evidence as possible when making probabilistic calculations, no matter what form the data initially may take (i.e. whether it comes pre-quantified or not; Lynch et al., 2008: 187–9).

For instance, Bayesians, unlike other statisticians, engage in limited forms of data collection, including, among other things,⁷ conducting interviews with experts in particular fields in order to work out initial probability distributions (priors) to ‘plug’ into their models. As with other uses of data, this process is explicitly formalized, conforming to set rules which govern how to conduct these ‘elicitations’ and appropriately quantify the results. Nonetheless, although Bayesians do formalize many aspects of the research process, much of what they routinely do is not explicitly set out in formal terms.

One member of the group, when asked about the nature of the work, explained, in an off-hand remark,⁸ that it consisted of 5 per cent elation, 5 per cent despair and 90 per cent tedium. Although those we worked with occasionally drew attention to that 90 per cent, and all were aware that it was central to the job, it was rarely given extended treatment in their publications despite making up the bulk of what they did, something replicated across the field – making it the ‘dark matter’ of statistical practice (see Chatfield, 2002). We found, as a result, that published work was a particularly poor guide for those not already initiated in what doing statistics involves, furnishing us with little by way of practical

instruction on how to make sense of the work we witnessed these researchers engaging in. Statisticians do not talk about routine practice in an extended way in their publications, and for good reason: it is not what interests their colleagues and collaborators and so is not *reportable* as research. Instead, it constitutes a ‘specifically uninteresting’ adjunct to it (Garfinkel, 1967; Davis, 1971). Despite this, in the course of our study of the work of this particular group, we were able to gain insights into some of these less discussed aspects of statistical practice.⁹

One of the first things to note is that statisticians are rarely directly involved in the process of gathering the data they seek to analyse – statisticians’ offices are largely free from the accompanying clutter of survey research: tape recorders, piles of completed questionnaires or interview transcripts, and so on. Instead, their work normally begins *after* the primary data has been collected.¹⁰ Statistical work is, as a result, deeply collaborative, resting on a division of labour between those who produce, manage and can provide access to datasets (who we will refer to as ‘domain experts’ following those we studied) and those with sufficient statistical skill to exploit their affordances. The statistician’s role, within these collaborations, is to provide well formed, ‘ready-made’ tools, i.e. models but not solely models, for answering domain experts’ questions. These are questions the domain experts would not have been able to answer in those ways were they to rely solely on their own statistical expertise but which the statisticians, equally, could not answer without access to the data. Neither can proceed without the other: statistician and domain expert form a ‘symbiotic’ relationship (Kish, 1978: 1). The first aspect of statistical practice we discuss relates to the work undertaken in the initial stages of this collaborative process.

Interpreting Quantitative Data

When members of the group first acquired access to the databases that they were to use as testing-grounds for developing and trialling their methodological tools, they came at them ‘cold’. That is, they had to gain familiarity with the idiosyncrasies of the given database through a range of often highly skilled craft techniques acquired and developed through extensive experience that are often glossed under the heading of ‘the initial examination of data’ (Chatfield, 2002). No database is perfect, but it was the specific character of the imperfections of whatever database the researchers were working on that absorbed their time and attention in a project’s initial phases. Gaining familiarity with a database by slowly working on it so as to make it manipulable in the requisite ways included exploratory ‘probing’, ‘cleaning’, ‘recoding’, ‘programming’ and ‘testing’, operations often (misleadingly) seen as low-level, mechanical and intellectually undemanding.¹¹ What such characterizations miss is that this initial preparatory work is actually among the most important

the researchers perform. These 'elementary' operations are necessary preliminaries to more complex ones, and they provide the researchers with the opportunity to see not only how their data is organized but also to reflect on what the data is 'about'. This is not just about grasping the formal properties of matrices of numbers but about grasping why certain items have come to be included as part of those matrices and to what ends – only in this way can the ground be laid for the work of modelling.

By painstakingly undertaking these exploratory analyses and preparatory work – work which, in the case we examined, proceeded hand-in-hand with, and was mutually informed by, the creation of bespoke, customized code in specialized statistical software programmes that was purpose-written for the task of analysing the dataset at hand – statisticians are able to take 'alien' datasets produced by others and make them their 'own', moulding them in line with their own slowly developing sense of the calculative possibilities they present. In the course of learning how the dataset has been set up, what it includes and how it is structured, they thus become manipulable 'contextures', numerical fields within which model building can temporally unfold as a series of demonstrably and accountably reasoned operations, available for others to see and inspect (Lynch, 1991b).

In the initial phases, coming to know and understand the dataset, exploring its structural potentials and opening it up for statistical manipulation involved finding out, among other things, what items in the database were designed to index and so gain a feel for them as materials relevant to the work. Approached with this intent, datasets were not just grids of numbers with discoverable mathematical and probabilistic relations (see Figure 1) but socio-cultural artefacts with discernible characters and peculiarities all of their own. Many features of the datasets which were immediately intelligible to those who produced them were not available to the statistician working on them at one remove, and making them intelligible meant grasping ordinary aspects of the socio-cultural context(s) in which they were produced. As a consequence, the researchers had to engage in investigative work, making specific, limited inquiries into the use of particular measures so as to find out information about the society and culture from which those particular measures had been taken and thereby determine why particular measures had been included and what the dataset could therefore be said to capture.

