Simpson S. and Dorling D. (1999) Statistics and 'the truth', Chapter 47 in D. Dorling and S. Simpson (eds) Statistics in society: the arithmetic of politics, London: Arnold.

47

Conclusion: Statistics and 'the Truth'

Stephen Simpson and Daniel Dorling

Social statistics are a social product

Statistics are a social product. In this chapter we discuss what this statement means for statistics and what it implies for people who want to change the world they live in. We draw on the chapters in this book to provide examples of how statistics are created and interpreted within society.

In a simple sense, statistics are a social product simply because they are produced by people. But they are also firmly located in the aims and tensions of the society that produces them – whether expressed by organisations of government, trade, or campaigns. In literature it is said that every text has a context. With statistics it is not just that what is discovered depends on the society from which those numbers are drawn. The methods and data that are used themselves are shaped by their context. Donald MacKenzie explores this context for the origins of modern statistics in Chapter 7.

Take homelessness. The official count of homeless people should, one would think, determine the amount of resources committed to tackle the problem. It is certainly taken by social commentators to indicate how the scale of the problem differs geographically. However, the official count of homeless people includes only those who apply successfully to local authorities for rehousing. This is reported quarterly to central government. As Rebekah Widdowfield shows in Chapter 22, it severely underestimates the scale of the problem while also overstating the numbers of people actually living on the street. The definition of these official statistics is being changed by government as we write, so we will have little idea of how the extent of homelessness is itself changing. It is very difficult to counter the official homelessness statistics, because other official statistics such as the national census also severely underestimate the homeless, as Ian Diamond shows in Chapter 2: those without a usual address are naturally difficult to count during any household enquiry. Homeless people can probably best recognise their conditions and count their number, but they have no incentive, nor the means, to do so. It remains a fact that homeless statistics do not estimate the number of homeless people.

The view that statistics are a product of their context is widely held. In *Demystifying Social Statistics* Dot Griffiths, John Irvine and Ian Miles (1979) distinguish

four views of critics of statistical practice and more generally of scientific practice:

First, an 'anti-science' approach would dismiss every statistic because it 'inevitably turns people into objects to be manipulated and controlled and is contrary to basic human values'. This view is widespread as a social attitude. Some people would reject out of hand attempts to measure individuals' or society's progress as being fundamentally restrictive.

Second, an 'alternative technology' approach charges itself with 'convincing people of the social, environmental and health hazards of such high technologies as supersonic aircraft, nuclear reactors, factory farms and automated production lines, and by providing them with alternatives . . . small-scale and controlled by the individual or community'. Within statistics perhaps the most successful alternative approaches have been developed by the Community Operations Research Network described by Charles Ritchie, Ann Taket and Jim Bryant (1994). Jeff Evans and Ivan Rappaport pursue the statistical needs of communities and the concerned citizen in this book (Chapters 9 and 10). In these cases, alternative statistics are seen not as replacing traditional statistics, so much as tools necessary for research in settings where traditional methods are not appropriate.

Next, 'social responsibility' in science sees statistics 'as ethically neutral, asocial, bodies of knowledge and techniques: they could be used for good purposes or abused for bad ones'. This is a widespread attitude among statisticians themselves. Socially responsible statisticians argue for appropriate codes of practice for statistical work, aiming to limit the political misuse of statistics. They criticise and poke fun at abuses of statistics, with classic published examples in WJ Reichmann's *Use and Abuse of Statistics* (1964) and Darrel Huff's *How to Lie with Statistics* (1973), as well as a regular column in the Royal Statistical Society's newsletter.

The fourth view, labelled 'radical science' and to which we would subscribe, recognises the success of the first three views in achieving changes to statistical policy or practice, in particular circumstances, but goes on to point out that if statistics are a product of society then they are never neutral. They cannot be ignored nor can they be substantially changed unless society itself is changed. People who wish to make statistics more open (Simpson and Dorling 1994, Dorling 1995) more democratically linked to the priorities of local and other communities, less dominated by government and trade, are allies of anyone who wants to change society in these ways.

