The Qualities of Statistics as Facts about Society

Ray Thomas, Research Fellow in Official Statistics
Faculty of Social Sciences, The Open University, Milton Keynes, U.K.; e-mail: r.thomas@open.ac.uk

Martin Sewell, Department of Computer Science, University College London, London, U.K.; email: m.sewell@cs.ucl.ac.uk

This paper aims to apply science to the evaluation of statistics as facts about society. The conventional treatment of quality as having only one dimension militates against such an evaluation. Evaluation of statistics as facts about society requires consideration of a range of qualities including the categorizations used, the population covered, and the forms of publication available. The paper argues that the education and training of statisticians should be extended to cover the social construction of statistics including the use of denominators. This extension should contribute to the proper scientific use of published statistics - both by professional statisticians and other users.

Used without question?

The year to March 2007 witnessed the most intense period of debate about the organization and management of UK official statistics since Winston Churchill set up a central statistical office in 1941. That wartime office was expected to manage the collection of coherent statistics from government departments in such a way that “the figures so collected should form an agreed corpus which will be accepted and used without question.” A challenge that is no less relevant today. (Statistics Commission, 2006, p 5.; the quotation is from Ward, R. and Doggett, 1991.)

This extract from the final report of the Statistics Commission captures attitudes to statistics that were widespread in the UK in first decade of the new century. Government statisticians, other professional statisticians, members of government departments and the members of the government itself want to believe that official statistics were objective and could be treated as if they belonged to the world of the natural sciences. In this period there was a crescendo of what might be called political correctness in the field of official statistics. The Commission’s assertion that figures should be used without question indicates an atmosphere in which many questions about official statistics that might have been asked were not articulated. Was the Statistics Commission suggesting that economic statistics, that are dominantly official statistics, should be treated as unchallengeable indicators of well being? Was the Commission suggesting that official statistics give a picture of society that is more accurate than that given by surveys conducted by non-government agencies?

The prevailing attitude is to see official statistics as like Caesar’s wife – as above suspicion. Caesar’s wife Pompeia was divorced in order to avoid investigation (Plutarch, n.d.) That attitude, coupled with lack of investigation, is not beneficial to the proper use or the development of statistics. This paper disputes the idea that statistics as facts about society, such as official statistics, should be regarded as objective. The aim of the paper is not to cast suspicion, but to call for investigation of the qualities of official statistics.

The paper points out that phrases such as “the figures so collected” and “an agreed
corpus” contribute to a portrayal of official statistics as objective evidence. The paper points out that statistics as facts about human beings and society are artifacts. They are the creations of organizations that have human managers and human staff. The organizations have specific functions and responsibilities. It cannot be assumed that differently constituted organizations with different staff would produce the same or even similar statistics for the same area of human activity. The paper points out that examination of the production processes is necessary for proper scientific use of official statistics – and other kinds of statistics about human activities. This paper argues that recognition of statistics as social products implies that official statistics, rather than be accepted without question, should be under constant examination by users, by social scientists and by representatives of the public.

The dual meaning of statistics
The term statistics has two meanings - singular and plural. The term is used here in the plural sense. The plural sense accords with that of the original Charter of the Royal Statistical Society formulated in the 19th Century. The original Charter specified the objectives of the Society as: “to collect, arrange, digest and publish facts, illustrating the condition and prospects of society in its material, social, and moral relations; these facts for the most part arranged in tabular forms and in accordance with the principles of numerical method” (Royal Statistical Society, 2007). Thomas (2003) labeled this plural meaning as Pstatistics. The capital letter P can be considered as standing for plural or population. Pstatistics corresponds to what is sometimes called descriptive statistics.

The plural meaning has become well established by developments in official statistics. The Statistics and Registration Service Act of 2007 uses the word statistics with a plural sense and meaning. The 2007 Act established The Statistics Board to monitor the production and publication of official statistics (House of Lords, 2007). Everyday usage tends to avoid the use of the word statistics in its plural sense. Except in phrases like “official statistics” and “government statistics” the plural form is avoided. Words like facts, figures or data are used to denote the plural meaning of statistics. Or, perhaps more common, a specific noun such as population, incomes, employment, is used to refer to particular series.