One of the main methodological problems these researchers were investigating was how to amplify the inferential potential of longitudinal studies, and they were trying to do this by creating a generic methodology for linking differently organized longitudinal datasets together around measures common to them all (principally measures of age, ethnicity, socio-economic status, levels of education, etc.). They had access to an international dataset containing a large number of longitudinal studies, and those they ended up working with covered 12 countries

| | 0 | 100 | 200 | 300 | 400 | 500 | 600 |
|----|-----|-----|------|------|------|------|------|
| 0 | 0 | 864 | 1728 | 2592 | 3456 | 4320 | 5184 |
| 1 | 8 | 872 | 1736 | 2600 | 3464 | 4328 | 5192 |
| 2 | 17 | 881 | 1745 | 2609 | 3473 | 4337 | 5200 |
| 3 | 25 | 889 | 1753 | 2617 | 3481 | 4345 | 5208 |
| 4 | 34 | 898 | 1762 | 2626 | 3490 | 4354 | 5216 |
| 5 | 43 | 907 | 1771 | 2635 | 3499 | 4363 | 5224 |
| 6 | 51 | 915 | 1779 | 2643 | 3507 | 4371 | 5232 |
| 7 | 60 | 924 | 1788 | 2652 | 3516 | 4380 | 5240 |
| 8 | 69 | 933 | 1797 | 2661 | 3525 | 4389 | 5248 |
| 9 | 77 | 941 | 1805 | 2669 | 3533 | 4397 | 5256 |
| 10 | 86 | 950 | 1814 | 2678 | 3542 | 4406 | 5264 |
| 11 | 95 | 959 | 1823 | 2687 | 3551 | 4415 | 5272 |
| 12 | 103 | 967 | 1831 | 2695 | 3559 | 4423 | 5280 |
| 13 | 112 | 976 | 1840 | 2704 | 3568 | 4432 | 5288 |

Figure 1. The database as a numerical grid.

Source: Wikipedia. This file is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license.

and included data on around 70,000 participants. Educational attainment, one of the variables they were most interested in, had been coded in a range of different ways: sometimes as a continuous variable, i.e. number of years of education completed, sometimes using country-specific standards, i.e. 'junior high school' versus 'key-stage five', to take US and UK examples respectively. Drawing on this jumble of data, the aim was to reclassify participants across the various studies as possessing a low, medium or high level of education by creating a more orderly and well-behaved variable, something which would enable associations with other variables to be modelled as part of the testing of the generic methodology they wanted to develop (see Figure 2).

The difficulty was that the researchers had no immediate way of knowing what anyone within the societies covered by the studies would know by virtue of their (often historical) membership within them: namely, what does 'has completed ten years of education' amount to in Finland, for example, and does it mean the same thing as it does in France? The measures were indexical – their meaning depended on the

| Education | CLS (Australia) | HOPE (Scotland) | LASA (Netherlands) | OCTO (Sweden) | SATSA (Sweden) |
|-----------|--------------------|--------------------|-----------------------|------------------|-------------------|
| Low | 1–9 yrs | 1–9 yrs | 1–9 yrs | 1–7 yrs | Level 1 |
| Medium | 10–12 yrs | 10–12 yrs | 10–12 yrs | 8–12 yrs | Level 2-3 |
| High | ≥13 yrs | ≥13 yrs | ≥13 yrs | ≥13 yrs | Level 4 |

Figure 2. Making heterogeneous data more tractable, classifying education.

cultural context in which they were used – and, as a result, the researchers had to find ways of translating the variously coded educational data into a common format. They did so on the basis of discussions with ‘natives’ and by consulting historical sources that described their practices (some of those in the studies had been born in the early decades of the 20th century and educational systems in their countries had changed considerably in the period since). These investigations into (restricted features of) the culturally indexed properties of the category ‘years of education completed’ thus established general grounds for treating individuals as possessing a low, medium or high level of education relative to others at particular moments in those societies’ histories.

This interpretive, ‘anthropological’ inquiry into cultural categories hinged entirely on exploring the conventional meaning of those categories and the results of that inquiry were themselves understood and used by the statisticians in those conventional terms (see also Desrosières, 2009). That is, they were not interested in whether, in any actual case, someone had been fairly categorized as possessing above or below average levels of education but how in general such characterizations were used independently of their accuracy in any given case. The derived variable was thus in no way meant as a putative description of actual cases but functioned as an ethnographic translation of culturally specific and culturally recognized ways of talking about such matters in the broad.

This form of investigative activity may be dismissed as trivial, at first glance, perhaps because of its circumscribed and limited character, perhaps due to its sheer intuitive obviousness. Moreover, it isn’t especially ‘interesting’, not even to those involved. It is not what people want to talk about, statisticians included, when they talk about what they do. This would be, however, to under-estimate our elementary case. In recent work, Kullenberg (2011) describes a Swedish statistical institute’s attempts to encourage participation in its surveys and the strategies used to get participants to cooperate. Kullenberg talks of those statisticians’ reliance on a ‘knowledge of practice’ which is ‘manifested in action’ as a form of ‘legwork sociology’ all about getting the job of survey research done, something which presupposes ways of ‘knowing

about society' that are not themselves quantified or even perhaps quantifiable, being more akin to recipes (2011: 74–5, see also Schutz, 1944). Developing Kullenberg's analysis, our case provides an encapsulated, site-specific and locally leveraged tutorial in what it means to talk of the interpretive grounds of ordinary quantitative work (Kuhn, 1977; Gephart, 1988), a 'Geertzian' exercise in explicating and communicating meaning with research materials through an exploration of the use of conventional cultural categories. It shows how interpretation unproblematically goes hand-in-hand with quantification and so reveals the interdependence of both. That is not to elevate the work involved but rather to draw attention to the ordinary forms of practical sociological reasoning that provide the 'grounds' for anthropological and sociological inquiry too (Schutz, 1944; Garfinkel, 1967; Moerman, 1969; Garfinkel and Sacks, 1970; Geertz, 1973; Sacks, 1992).