In order to understand why some statistics are believed and acted upon, while others are ignored, we need to know something about the process by which statistics in society are *paid* for. Four things are necessary to produce a well-known statistic:

- The statistic has to have a *purpose*, be seen to be lacking, be wanted, be commissioned, be called for, asked for, needed, demanded, by someone with the power and the money to express a need for it and ensure that it will be provided.
- The data have to be *assembled*, collected, brought together, gathered; this often requires resources, or control over an administrative system to extract the statistics from information already recorded.
- The statistic has to be *interpreted*, analysed, reasoned, understood, by researchers who use methods developed to answer previous questions, or who

develop new methods for the purposes of those who commission statistical work.

- The statistic has to be *described*, reported, presented, disseminated or communicated by the researchers or the commissioners to an audience which must be convinced of the results in order to support policy decisions.

These four processes are all required to produce a statistic. Social statistics have to be *paid* for. The processes are not always so clearly divided into four stages, but the reality is illustrated in detail in most of the chapters of this book; for example Paul Dunne in Chapter 43 describes the construction of statistics used to determine the development of military strength, while Alan Freeman in Chapter 41 reanalyses national economic accounts with a new purpose. Jeff Evans and Ivan Rappaport in Chapter 9 discuss the statistical needs of communities, while Ian Plewis in Chapter 31 seeks statistics that could make sense of educational achievement.

Statistics in the natural sciences are also a product of these processes. Statistics about society are more often used in the political processes of debating the formulation and implementation of policy. Inevitably they are themselves political. To talk of statistics as political because they are a social product does not invalidate them. In particular, we are not saying that statistics are invented to reflect someone's narrow interest: examples of inventions within research, such as Cyril Burt's twin studies to 'prove' hereditary intelligence, are famous but few (Norton and Mackintosh, 1995). In a democratic society different priorities call for different types of statistics to be collected, analysed and reported. Those who demand genuine and useful debate will also recognise that apparently conflicting statistics can be judged only if they are honestly collected and reported, and the processes of their construction are fully described.

The issues raised by this book are therefore of importance to research workers as much as to those who seek to understand the results of research. It seems useful to summarise those issues within the four stages of statistical research set out above; many of the issues may in fact apply to more than one stage.

Purpose: wanted, commissioned, called for, asked for, needed, demanded

Government statistics, by their nature, address existing government policies. They do not monitor new social issues, nor past and present issues that are not addressed by current policy. In this book Alison Macfarlane, Jenny Head and Allyson Pollock (Chapters 26 and 29) give examples from health and social care, Humphrey Southall, David Gordon and Pauline Heslop (Chapters 20 and 40) from poverty, and Christina Pantazis and David Gordon (Chapter 24) from crime. This is also the reason that only paid work is measured in the Labour Force Survey and the national accounts, and why the majority of government statistics are about business and trade.

People's priorities for statistics are often different from those of government. For instance, a common understanding of the local need for jobs would be based on the number of people in an area, while a company will focus on the skills that it needs. The difference between the jobs people want and the people that

employers want creates unemployment, commuting and labour shortages, which are examined by government planners as problems without resolution as Jamie Peck makes clear (Chapter 39).

Government and other bodies can decide to commission a statistic as a means of doing nothing, or to give an image of doing something. While it is wise to plan research very carefully, feasibility studies and investigations are also a means to delay implementation of new approaches while the political support for them is gauged. How often is 'We shall investigate' the cry of those who should have known all along?

Assembled, collected, brought together, gathered

There is a huge degree of control by government over access to the raw materials needed to construct social statistics once they have been deemed to be lacking. Government administrative registers are often the only counts of the population; but they are also often particularly poor representations of social conditions. In this book Cecilio Mar Molinero (Chapter 35) discusses examples in special education. Ray Thomas and Anne Green (Chapters 36 and 37) demonstrate that unemployment statistics do not measure the number of people who need a job. David Gordon and Pauline Heslop (Chapter 20) refer to the poverty of poverty statistics, while Richard Bate and Rebekah Widdowfield (Chapters 42 and 22) show that housing statistics do not measure homelessness or housing need. Government undertakes a number of national social surveys. It is the only body to have the resources to do so regularly. Charlie Owen and Angela Dale (Chapters 3 and 4) discuss these government surveys in this book.