Everyday usage has connotations of expertise, professionalism, statistical method, etc. Such connotations derive from developments in the activities and influence of the profession of statistics and Royal Statistical Society.

The RSS was granted a new Charter in 2006 very different from the Charter of the 20th Century and earlier. The new Charter specifies the objectives of the Society as:
the advancement of the science of statistics and its application, and the promotion of the use and awareness of statistics, in particular by: fostering and encouraging the growth, development and application of such science in all areas of activity which can benefit from it (Royal Statistical Society, 2007b). The use of the word “statistics” in the phrase “the science of statistics” betokens a singular meaning of the word.

Thomas (2003) labeled the singular meaning as Mstatistics. The capital letter M standing for method and mathematics. Mstatistics corresponds to what is commonly called inferential statistics. The singular meaning of statistics has largely dominated since the 19th Century and the establishment of the statistics profession.

The implications of these two meanings have not been widely explored. But, as Thomas (2003) pointed out, the dominance of Mstatistics has led to the neglect of various aspects of Pstatistics.

The importance of differences between the singular and plural meanings is not often
emphasised. But the difference becomes apparent in most matters related to official statistics. In Welsh the Statistics Board created by the 2007 Statistics Act is known as Y Bwrdd Ystadegau. The word ystadegau means statistics with its plural sense and meaning. The Welsh language has a different word ystadegaeth for statistics as a body of techniques. The Welsh language usage poses a question. Is the subject of statistics defined by the techniques of statistical method? Or is the subject matter defined by the body of knowledge created by the application of the techniques in the form of published statistics?

The limited acknowledgment of the differences between the singular and plural meanings of the word statistics provides the justification for and demonstrates the need for this paper. The pursuit of the science of statistics by the RSS and by professional statisticians has given emphasis to those aspects of statistics that relate to experiments and to the analysis of statistical outputs of various kinds. The statistical outputs come from specially conducted surveys and from a variety of administrative systems - such as medical records as well as governmental systems. The focus on the use of outputs has led to relative neglect of the procedures and processes involved in the creation of such outputs.

The neglect of examination of such processes and procedures means that the existing body of knowledge associated with the evaluation of statistics as facts about society is shallow. The paper aims to give a coherent account of a range of factors that should be taken into account in any serious examination of official statistics and other published statistics about society. The aim of this paper is to apply science to statistics as facts about society.

The professional statisticians’ neglect of qualities
A necessary stage in the application of science to published statistics is recognition of the many different qualities of published statistics. Government statisticians and the Royal Statistical Society avoid consideration of the qualities of official statistics by treating quality as if it has only one dimension. The 2007 Statistics Act expressed this in legislation by specifying that the function of the new Statistics Board is to promote and safeguard the ‘quality’ of official statistics.

Other clauses of the Statistics Act indicate how this focus on quality as if quality had only one dimension oversimplifies. The Act refers to accessibility, impartiality, accuracy, relevance and coherence with other official statistics. All of these characteristics can be considered as aspects of quality. The Act does not offer any explanation as to how these characteristics might be combined in order to assess overall quality.

The 2007 Act is consistent with earlier papers examining quality from an international perspective. Lievesley (2001), for example, listed twelve components of quality – validity, reliability, relevance to policy, potential for disaggregation, currency, punctuality, coherence across different sources, clarity and transparency, accessibility and affordability, comparability to internationally agreed standards, consistency over time and space, efficiency in the use of resources. Lievesley’s starting point was quality as defined in ISO 8402, “the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs” (2001, p.4). ISO 8402 is now superseded by ISO 9000. But it is not clear that a definition of quality that appears to be formulated with manufactured goods or the provision of private services in mind is appropriate for published statistics. Lievesley’s paper does not offer any explanation for limiting the definition of quality to the ‘totality’ of features and characteristics.