However, despite highlighting qualitative dimensions of quantitative work, we would also do well not to go too far the other way and cede analytical priority to interpretation. If we fixated on interpretation we would lose sight of the interesting question of what goes without saying, what requires no interpretation at all. For one thing, note that the investigation of equivalences across dissimilar datasets presupposed that equivalences were there to be drawn. That is, this particular analysis assumed that the societies and cultures across which the data was collected were organized in sufficiently similar ways to make the search for equivalences meaningful. They looked, in other words, because they knew there was something to be looked for. No investigation was necessary to establish that, for example, the education level of children, again as conventionally understood, does not decrease as they make their way through school. Note too that no equivalent concerns were raised about the use of numbers across these societies and cultures. The fact that 10 could be taken to mean the same thing from one place to another for their practical investigative purposes was not something that *needed* to be questioned in these particular circumstances. As Wittgenstein reminds us with respect to number symbols on railway time-tables: 'What happens is not that this symbol cannot be further interpreted, but: I do no interpreting... As I have a railway time-table and use it without being concerned with the fact that a table can be interpreted in various ways' (1974: Part I, §99). In our example, it is not that the number systems in play could not have been interrogated in all manner of heterodox ways – the statisticians involved were certainly capable of such interrogation – but that the nature of the activity in which they were engaged did not call for it, something they knew without having to ask.

However, they were not *merely* or *blindly* presupposing such things: in taking those surroundings for granted, in treating them as grounds for further inference and action themselves in no need of independent justification, they were also putting them to work for their own specific

investigative purposes. In these and endlessly many other ways, statisticians rely on ‘commonsense knowledge of social structures’ (Garfinkel, 1967) in order to put those structures, here educational structures, on display and thereby make them observable-reportable. It is the specifics of such uses that are of interest. It is precisely because interpretation has its limits, that it is not boundless but bounded, that such analyses become intelligible – and recognizable – as accounts of social and cultural affairs at all.

Statistical work does not, then, involve the evacuation of meaning – an accusation commonly levelled at quantitative research – but efforts to capture meaning. In our case, the statisticians themselves were involved in the interpretive task of unravelling the interweaving of cultural and technical reasoning in the production of the artifacts they were working with. Moreover, the results of that inquiry were themselves understood and put to work by the statisticians in those very terms, with the same taken-for-granted assumptions employed in producing the data also used in the analysis. Putting society on display, then, is not merely a matter of calculation but of artfully exhibiting membership in social worlds inhabited and known in common (Schutz, 1944), something Sacks identified as central to ‘cultures’, as sets of procedures for producing recognizable actions, more broadly (Sacks, 1992: 227). It is precisely because, in being made use of, they *are* assumed – indeed, it is *in* their being assumed – that such things as ‘levels of education’ stand as features of a world in which it is possible, under at least some circumstances, to investigate regularities across societies despite cultural variation. The specific forms of social and cultural inquiry that we have drawn attention to in this section, forms of ‘ethnographic’ investigation that provide the backdrop to and which directly inform and shape the ‘technical’ work, are thus integral to the success of the statistical enterprise as it is in and through them that their analyses can be seen to unproblematically connect up with, and be seen to be speaking of and to, the everyday social and cultural worlds from which their materials derive.

Statistical Publics and Theatres of Proof

Our second example deals with another kind of ‘*verstehen*’ work that forms part of the statisticians’ craft but again is rarely explicitly remarked upon. Despite having their own areas of substantive expertise, members of the statistical unit whose work we studied found their status as methodologists was foregrounded in the course of their collaborations with other investigators (particularly collaborations with those outside their ‘home’ fields, as was the case in most of the projects we had opportunity to follow). As noted, statistical research typically involves the interplay of primary research and statistical analysis, and this is usually the responsibility of different personnel with specialized skills.

Beyond the level of co-production examined in the discussion above, another level of interplay is involved.¹²

In order to create statistical tools that are useful, usable and will get used, statisticians have to find ways of aligning others' analytical interests with their own. They need to see problems encountered in practice from their collaborators' perspectives, as this is part of devising solutions that anticipate and speak to the collaborators' problems while producing work that is statistically innovative. However, this means that to tell methodological stories, statisticians have to be in a position to tell substantive stories. They have to show that they can account for real-world phenomena in terms of any given method (or model) and so show others what the tools they have produced can do for them. This involves developing applications of instrumental value to non-statisticians who may not be in any position to 'appreciate' the novelty or sophistication of the statistical work itself. This is a far from straightforward task, and involves significant 'understanding work' in its own right.

As Winch puts it, the relation of social scientists 'to the performers of...[an] activity cannot be just that of observer to observed. It must rather be analogous to the participation of the natural scientist with his fellow-workers in the activities of scientific investigation' (1958: 87–8). And the problem of understanding 'fellow-workers' can be particularly acute, involving them in 'doing, recognizing and using ethnographies' (Garfinkel, 1967: 10) in related though nonetheless distinct ways to that discussed in the previous section. Alongside their ethnographic forays into the social, cultural and technical production of statistical data, statisticians also have to develop an understanding of those who would use the outcomes of their work in order to understand the sets of relevancies, often quite different to their own, within which their work might find a place or fit. This requirement means they must produce solutions that not only handle defined statistical problems but do so in demonstrably relevant ways, enabling others, as in the case of the unit we studied, to make sense of the three-dimensional growth of cancer cells in the human body, of spatio-temporal differentiation in income patterns or in voting by social and ethnic group across small areas. Models had to interest, to capture the attention: they had to offer a ready answer to the question, 'what will this do for me?'. Developing applications that others could see the demonstrable relevance of was far more of an art, a matter of trial and error, than a science. Technical sophistication or novelty on its own was not enough. The capacity to pinpoint arresting applications among a large range of potential applications (which would be reviewed and discussed at the start of a project in conjunction with the databases to which the statisticians had access to or had worked on before) was acquired through hard work and often difficult first-hand experiences of failure, as we shall see in the example that follows.