Social problems often become fragmented by specialist professional approaches. Statistical analysis is one of those professional approaches. Bound by rules developed explicitly in textbooks and journals, it provides some independence from political interference. This specialisation also removes statisticians from an understanding of the people and the issues they are investigating. Even more frequently, the researcher is much more familiar with a computer screen than with the human subjects of research. One of the most important conclusions of several chapters of this book is that a variety of research approaches are required to understand social development and needs. Dan Wright (Chapter 8), in discussing psychology, stresses that qualitative data about private events must supplement quantitative data that are more easily recorded, Diane Perrons (Chapter 14) does the same when talking of gender relations, and Mary Shaw (Chapter 18) regarding food surveys.

Interpreted, analysed, reasoned, understood

Results are not always categorised in the ways that would be most useful to policy consideration. Mel Bartley, David Blane and George Davey Smith (Chapter 27) illustrate how problems of categorisation influence studies of health. Nicola Brimblecombe (Chapter 34) shows how crude decisions about the categorisation of teachers determine official reports on schools. Ann Singleton (Chapter 19) shows how the categorisation of different people influences European migration statistics, Ian Miles (Chapter 5) shows how the categorisation of research

statistics is used to disguise military funding. Peter Lee (Chapter 21) shows how different parts of Britain are labelled as deprived depending on the analysis used. Ray Thomas (Chapter 37) shows how choices of categories makes it appear that unemployment is either rising or falling. And Waqar Ahmad (Chapter 16) discusses the general problem of categorisation in relation to ethnicity.

Analysis of data always involves assumptions. Conservative assumptions are the norm because where there is doubt, it is often convenient to assume something acceptable to the status quo. Thus Ian Diamond (Chapter 2) refers to the fact that those missed from enquiries are often assumed to be the same as those who were not missed. Jamie Peck and Richard Bate (Chapters 39 and 42) refer to forecasts that assume that the future will be the same as the past, and Jay Ginn and Charlie Owen (Chapters 15 and 3) refer to the assumption that resources and consumption are spread evenly within families. Results are rarely tested for their robustness to alternative reasonable assumptions, but it is always reasonable to ask of research 'Would the conclusions have been different if the assumptions had been different?'

Statistics never reveal some precise truth simply waiting to be discovered, though this is how they are commonly perceived. The universal role of uncertainty is rarely recognised, yet it is a central part of any statistical study that relationships between variables are approximate and that true values are not known exactly. Harvey Goldstein (Chapter 32) illustrates the folly of assuming that pupils' school performances can be measured exactly. James Cornford (Chapter 44) shows how statistics about the Internet are riddled with uncertainty. Roy Davies (Chapter 38) shows how the labelling as sick of many unemployed men in the Welsh valleys adds uncertainty to economic indicators there. Ron Johnston, Charles Pattie and David Rossiter (Chapter 45) show how the margin of uncertainty in the British electorate and voter registration is sufficient to affect national election results. Theo Nichols (Chapter 30) illustrates how uncertainty over the meaning of industrial injury statistics allows for official inaction. Recognition of uncertainty would (certainly!) affect policy.

Disseminated, reported, presented, described, told, informed, communicated

Research reports often reflect the commissioners' view of the solution as well as of the problem and can for this reason be very restricted. Walid Abdul-Hamid (Chapter 23) shows how statistics have been used to make homeless people appear to have more of a mental health than a housing problem. Jane Elliott (Chapter 13) shows how statistical models are taken too literally. Ian Parker (Chapter 11) discusses how subjective studies are made objective, David Sibley (Chapter 12) argues that the use of statistics can damage understanding, while Pauline Davis (Chapter 33) shows how statistics about children's ability to read are seen all too readily as realities.

One cannot expect every report of research to include all the information that would allow each reader to judge the appropriateness of the design, the quality of the sample and collection, and whether different analyses might have given different conclusions. The demand from funders and media alike for concise and influential reports contradicts the tentative nature of most research results. It is

reasonable to expect all research reports to highlight the limitations of their results, and to be treated with a large dose of scepticism if they do not.

The dissemination of social statistics does not reach very far. Headlines dominate. While technological advance makes vast sets of data available to professional researchers, very little is accessible to individuals and community associations. Access to data increasingly requires resources and computing skills. A prerequisite to wider access is clearer dissemination, explanation and interpretation of datasets. Government could stimulate such clarity by making data from its censuses and surveys more cheaply and easily available, and by offering alternative commentaries on social trends. The debate over the creation of a national statistical service independent of government should have access and criticism high on its agenda (*see* p. xxvi).