Eurostat adopts a broadly similar approach. But Eurostat does give emphasis to users (Who they are? How many? How
important is each one? What are their needs? etc.). It would hardly be possible to answer such questions without consideration of a range of qualities of statistics,

Apart from this acknowledgement by Eurostat of the importance of users, the emphasis given to the singular form of quality denies the existence of the wide range of qualities that may be important for users and for the public. The term quality in its singular form becomes a slogan or rallying-cry supporting the profession of statistics and government statistics in general rather than providing a basis for evaluation of statistics for any specific function, purpose or use. The value of statistics for any specific use depends upon the qualities of the statistics. The qualities depend in turn upon a variety of factors involved in the production of the statistics.

This paper picks out the factors affecting qualities under three headings - the categorizations adopted, the production procedures followed, and the statistical processes and the forms of presentation involved. The nature of the factors affecting the qualities of statistics largely depends upon the nature and functions of the organization, usually part of the government, that produce the statistics. The involvement of professional statisticians in the production process may be very limited.

The categorizations of statistics

The subject matter of a statistical series or a dataset is defined by the categories specified. Categorization is a fundamental stage in the production of statistics. According to the classic view, categories should be clearly defined, mutually exclusive and collectively exhaustive. But statistical categorizations typically involve factors that may conflict with the image given by the classical view.

The labels used to categorize statistics about human activities usually stem from societal or cultural sources. But the range of societal and cultural meanings is usually wide. Most commonly categorizations are made operational, that is applied to a dataset, in accordance with governmental requirements. Typically that process involves specifying boundaries that exclude many social and cultural meanings and have boundaries that in their detail appear to be fairly arbitrarily drawn.

The range of meaning covered by the word ‘work’ for example is almost unlimited. But the governmental interest in work is primarily as a source of taxation revenue. The usual statistical definition is limited to work for pay or profit. The lower boundary of the statistical definition is more generous – but arbitrary. The international definition counts all who work for one hour a week or more as in employment.

This example may be typical in giving a statistical label that is far from widely understood meanings. Typically it appears that categorizations are determined by a combination of governmental need and decisions taken by the statistics-producing agency often tempered by decisions taken at international conferences of statisticians and other government officials. It is unclear what body of knowledge or professional activity is relevant to such categorizations. Paul Starr (1987) used the phrase ‘cognitive commitment’ to describe the mostly uncritical use of the statistics produced in accordance with official categorizations. Starr suggested that official categorizations have a powerful influence beyond statistics:

Once the Census Bureau adopted the category Hispanic, American society became cognitively committed to it. In France, the use of the category “cadres” in official statistics helped crystallize the class identity of those in middle-level technical and managerial posts. The groupings for social insurance in Germany particularly the split between Arbeiter and Angestellten (wage workers and salaried employees, are said
to have helped shape class consciousness (Starr, 1987). Statisticians can generally be expected to say that the client (i.e. the body paying for the production of the statistics) should expect to take responsibility for the application of the categorizations. But in the case of official statistics the client includes the minister of the department, members of the government statistical service and civil servants within the department who would be the principal users of the statistics. There is no reason to assume that any of these parties has any special knowledge relevant to categorization. Nor is their reason to suppose that they are aware of what may be the consequences of their decisions as suggested by Starr.

Statisticians have no special expertise about the categorizations used for the production of statistics from surveys or from administrative sources. Nor does it appear that statisticians have special knowledge about making the labels operational - that is applying the labels to a data set. The justification for statistical categorizations largely belongs to statistics producing agencies. In the case of official statistics the statistics-producing agency is not clearly defined.

Production procedures involved
The coverage and many other qualities of published statistics depend upon the production procedures adopted. A variety of factors influence both the nature of the population covered and the coverage of that population.