The project from which that second example is taken was an attempt to offer new ways of modelling bias due to survey non-response. The problem at its core was how to develop a systematic approach to analysing the problem of missing data. Missing data is regarded as a species of selection bias which arises due to partial or total non-response in large-scale surveys: i.e. respondents either skipping particular questions or not responding to the survey at all. The stock solution has long been list-wise deletion, simply erasing the entry, i.e. the record of those persons who have failed to answer, containing wholly or partly missing information. This is unsatisfactory for various well-recognized statistical reasons. Are the people who do not answer of a kind? Does their erasure mean inferences drawn from what remains will be distorted in unknown but potentially serious ways? The answer to both questions is thought to be yes, but the standard approach leaves this impossible to determine, with suspicions of bias haunting the conclusions made. The approach taken on this particular project was to develop a different kind of solution to the problem. That solution turned on treating non-response itself as a variable and building hierarchical models that drew on multiple over-layered data-sources to 'fill in' or 'impute' missing variables, using complex random simulations to build synthetic entries for those who would otherwise have been consigned to the statistical void. The overall goal was to communicate a series of methodological messages, primarily:

- missing data matters;
- doing nothing is doing something (i.e. it has effects on the inferences drawn);
- there are a variety of ways of dealing with the problem, not simply list-wise deletion;
- Bayesian modelling compares very favourably with the most effective of the alternatives available.

Nonetheless, despite the methodological if not reflexive aspect of the project (a statistical researcher studying the practices of statistical research statistically), the study had to remain anchored to substantive questions. In order to make the methodological point (and hence make the case for the Bayesian solution to the problem of missing data), the researchers involved crafted demonstrations that examined the effects of non-response on analyses of exposure to certain chemicals and birth-weight as well as the relationships between ethnicity, education, marital status and income. It was these demonstration cases that enabled the researchers to take the research 'public' (Gieryn, 2002), and stage their own 'theatres of proof' designed to persuade quantitative domain experts of the relevance of their work (Shapin and Schaffer, 1985).

At the time of our research, this particular study had been underway for around a year and a half. It was the most methodologically focused of all the projects we studied and involved a great deal of what other statisticians, let alone domain experts, regarded as particularly high-level and

esoteric modelling. In the phase we witnessed, the focus was switching from modelling over to demonstrating the practical gains of that modelling work. We were able to attend a series of preparatory meetings in the run-up to one of the first public presentations of results from the research, a workshop held by the principal external collaborator the team were working with as part of this project as well as, finally, the reporting-back/debriefing session which followed it.

A major aspect of the build-up to the presentation of this work was the selection of graphs and plots that were thought to best express the power of the solution the researcher was proposing. Statistics are often presented as 'speaking for themselves'. However, such 'speaking' is of course carefully designed, i.e. requires work (cf. Tufte, 2006). This is exactly what we observed in this and other project meetings. In an average meeting, 50-plus 'plots' would be discussed. These discussions did not hinge on the 'correctness' of the graphic displays (which was assumed) but on their usefulness for making the character of the work they were to exhibit perspicuous in certain ways, where the question of what could be seen in any given display was itself clarified in the course of comparing and contrasting the merits of a range of candidate displays. The purpose of presenting these displays was a central consideration in this: it was not that, in isolation, these were 'good' or 'better' than other graphs and plots but that these were 'good' or 'better' relative to the occasion, audience and message the researchers were trying to convey. If no viable candidates could be found for those particular purposes, the researchers would go back to the statistical drawing board. Graphic displays of statistical findings were, then, being continually developed and revised with an eye to how they would be or had been responded to. Potential users therefore appeared as 'scenic features' of the development process over its course (Sharrock and Anderson, 1994).

Following a run-through with members of the unit which threw up no particular problems, the workshop where the work was presented 'for real' did not go particularly well. The researcher received limited feedback from the audience and a less than enthusiastic response from the external collaborator. It was not that the fundamentals of the modelling had been challenged. Worse, they had not really come up for discussion at all. The research had not spoken to those at the workshop (it did not 'speak their language'; Bowman, 2007). As a consequence, the researcher voiced a keen sense of what we might call 'professional disappointment'. The audience had been unwilling to engage with what made the research innovative in methodological terms because they were unable to identify any particularly visible substantive reward in doing so – they could not see what was in it for them. As a result, the discussion focused on what the researcher saw as the least statistically interesting or significant parts of what they were attempting to do. The response, perhaps

unsurprisingly, was read as a sign of indifference to the central methodological messages.

The debriefing raised interesting questions related to the entanglement of methodology, application and audience in statistical research. In the course of their discussion, the unit's director assured the researcher whose work this was that it was a common experience, highlighting the need to see misunderstandings as themselves an issue for interpretation in the process. As the director put it:

That sounds like experiences I've had. There's some discussion of the application but no interest in the methodology... It's quite hard, and this comes back to what we've said before, to respond to the 'so what' [question]. Because you don't demonstrate huge impacts of dealing with missing data [in the presentation], people say 'well, at the end of the day there's an awful lot of complex modelling involved and what's the gain?'. I know you've tried to think about that and say you've got to understand the assumptions you're making, but it's difficult.