Finally, the authors of this book have raised some pressing issues for debates about social statistics in Britain today. Should we have open government and more access to data (Michael Blakemore, Chapter 6)? Do we need a question on income in the next census (James Nazroo, Chapter 25) and should we be asking about religion (Joanna Southworth, Chapter 17)? How vital are the General Household Survey and other government surveys (Charlie Owen, Chapter 3)? How should issues of inequality be explored and addressed through statistics (Anthony Staines, Chapter 28)? This book also illustrates how social statistics can be represented graphically. Graham Upton's chapter (Chapter 46) illustrates some innovative new methods of displaying multi-dimensional data. The presentation of clear and informative graphics is essential to good statistical dissemination (Dorling and Fairbairn, 1997).

Conclusion

This book has shown how a series of critical views can open up the black box of social statistics. Politicians quote statistics relentlessly and rely on them for their work. While some may abuse statistical practice with narrow and mystifying jargon, the origins and meaning of social statistics are usually capable of straightforward interpretation. With a little thought and a little knowledge you can produce statistics which are fair and just, or discredit other statistics which are neither fair nor just. Social statistics is not an area of knowledge best left to the social statistician, numerical sociologist or epidemiologist. The numbers which are produced under its heading concern, and are concerned with, all of us (*see* Chapter 1). The errors that are often made in their construction are obvious to many of us and have implications for most of us.

Statistics can also be developed to uncover injustice and prevent bad decisions being made. They can be used to change society as much as society changes them.

The critiques in this book attempt to clear away some of the mystery around social statistics. Clarity about the concepts and theories that are used to frame research questions is as important as clarity in the analysis and presentation of statistics. Faced with a statistical claim, it is fair to ask:

- Is that true? How do I know it is true? Where did it come from? Who wanted me to know that? What are the alternative explanations? Could I have done better?

Faced with a social problem that requires action, it is fair to ask:

- Who would want change and who would not? Who needs to be convinced of the nature of the problem? What extra information is needed to clarify the problem? What information is needed to clarify possible solutions? How would it best be collected and analysed so as to leave no doubt?

Appendix: a list of Radical Statistics publications

Books and pamphlets

Irvine J, Miles I and Evans J (eds) (1979): *Demystifying Social Statistics*. London: Pluto.

Jones D, Smith P, Kinnersley P and Radical Statistics Health Group (1982): *Two Statistical Methods for Detecting Hazards at Work*. London: Radical Statistics.

Radical Statistics (1978): *RAW(P) Deals: A Critique of 'Sharing Resources for Health in England'*. London: Radical Statistics.

Radical Statistics (1978): *Social Indicators: For Individual Well-Being or Social Control? (The Case of OECD)*. London: Radical Statistics.

Radical Statistics Education Group (1982): *Reading between the Numbers: A Critical Guide to Educational Research*. London: BSSRS Publications.

Radical Statistics Education Group (1987): *Figuring Out Education Spending: Trends 1978–85 and Their Meaning*. London: Radical Statistics. 20pp.

Radical Statistics Health Group (1976): *Whose Priorities? A Critique of Health and Personal Social Services for England*. London: Radical Statistics.

Radical Statistics Health Group (1977): *In Defence of the NHS: An Attack on Fee for Service Payments in Medical Care*. London: Radical Statistics.

Radical Statistics Health Group (1981) *The Unofficial Guide to Official Health Statistics*. London: Radical Statistics.

Radical Statistics Health Group (1985): *Unsafe in Their Hands: Health Service Statistics for England*. London: Radical Statistics.

Radical Statistics Health Group (1987): *Facing the Figures: What Really Is Happening to the National Health Service?* London: Radical Statistics. 191pp.

Radical Statistics Health Group and Local Radio Workshop (1980): *A Better Start in Life? Why Perinatal Rates Vary in Different Parts of the Country*. London: Radical Statistics. [audio cassette]

Radical Statistics Nicaragua Statistics Fund Subgroup (1987): *Statistics in Nicaragua: A Report*. London: Radical Statistics. 22pp.

Radical Statistics Nuclear Disarmament Group (1982): *The Nuclear Numbers*

Game: Understanding the Statistics behind the Bombs. London: Radical Statistics. 95pp.

