Prioritizing the User Perspective on Statistical Quality

Richard Alldritt*
Consultant, London, United Kingdom – Alldritt.consulting@gmail.com

Abstract

It is a much quoted tenet of management theory that you can only control the things you can measure. But it is also often said that the issues that are most important, in the long term, are the ones that cannot be measured in advance. Statistical offices worldwide, but notably in Europe, are under pressure to establish formal systems to document and monitor quality. This tends to be translated into practice in terms of 'process quality'; and that sometimes means little more than conformity to a preordained process specification. That in turn promotes a culture in which effective process control is seen as the goal. Process control is not a bad thing in itself but it risks creating an inward focus and an impediment to product innovation. And that constitutes a major risk for any modern public service.

Ultimately, the quality of any statistical product or service must be contextualised by whether it serves a valid purpose – whether it does any good. The 'shelf life' of many official statistics is long, potentially 100 years or more. During that period the statistics will be used countless times in diverse ways, not all of them predictable and many of uncertain benefit to society. The utility of statistical outputs is the thing that cannot be measured in advance, and it really is the most important issue in the long term. Not only is this the right approach to quality, it is fundamental to the business case for any statistical function. Statistical offices around the world are struggling in the competition for public funding. They must – in their own interests - find ways to build the actual and potential use of their outputs into their quality frameworks and, by this means, focus attention on the social and political value of their products. To do this effectively requires a standard approach by which to identify distinct and valuable uses of statistics and indicate the value to society associated with each use.

This will never be an exact science, but a less than perfect approach may still serve to push organisational development in the right direction. There are already some instances where the need for a formal business case obliges a statistical office to make explicit assumptions about the future use and value of proposed new statistics - and this is powerful where it is done. But extending that obligation to all official statistics, whether existing or planned, should be seen as a prerequisite for effective quality monitoring.

Keywords: quality management; utility; resources.

1. Introduction

This paper is not written from the perspective of any national or international statistical office. Nor is it a plea on behalf of users of statistics. It is instead a reflection on why the current emphasis on formal systems of quality management is failing to drive real quality; and failing to inspire confidence on the part of the governments that fund statistical offices. And why funding for work in the field of official statistics will continue to be under pressure until this is resolved.

2. Quality confusion

The great 20th Century statistician, W. Edwards Deming is quoted as saying *"The most important things cannot be measured."* And in another version *"The most important figures that one needs for management are unknown or unknowable, but successful management must nevertheless take account of them."*

When we consider the quality of official statistics, the 'most important things' are not immediately obvious. Words such as accuracy, robustness, comprehensiveness, relevance, timeliness are used. But none of these concepts is as fundamental as utility. The instinctive warmth that statisticians feel for things that can be measured, coupled with the perceived need for objective indicators, has sucked the debate over quality and quality management in the direction of process control – if a statistical process follows the approved description of the process, then quality is good. Unfortunately, that statement is profoundly not true, as Deming would have been the first to point out. Processes should certainly be monitored but conformity to them is not of itself quality.

To establish what the 'most important things' for official statistics are, it is helpful to start with why they are produced at all. To many European statistical offices the answer to that question seems to be mainly to meet the requirements of European statistical regulations, and with that one leap they seek to absolve themselves of any real responsibility for answering it.

But given that official statistics are produced very largely at public expense, there can only be one fundamental answer to why we produce them – and that is that they have the potential to enhance national wellbeing, or make life better for the citizen. Or more universally, that they contribute to making the world a better place. If they did not do this, then our governments should not fund statistical work but rather spend the money on hospital beds, or public transport, or whatever seems a greater social priority. Indeed, to really justify spending public money on producing statistics, we must argue not just that they make the world a better place but that they make it better by a greater degree than a corresponding increase in any other field of public expenditure. At times of economic austerity, that is a tough case to argue with politicians whose priorities usually lean towards ways of spending money that resonate more naturally with voters. Spending more money on statistics has never been a sure-fire vote winner.

And unfortunately for official statisticians worldwide, in times of economic difficulty governments sometimes see statistical data as a source of bad news, and are even less inclined to see them as a priority. Faced with these challenges, statistical offices must make a convincing case that official statistics really do make the world a better place. And where they fail to do that currently, it is at least in part because they have been led down the wrong path by employing concepts of quality and performance that ignore what is really important.

National statistical offices do make some effort to get the message right. Most of their mission statements recognise that statistics are produced to inform decisions and actions, although they often seem to imply that only government decisions and actions are of any importance - a stance that undervalues statistics substantially. Some do better, recognising that spending public money on official statistics is justified by the aggregate of all uses of them that are of public value. This principle is captured in the words of the first of the UN Fundamental Principles of Official Statistics:

"Official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information."

As a general statement, that is hard to beat. The 'test of practical utility' is the right test. The only real justification for producing official statistics is that they pass that test. We can assume here that 'practical utility' includes all uses that serve the public good and excludes ones that are contrary to the public good. The use of official statistics by criminals looking to target their next fraud does not count as added 'utility'. What the principle says is that statistics that pass the test 'are to be produced'; it is a profound obligation, not an option to be determined by political whim.

National and international quality frameworks often fail to incorporate this concept adequately. However, a good model is offered in the 'Quality Framework for Statistical Activities' of the Organisation for Economic Cooperation and Development (OECD):

"Quality is defined as 'fitness for use' in terms of user needs. This definition is broader than has been customarily used in the past when quality was equated with accuracy. It is now generally recognised that there are other important dimensions. ... The quality characteristics of most importance depend on user perspectives, needs and priorities, which vary across groups of users."