Without a persuasive substantive story to tell, the methodological point was not taken up and recognition of this meant substantially reworking the analysis (over a period of some months) in order to make it seeably relevant to others. This involved anticipating better the tasks those who might use the techniques would want them to do and finding a new application which made its utility more immediately apparent. This provides another sense in which understanding work is crucial to the enterprise of quantification. Not only do statisticians have to think carefully about the perspective from which others gauge their work but, as this case also shows, ultimately the models have to do the talking in research of this kind: they have to demonstrably deliver. Where models do deliver, the understanding work embedded in the model is successful precisely where there is no need for it to be made explicit. Failure is most apparent where the import of what has been done, like a joke that has fallen flat, has to be explained at length to those it was directed to.

Once again, we see that statisticians must possess a grasp of social and cultural worlds, this time the worlds of those who would use the outcomes of their work, a set of competencies that are acquired as part of doing statistical work. Once again, and with a slight shift of emphasis, we see a set of methods for exhibiting commonsense knowledge of social structure in the making of 'successful research' – knowledge which is ubiquitous, pervasive and relied upon without explicit formulation or discussion. Indeed, 'old hands' writing on this very subject have denied such knowledge of the 'context of justification' could be formalized, treating it instead as a craft form central to the practice, one which resists formalization and is only capable of being acquired through engaging in

that practice (Chatfield, 2002). In sum, orienting to others in the production of statistical ‘accounts’ points to another less visible ‘Winchean’ way in which understandings of ‘social structure’ are also put on display in the statistician’s work, as well as the different kinds of social and cultural investigation it is developed through.

Conclusion

We have presented two examples of the ways in which ‘the social structures of everyday activities’ are exhibited in and through statistical work: via, firstly, displayed understandings of the social and cultural contexts – the everyday social and cultural worlds that statisticians are also members of – within which databases as numerical artifacts are produced; and, secondly, displayed understandings of the users of, and uses to which, statistical models may be put. The capacity to do either well was tied to a diffuse set of competencies involving what we have termed ‘understanding work’, lines of social and cultural inquiry undertaken as part of and not simply in parallel with model and application building, with one informing the other. We would suggest that the specificities of this work – what made it the work it was – are to be located here, not in whether it was more or less qualitative or quantitative, more or less interpretive or calculative. Those labels, while useful under some conditions, have little value as guides to what these and, we would venture, other researchers routinely get up to. If we are to develop studies of the ‘social life of methods’ as part of a more empirically-oriented and descriptive ‘sociology of social scientific knowledge’ (Maynard and Schaeffer, 2000; Savage, 2009, 2013), we would suggest that we cannot proceed by taking ready-made distinctions, such as that between qualitative and quantitative research, for granted, even for critical purposes, as the starting point for our inquiries. Instead, following Garfinkel, we suggest that a more profitable way of proceeding is to treat such distinctions as problematic from the off and attempt to ascertain what, if any, relevance they have at the level of practice, something those interested in the social life of method have had little concretely to say about as yet.

Investigating the praxiological grounds of ‘methods work’ is not necessarily a straightforward matter. The ‘low-level’ (we might say ‘foundational’) aspects of such things as ‘making a database your own’ and ‘developing analyses demonstrably relevant to others’ will be pervasive features of statistical research more broadly, Bayesian or not. However, the commonplace character of that work means it will also go (at least in large part) ‘seen-but-unnoticed’, liable to be overlooked by those coming at researchers’ work through the accounts given in their publications. This is not a mark of the researchers’ duplicity or negligence – it shows that such practices, while central to what researchers do, are not constituent of the relevant disciplinarily reportable parts of their work.

Instead, they are firmly within that domain of 'knowledge' that, for example, every competent statistician knows without having to be told and which there is much more to than can be said 'in so many words' (Garfinkel, 1967).

The examples we drew from our case study also show that there is much more to 'method' than is typically allowed. Perhaps such practices have been overlooked to date because they simply have not seemed 'interesting' enough to be explored in any great depth (Davis, 1971). Precisely because they are treated as 'banal' – the chores of statistical work – by statisticians themselves, it is easy to miss the different ways in which statistical work, for instance, presents and proceeds on the basis of an accomplished understanding of social practice. Such understandings are seamlessly stitched into and exhibited by quantitative practices – they generate no friction, they are not 'out of place' but are entirely 'in place' within them. Statisticians do not worry about this aspect of their work, by and large, but can get on with the job in the knowledge that competence in them can be taken on trust – until, that is, things go wrong, as in our second example. The ways in which statistics put society on display are thus quite deeply embedded in statistical practice – it is for that very reason that investigations of them open up unexplored methodological worlds.

Our elementary examples are thus important in the first instance, we would argue, because they allow us to gain an appreciation of the specificities of statistical work as work. They also, however and as a result, allow us to see the limits of current work on the social life of methods, particularly Law, Ruppert and Savage's analysis of their 'doubly social' character. On the basis of what was only a small study, a prolegomenon of sorts to a more extended and in-depth body of research on these issues, we have been able to start to point to the heterogeneous collection of activities upon which statistical method rests, just two aspects of which we have discussed here. Methods are social; they are both in and of societies, as Law, Ruppert and Savage suggest, but not in the same way. Law, Ruppert and Savage establish a problematic, but it is one the pursuit of which leads rather quickly to its own dissolution because we soon find, when we go in search of method, that no such overarching thing exists in any simple sense. For that reason, recognizing that statistical research, like other forms of social scientific work, is social through-and-through can only be a starting point, a heuristic – it does not itself constitute an account of the practices in question.