Runnymede Trust and Radical Statistics Race Group (1988): *Britain's Black Population*. London: Heinemann Educational Books. 160pp.

Broadsheets

Radical Statistics (1997): *More Money for Running Costs?* London: Radical Statistics.

Radical Statistics (1997): *More Money for Capital Projects?* London: Radical Statistics.

Radical Statistics Groups on Performance Indicators and Education (1992): *Choosing Schools by Numbers? The Statistics of School Performance*. London: Radical Statistics.

Radical Statistics Health Group (1992): *A Growing Health Service?* London: Radical Statistics.

Radical Statistics Health Group (1997): *More Patients Treated?* London: Radical Statistics.

Radical Statistics Health Group (1997): *More Nurses and Doctors?* London: Radical Statistics.

Journal articles: 'Watching the figures' pages in *Health Matters*

Radical Statistics Health Group (1991): Counting the cost of young lives. *Health Matters*, **6**, 11.

Radical Statistics Health Group (1991): Residential care: a private takeover? *Health Matters*, **7**, 17.

Radical Statistics Health Group (1991): Ignoring inequality. *Health Matters*, **8**, 7.

Radical Statistics Health Group (1992): Waiting lists: the long and the short of it. *Health Matters*, **10**, 12–13.

Radical Statistics Health Group (1992): A growing health service? *Health Matters*, **11**, 18–19.

Radical Statistics Health Group (1993): Fewer beds, but more patients? *Health Matters*, **14**, 18.

Radical Statistics Health Group (1995): Figuring out league tables. *Health Matters*, **20**, 5.

Radical Statistics Health Group (1996): Shorter waits for the clinic? *Health Matters*, **25**, 17.

Radical Statistics Health Group (1996): Nurses up, managers down? *Health Matters*, **27**, 7.

Radical Statistics Health Group (1997): Things can only get better? *Health Matters*, **31**, 6.

Other journal articles

Radical Statistics Health Group (1991): Missing: a strategy for the health of the nation. *British Medical Journal*, **303**, 299–302.

Radical Statistics Health Group (1991): Let them eat soap. *Health Service Journal*, **14**, 25–7.

Radical Statistics Health Group (1992): NHS reforms: the first six months – proof of progress or a statistical smokescreen? *British Medical Journal*, **304**, 705–9.

Radical Statistics Health Group (1995): NHS 'indicators of sucess': what do they tell us? *British Medical Journal*, **310**, 1045–50.

Report

Public Health Alliance and Radical Statistics Health Group (1992): *The Health of the Nation: Challenges for a New Government*. Birmingham: Public Health Alliance.

Journal

Radical Statistics, published approx. 3 times a year. Current issues are available from The Radical Statistics Group, care of 10 Ruskin Avenue, Heaton, Bradford, DB9 6EB.

Date	*No.*	*Date*	*No.*	*Date*	*No.*
January 1975	1	June 1982	24	Winter 1990	47
June 1975	2	October 1982	25	Spring 1991	48
August 1975	3	February 1983	26	Winter 1991	49
October 1975	4	May 1983	27	Spring 1992	50
Spring 1976	5	October 1983	28	Summer 1992	51
Summer 1976	6	February 1984	29	Autumn 1992	52
Autumn 1976	7	July 1984	30	Chinese NY '93	53
Winter 1976	8	Spring 1985	31	Summer 1993	54
February 1977	9	Summer 1985	32	Autumn 1993	55
May 1977	10	Autumn 1985	33	Spring 1994	56
October 1977	11	Winter 1985	34	Summer 1994	57
Spring 1978	12	Spring 1986	35	Autumn 1994	58
Summer 1978	13	October 1986	36	Winter 1994	59
Autumn 1978	14	February 1987	37	S/Summer 1995	60
Winter 1978	15	June 1987	38	Winter 1995	61
September 1979	16	September 1987	39	S/Summer 1996	62
Winter 1979	17	January 1988	40	A/Winter 1996	63
May 1980	18	May 1988	41	Spring 1997	64
October 1980	19	October 1988	42	Summer 1997	65
February 1981	20	September 1989	43	Autumn 1997	66
May 1981	21	Winter 1990	44	Spring 1998	67
October 1981	22	Spring 1990	45	Summer 1998	68
February 1982	23	Autumn 1990	46	Autumn 1998	69