In a few areas coverage is defined by the system itself. In the UK registered unemployment or the Count of Claimants, for example, is a count of the numbers receiving unemployment benefit. The figures are produced by local employment offices on the basis of the numbers claiming at that office. The figures are available for local areas within a month and the coverage is effectively 100%.

But the number of registered unemployed according to the Count of Claimants is not the only official measure of unemployment.

The UK also produces statistics for unemployment according to International Labour Office definitions. These figures come from a very different kind of source - the Labour Force Survey. The LFS is a large sample survey in which around 10,000 people aged 16 and over are interviewed each week. In order to produce estimates from the survey, the LFS sample data are scaled up (weighted) to ONS population estimates. The large sample size does not support the production of reliable statistics that approach the Count of Claimants in terms of accuracy or up to dateness. But the LFS/ILO figure is supported by UK statistical authorities in the UK and elsewhere because it aims for international comparability.

Estimates of the total population provide a quite different type of example - near the other extreme from registered unemployment statistics. A census of the population is the traditional solution. But the coverage of a census is always uncertain and may be becoming less certain. People have a wide range of motives for not wishing to be found by census enumerators and for not giving them all the information they seek. These motives are often connected with migration and economic status.

It is costly for governments to organize censuses. After the failure of the 2001 Census to provide the best basis for estimating the distribution of population it seems likely that the idea of a 100% census will be abandoned. Increasingly reliance will be placed on estimates based on administrative records and claims made by local authorities.

The coverage of statistics relating to incomes cannot be expected to be high or exact even where they are based on administrative sources. People have ways of avoiding tax - and ways of providing information for
administrative systems that support taxes on income.

A growing proportion of statistics at least in Britain are from household surveys rather than administrative sources. Where the subject of the statistics relates to the workplace, such as earnings, household-based statistics may lead to further losses of coverage or accuracy.

The major problems with survey statistics relate to response rates. In Britain at least survey response rates have been slowly declining for some years. Statisticians have not provided explanations for this decline beyond saying such things as that it happens everywhere and that it is associated with a decline of respect for authority and government. The response rate is usually based on the estimated numbers who declined to fill in a form or answer interviewers’ questions.

A parallel problem is what might be called the ‘avoidance rate’. Just as illegal immigrants, for example, take steps to avoid being counted in a census so can individuals and households take steps to avoid being counted in a survey. National statistical offices do not appear to find it easy to make estimates of the avoidance rate.

**Statistical processes and presentation**

It seems likely that most users of published statistics assume that professional statisticians have involvement in the forms of presentation of published statistics. But there is little evidence to support that assumption. It appears that professional statisticians see presentation as a matter for the statistics-producing agency. It appears that professional statisticians see presentation as involving only very basic statistics.

The education and training of the statistician covers only the basic elements relevant to descriptive statistics and the presentation of statistics. Any further knowledge among statisticians is assumed rather than attested. The RSS syllabus for its entry-level Ordinary Certificate includes a section covering a number of descriptive measures. “Arithmetic mean, median, mode, percentiles, range, inter-quartile range, variance, coefficient of variation; their uses and limitations as measures; their calculation from frequency tables and raw data; graphical methods of estimation”. But there is no other mention of descriptive statistics in the RSS syllabus for any of its examinations up to and including to the Graduate Diploma level that is seen as qualifying a professional statistician. The RSS examination system assumes, in other words, that knowledge of this range of measures at this entry-level Certificate is adequate to meet the needs of the professional statistician.

The gap between the professional statisticians’ perspective and the needs of statistics publications is starkly illustrated by the omission from the RSS syllabus of any mention of the term denominator. A high proportion of official statistics are expressed in terms of rates – for the incidence or prevalence of a condition in the population. The use of denominators supports comparison over time, between areas or between age groups.

The choice of denominator is often far from straightforward. There is for example often a choice between using the total population of geographical area/age/sex group or a more precisely defined population-at-risk. Quite different pictures of the incidence and prevalence of a condition can be given according to choice made between the use of whole population or population-at-risk. The particular criteria used to define population-at-risk can also be a major influence.