Acknowledgements

This research was made possible by the ESRC through funding to the NCRM, project code RES-576-25-0022. We would like to thank the editor and the anonymous reviewers for their helpful comments on the article, which we have taken up, and the statisticians

whose work we focus on for their time, their patience and their generosity in facilitating our study.

Notes

1. We cannot do justice to the volume and variety of these studies but a brief survey, starting in the 20th century, would include reference to Husserl (1936), Hacking (1965, 1975, 1990; but see also 2001, 2002), Foucault (e.g. 1975), MacKenzie (1981), Livingston (1986), Porter (1986, 1996), Latour (1987), Gephart (1988; Gephart and Smith, 2009), Lynch (e.g. 1991a; 1991b; 1993), Potter, Wetherell and Chitty (1991), Desrosières (e.g. 1998), Bowker and Star (1999) and, more recently, to the likes of Stigler (2002), Kirby (2005), Ruppert (e.g. 2007), Espeland and Stevens (2008), Law (2009), Saetnan, Lomell and Hammer (2011) and Courgeau (2012), among others.
2. The hallmark of ‘as if’ (or ‘might as well’) accounts in this area is the reconstruction (or ‘redescription’) of statistical or mathematical practice on the basis of the products of that practice rather than a study of the practice itself. John Law’s ‘Seeing Like a Survey’ (2009) is an archetype of a ‘recovered’, ‘imagined’ or ‘projected’ description, i.e. one which proceeds ‘as if things had been (or ‘might as well have been’) done in ways the analyst regards as plausible (see Sacks, 1963).
3. For example Sacks (1992); Cicourel (1964); Livingston (1986, 1987); Gephart (1988); Goodwin (1994); Greiffenhagen and Sharrock (2006); Lynch et al. (2008); Martin and Lynch (2009); Gephart and Smith (2009).
4. Features of their practices which, crucially, statisticians *themselves* use to help others orient to their work as work – in our case, these examples are what we were pointed to so that we could gain a better appreciation of what the statisticians whose work we were following were actually doing.
5. But for an introduction see Hacking (2001), and for applications see Lynch et al. (2008).
6. And was both data and computationally intensive long before it became fashionable.
7. See Lynch et al. (2008: 190–219) for a particularly interesting example of a Bayesian approach to the quantification of ‘commonsense knowledge’.
8. We would echo Gilbert and Mulkey’s (1984) insistence that ‘off-hand’, ‘light-hearted’ or ‘joking’ comments often say a great deal about the ‘serious’ work they comment on.
9. Note, however, that while they did not talk about it, that does not mean they did not see, recognize, understand, appreciate, think about, reflect upon or otherwise take time to reason such matters through. Failures to do so were accountable, i.e. they were picked up and were both visible-reportable *and* correctable so that consideration of them could be eliminated from publication – because, again, they were ‘not interesting’ rather than because they needed to be concealed. It was not, then, that they had no conception of such things, they just did not discuss it publicly or reflectively topicalize the presence of these matters in their work. Not all absences constitute failures. There is not enough time in the world to say everything, and were anyone to try, no-one would stick around to listen – ethnomethodologists partially excepted.

10. Including expert elicitation, which Bayesians undertake as part of modelling already gathered data.
11. Chatfield (2002) is very good on this point, underscoring the overlooked importance of the pragmatic side of statistical work to the stability of the enterprise as a whole.
12. See Desrosières (1997, 2001, 2009) for discussions of closely related sets of issues, discussions we have drawn on throughout this section but develop in a different way.

References

- Adkins L and Lury C (eds) (2012) *Measure and Value*. London: Wiley-Blackwell.
- Back L and Puwar N (2012) *Live Methods*. London: Wiley-Blackwell.
- Bowker G and Star SL (1999) *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Bowman AW (2007) Interdisciplinary research: The importance of learning other people's language. *AStA: Advances in Statistical Analysis* 91(4): 361–365.
- boyd d and Crawford K (2012) Critical questions for big data: Provocations for a cultural, technological and scholarly phenomenon. *Information, Communication and Society* 15(5): 662–679.
- Bryman A (1988) *Quantity and Quality in Social Research*. London: Unwin Hyman.
- Chatfield C (2002) Confessions of a pragmatic statistician. *The Statistician* 51(1): 1–20.
- Cicourel AV (1964) *Method and Measurement in Sociology*. Glencoe, IL: The Free Press.
- Courseau D (2012) *Probability and Social Science: Methodological Relationships between the Two Approaches*. London: Springer.
- Davis MS (1971) That's interesting: Towards a phenomenology of sociology and a sociology of phenomenology. *Philosophy of the Social Sciences* 1(4): 309–344.
- Desrosières A (1997) The administrator and the scientist: How the statistical profession has changed. *Statistical Journal of the United Nations Economic Commission for Europe* 14(1): 31–50.
- Desrosières A (1998) *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.
- Desrosières A (2001) How real are statistics? Four possible attitudes. *Social Research* 68(2): 339–355.
- Desrosières A (2009) How to be real and conventional: A discussion of the quality criteria of official statistics. *Minerva* 47: 307–322.
- Diaconis P (2009) The Markov Chain Monte Carlo revolution. *Bulletin of the American Mathematical Society* 46(2): 179–205.
- Efron B (2013) A 250-year argument: Belief, behavior, and the bootstrap. *Bulletin of the American Mathematical Society* 50(1): 129–146.
- Elliot M, Purdam K and Mackey E (2013) *Data Horizons*. London: ESRC.
- Espeland WN and Stevens ML (2008) A sociology of quantification. *European Journal of Sociology* 49(3): 401–436.

- Fienberg SE (2006) When did Bayesian inference become 'Bayesian'? *Bayesian Analysis* 1(1): 1–40.
- Foucault M (1975) *Discipline and Punish*. London: Penguin.
- Gane N (2012) Measure, value and the current crises of sociology. In: Adkins L and Lury C (eds) *Measure and Value*. London: Wiley-Blackwell, pp. 151–173.
- Garfinkel H (1967) *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
- Garfinkel H and Sacks H (1970) On formal structures of practical actions. In: McKinney JC and Tiryakian E (eds) *Theoretical Sociology*. New York: Appleton Century Crofts, pp. 338–366.
- Geertz C (1973) *The Interpretation of Cultures*. New York: Basic Books.
- Gelman A and Shalizi CR (2013) Philosophy and the practice of Bayesian statistics. *British Journal of Mathematical and Statistical Psychology* 66(1): 8–38.
- Gephart RP (1988) *Ethnostatistics: Qualitative Foundations for Quantitative Research*. London: SAGE.
- Gephart RP and Smith RS (2009) An invitation to ethnostatistics. *Sciences de Gestion* 70: 85–102.
- Gieryn T (2002) Three truth-spots. *Journal of the History of the Behavioral Sciences* 38(2): 113–132.
- Gilbert N and Mulkey M (1984) *Opening Pandora's Box: A Sociological Analysis of Scientists' Discourse*. Cambridge: Cambridge University Press.
- Glymour C (1981) *Theory and Evidence*. Princeton, NJ: Princeton University Press.
- Goodwin C (1994) Professional vision. *American Anthropologist* 96(3): 606–633.
- Greiffenhagen C and Sharrock WW (2006) Logical relativism: Logic, grammar, and arithmetic in cultural comparison. *Configurations* 14(3): 275–301.
- Greiffenhagen C, Mair M and Sharrock WW (2012) From methodology to methodography: A study of qualitative and quantitative reasoning in practice. *Methodological Innovations Online* 6(3): 93–107.
- Greiffenhagen C, Mair M and Sharrock WW (forthcoming) Methodological troubles as problem and phenomenon: Ethnomethodology and the question of 'method' in the social sciences. *British Journal of Sociology*.
- Hacking I (1965) *The Logic of Statistical Inference*. Cambridge: Cambridge University Press.
- Hacking I (1975) *The Emergence of Probability*. Cambridge: Cambridge University Press.
- Hacking I (1990) *The Taming of Chance*. Cambridge: Cambridge University Press.
- Hacking I (2001) *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.
- Hacking I (2002) *Historical Ontology*. Cambridge, MA: Harvard University Press.
- Husserl E (1936) *The Crisis of the European Sciences and Transcendental Phenomenology: An Introduction to Phenomenology*, trans. Carr D. Evanston, IL: Northwestern University Press.
- Kirby V (2005) Just figures? Forensic clairvoyance, mathematics and the language question. *SubStance: A Review of Theory and Literary Criticism* 35(2): 1–24.

- Kish L (1978) Chance, statistics and statisticians. *Journal of the American Statistical Association* 73(361): 1–6.
- Kuhn T (1977) *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago, IL: University of Chicago Press.
- Kullenberg C (2011) Sociology in the making: Statistics as mediator between the social sciences, practice and the state. In: Saetnan AR, Lomell HM and Hammer S (eds) (2011) *The Mutual Construction of Statistics and Society*. London: Routledge, pp. 64–78.
- Latour B (1987) *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press.
- Latour B (2010) Tarde's idea of quantification. In: Candea M (ed.) *The Social after Gabriel Tarde: Debates and Assessment*. London: Routledge, pp. 145–162.
- Law J (2009) Seeing like a survey. *Cultural Sociology* 3(2): 239–256.
- Law J, Ruppert E and Savage M (2011) The double social life of methods. *Working Paper Series: Working Paper No. 95*. Milton Keynes: ESRC Centre for Research on Socio-Cultural Change (CRESC).
- Livingston E (1986) *The Ethnomethodological Foundations of Mathematics*. London: Routledge & Kegan Paul.
- Livingston E (1987) *Making Sense of Ethnomethodology*. London: Routledge & Kegan Paul.
- Lury C and Wakeford N (eds) (2012) *Inventive Methods: The Happening of the Social*. London: Routledge.
- Lynch M (1991a) Method: Measurement – ordinary and scientific measurement as ethnomethodological phenomena. In: Button G (ed.) *Ethnomethodology and the Human Sciences*. Cambridge: Cambridge University Press, pp. 77–108.
- Lynch M (1991b) Laboratory space and the technological complex: An investigation of topical contextures. *Science in Context* 4(1): 51–78.
- Lynch M (1993) *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science*. Cambridge: Cambridge University Press.
- Lynch M, Cole S, McNally R and Jordan K (2008) *Truth Machine: The Contentious History of DNA Fingerprinting*. Chicago: University of Chicago Press.
- Machiavelli N (1910 [1532]) *The Prince*, trans. Thomson NH. New York: P.F. Collier & Son.
- Mackenzie D (1981) *Statistics in Britain, 1865–1930: The Social Construction of Scientific Knowledge*. Edinburgh: Edinburgh University Press.
- Mair M, Greiffenhagen C and Sharrock WW (2013) Social studies of social science: A working bibliography. *National Centre for Research Methods Working Paper Series*, No. 3219.
- Majima S and Moore N (2009) Introduction: Rethinking qualitative and quantitative methods. *Cultural Sociology* 3(2): 203–216.
- Martin A and Lynch M (2009) Counting things and people: The practices and politics of counting. *Social Problems* 56(2): 243–266.
- Maynard DW and Kardash T (2006) Ethnomethodology. In: Ritzer G (ed.) *Encyclopaedia of Sociology*. Oxford: Blackwell, pp. 1483–1486.
- Maynard DW and Schaeffer NC (2000) Toward a sociology of social scientific knowledge: Survey research and ethnomethodology's asymmetric alternates. *Social Studies of Science* 30(3): 323–370.