In some areas such as the labour market the choice of denominator belongs to tradition. The denominator usually used for employment and economic inactivity, for example, is the population of working age. That choice of denominator seems logical. But why then is the denominator usually used for unemployment the total of employment and unemployment? What are the consequences of this choice?
This paper argues that users should not assume that indications of the involvement of professional statisticians in published statistics guarantee that the form of publication is optimum. The paper suggests that particularly in the use of denominators in labour market statistics the denominator used in published statistics is misleading or difficult to justify.

**Why statisticians offer little support**

Why have statisticians neglected these areas that are relevant to the evaluation of published statistics? The limited contribution by statisticians to the application of science is associated with professionalization. In the 19th Century the Royal Statistical Society was a kind of gentleman’s’ club. The members of the middle classes established the RSS as a learned society devoted to investigation of poverty and the horrors of rapid urbanization that characterized Britain in the 19th Century. People became members of the RSS by nomination and were called Fellows of the Society.

That situation began to change in the 20th Century with developments such as a growth in official statistics, the introduction of sample surveys and significant advances in methods of statistical analysis. It was no longer sufficient to be a gentleman educated in the humanities to be able to use statistics to develop a good understanding of how society was changing. A focus on inference was replacing purely descriptive uses of statistics. At least a modicum of numeracy was necessary.

The Institute of Statisticians (IOS), an organization expressly dedicated to the interests of professional statisticians, was formed in 1948. In 1993 there was a merger between the IOS and the RSS. The main victor in this merger was not the professional statistician or the RSS but mathematics. Mathematics and statistics have become inseparable. But that unification has been associated with growing distance between professional statisticians and the world of published statistics.

The Royal Statistical Society and the Statistics profession are generous in their efforts to introduce members of other professions and members of the public to the possibilities offered by statistical analysis. A wide variety of courses are offered. But typically these courses focus on a particular statistical technique. The courses do not necessarily assume that attendees have any background in statistical method but they do typically assume that attendees are going to apply method to sets of statistics that they are familiar with.

The statisticians giving such courses typically assume that attendees are already familiar with the categorizations adopted, the production procedures followed, the statistical processes and the forms of presentation of the statistics to which they will be applying statistical method. In other words, statistical method is constantly being defined in terms that exclude the aims and content of this paper.

**What should be done?**

At frequent intervals the Economic and Social Research Council in Britain organizes courses and meetings on statistical training for the social scientist. It is hoped that this paper indicates the need for some training in the other direction. The paper has highlighted a number of areas associated with official statistics where there is a need for evaluation that is not provided by the education or training of the professional statistician. The recognition of statistics as a social product with a range of different qualities is necessary for making that evaluation.

Journals published by the Royal Statistical Society rarely include articles that use the term social product, or social construction of statistics, or any closely related terms. Professional statisticians habitually define the qualities of statistics as belonging to the organization producing the statistics. This
This paper aims by discussing the qualities of statistics to indicate the importance of treating statistics as a social product. By focusing on the qualities of statistics, rather than the quality of statistics, this paper aims to remedy a gap in the body of knowledge associated with the use of official statistics and other published statistics. It is hoped that this paper, by remedying this gap, will encourage the proper scientific use of published statistics – both by professional statisticians and other users.

References

http://www.opsi.gov.uk/acts/acts2007/ukpga_20070018_en_1


Ray Thomas is an Honorary Research Fellow in the Faculty of Social Sciences of the Open University in Milton Keynes in England. He is a Fellow of the Royal Statistical Society and has conducted research on the functions of published statistics and in the areas of new towns, urban development and international trade.

Martin Sewell, BSc (maths), MSc (computer science), is currently completing a PhD on the application of intelligent systems to financial time series analysis in the Department of Computer Science at University College London. He has previously worked on financial statistics in the City of London and at Cranfield University.

This article is published under the Creative Commons license: Attribution-Noncommercial-Share Alike 3.0 Puerto Rico.