- Mayo DG (2011) Statistical science and philosophy of science: Where do/should they meet in 2011 (and beyond)? *Rationality, Markets and Morals* 2: 79–102.
- McGinn M (2010) Learning to use statistics in research: A case study of learning in a university-based statistical consulting centre. *Statistics Education Research Journal* 9(2): 35–49.
- Moerman M (1969) A little knowledge. In: Tyler S (ed.) *Cognitive Anthropology*. New York: Holt, Rinehart and Winston.
- Mohr JW and Ghaziani A (2014) Problems and prospects of measurement in the study of culture. *Theory and Society*. DOI: 10.1007/s11186-014-9227-2.
- Moretti F (2013) *Distant Reading*. London: Verso.
- Porter TM (1986) *The Rise of Statistical Thinking, 1820–1900*. Princeton, NJ: Princeton University Press.
- Porter TM (1996) *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
- Potter J, Wetherell M and Chitty A (1991) Quantification rhetoric: Cancer on television. *Discourse and Society* 2(3): 333–365.
- Rogers R (2013) *Digital Methods*. Cambridge, MA: MIT Press.
- Ruppert E (2007) Producing population. *Working Paper Series: No. 37*. Milton Keynes: ESRC Centre for Research on Socio-Cultural Change (CRESC).
- Ruppert E (2013) Rethinking empirical social sciences. *Dialogues in Human Geography* 3(3): 268–273.
- Ruppert E and Savage M (2011) Transactional politics. *Sociological Review* 59(s2): 73–92.
- Ruppert E, Law J and Savage M (2013) Reassembling social science methods: The challenge of digital devices. *Theory, Culture & Society* 30(4): 22–46.
- Sacks H (1963) Sociological description. *Berkeley Journal of Sociology* 8: 1–16.
- Sacks H (1992) *Lectures on Conversation*. Oxford: Blackwell.
- Saetnan AR, Lomell HM and Hammer S (eds) (2011) *The Mutual Construction of Statistics and Society*. London: Routledge.
- Sandelowski M, Voils CI and Knafl G (2009) On quantitizing. *Journal of Mixed Methods Research* 3(3): 208–222.
- Savage M (2009) Contemporary sociology and the challenge of descriptive assemblage. *European Journal of Social Theory* 12(1): 155–174.
- Savage M (2013) The social life of methods: A critical introduction. *Theory, Culture & Society* 30(4): 3–21.
- Savage M and Burrows R (2007) The coming crisis of empirical sociology. *Sociology* 41(5): 885–899.
- Schutz A (1944) The stranger: An essay in social psychology. *American Journal of Sociology* 49(6): 499–507.
- Shapin S and Schaffer S (1985) *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ: Princeton University Press.
- Sharrock W (2000) The fundamentals of ethnomethodology. In: Smart B and Ritzer G (eds) *Handbook of Sociological Theory*. London: SAGE, pp. 249–259.
- Sharrock W and Anderson B (1994) The user as a scenic feature of design space. *Design Studies* 15(1): 5–18.
- Stigler SM (2002) *Statistics on the Table: The History of Statistical Concepts and Methods*. Cambridge, MA: Harvard University Press.

- Toulmin S (1953) *The Philosophy of Science: An Introduction*. London: Hutchinson's University Library.
- Tashakkori A and Teddlie C (eds) (2010) *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks: SAGE.
- Tufte ER (2006) *Beautiful Evidence*. Cheshire, CT: Graphics Press LLC.
- Verran H (2012) Number. In: Lury C and Wakeford N (eds) *Inventive Methods: The Happening of the Social*. London: Routledge, pp. 110–124.
- Winch P (1958) *The Idea of a Social Science and Its Relation to Philosophy*. London: Routledge.
- Wittgenstein L (1974) *Philosophical Grammar*. Oxford: Blackwell.

Michael Mair is Senior Lecturer in Sociology at the University of Liverpool. His most recent research falls into two main areas: politics, government and the state; and the methodology and philosophy of research. The focus of that work includes the politics of accountability in different settings, including military investigations into friendly fire deaths, and work on methodological practice in the social sciences, incorporating research on qualitative, quantitative and 'digital' methods. Recent publications include a bibliographical review of *Social Studies of Social Science* for the ESRC's National Centre for Research Methods' Working Paper Series and a 2015 article on 'Methodological troubles as problem and phenomenon' in the *British Journal of Sociology*.

Christian Greiffenhagen is Senior Lecturer in Sociology at Loughborough University. His research is concerned with studying scientific work and new technologies 'in use', e.g. the introduction of educational software into classrooms, the workaday practices of professional mathematicians, or the use of public access wi-fi services to address digital exclusion. Recent publications include 'The materiality of mathematics' (*British Journal of Sociology*), 'Visual grammar in practice' (*Semiotica*), and 'Does mathematics look certain in the front, but fallible in the back?' (*Social Studies of Science*).

W.W. Sharrock is Professor in Sociology at the University of Manchester. In a career spanning nearly 50 years at Manchester, he has explored two central themes – the relevance of fieldwork and an understanding of ordinary language for an understanding of social practice and the re-specification of social theory – pursuing them across a huge variety of settings, from ordinary scenes of everyday social life through to complex domains of practical action and reasoning in various academic and industrial work situations, including sociological research.