The European Statistical System offers a rather more laboured version:

> "In line with the last five European Statistics Code of Practice Principles, output quality in the ESS is assessed in terms of the following quality criteria: *Relevance: Accuracy and Reliability: Timeliness and Punctuality: Coherence and Comparability: Accessibility and Clarity.*

Here the concept of meeting the test of practical utility has metamorphosed in to *Relevance* and is presented as just one subsidiary element of quality, on a par with timeliness. *Relevance* is in turn described in the European Statistics Code of Practice, and in the Eurostat Quality Assurance Framework, in terms of three indicators:

- Processes are in place to consult users, monitor the relevance and utility of existing statistics in meeting their needs, and consider their emerging needs and priorities.
- Priority needs are being met and reflected in the work programme.
- User satisfaction is monitored on a regular basis and is systematically followed up.

Buried in this formulation is the requirement to monitor the 'utility of existing statistics' in meeting user needs. But the implication of the surrounding text is that one only needs to ask the users; that they will tell us whether the statistics are of value and that is all we need to know. It isn't.

But even this rather limited commitment to utility, fails to find its way in to the European Statistical System Quality and Performance Indicators. Here are the main headings:

R1. Data completeness - rate
A1. Sampling error - indicators
A2. Over-coverage - rate
A3. Common units - proportion
A4. Unit non-response - rate
A5. Item non-response - rate
A6. Data revision - average size
A7. Imputation - rate
TP1. Time lag - first results
TP2. Time lag - final results
TP3. Punctuality - delivery and publication
CC1. Asymmetry for mirror flows statistics - coefficient
CC2. Length of comparable time series
AC1. Data tables – consultations
AC2. Metadata - consultations
AC3. Metadata completeness - rate

This sadly technical list of indicators is at best a set of process control indicators of rather limited application. It has little to do with the concept of quality as defined, for example, in the OECD Quality Framework.

We need to take a long hard look at some of the assumptions implicit in the text of the various statements. 'Relevance' is how things look from the *producer* perspective, not the user perspective. It is a property that a particular statistical series may have – that it seems to those who produce it to be relevant to the decisions that other people make. But that is not of itself enough to justify the commitment of public money. Perhaps the user would find some other statistics, not currently produced, even more relevant or more useful. Perhaps the user, whilst enthusiastic and vocal, does not really deliver much public value by their use of the statistics.

A statistical series may be relevant but the uses and users to which it is relevant may not actually be of much importance in terms of current or future public value. And this scenario is common. Many of the most vocal and demanding users of official statistics do not actually do very much of lasting value with them. European statisticians may wonder if the European Commission itself comes in to that category – vocal and demanding but not necessarily making much use of the statistical data that it requires Member States to produce. It is up to the European Commission and the other EU institutions to demonstrate otherwise.

Statistical offices need to understand the actual and potential use than can be made of their statistical products, not just whether they have 'relevance' to those uses. The emphasis on relevance is a poor proxy for the concept of utility, and potentially it fails to meet the main political imperative which is to prove the value of official statistics to those who control the supply of funds for statistical activities.

The argument that there is a statutory requirement to produce many of the statistics produced by EU countries is not of itself a good enough argument. Without the support of a clear rationale in terms of utility, the fact that there is a legal requirement simply encourages politicians to see statistics as a bureaucratic overhead – something to be done as cheaply as possible to no higher standard than absolutely necessary. That is a recipe for statistics that actually have very little utility and a statistical system that is regarded by the state as little more than a burden.

3. Evidence of utility

Statisticians are not the sort of people to make bold claims about the importance or value of their work. They want to see the evidence and have it available to support anything they do say. It is rightly expected of the profession to demonstrate a certain caution in reaching conclusions, to brigade and examine the objective evidence before making public statements.

But if statisticians, and statistical offices, are to fight their corner effectively, both in promoting their products to potential users and in making the case for public funding to support their work, they need to get the maximum impact from whatever evidence they have about the use of their products. The first step in this is get hold of as much evidence as possible. And that requires the right questions to be asked of users of official statistics. Too often at present, user satisfaction surveys seem to ask 'have you heard of the statistics office? Do you visit the website? Do you find it useful? These questions are a waste of time – they are the wrong questions to the wrong people.

The questions that must be asked are more like this: 'what decisions do you take?' 'What statistical information do you need to inform those decisions?' And 'what use do you make, or could you make, of the statistics we can provide you with?'

The people, or institutions, of whom such questions should be asked are mostly not the people who download tables from the website. Those individuals may have little idea what use will ultimately be made of the statistics they find. It is necessary instead to hunt out the, usually, senior figures who actually understand the decision making processes of an organisation and engage them directly. This is more work but is likely to lead to much stronger conclusions. And the aim must be to find the persuasive examples of socially or economically beneficial use, whether that be among government agencies, public bodies, businesses, not-for-profit organisations or the general public. Of course this will not reveal a comprehensive picture of use; nor will it allow the creation of a quantified indicator or index that can be tracked over time. But it will serve to fill the void left by the absence of those things currently. The evidence compiled will be so much stronger than an index of website hits, which is largely meaningless.

In seeking the persuasive examples of use, statisticians need also to speculate about the future, and about future uses of their products. It may not come naturally to their cautious, evidence-based, mind-set but they should nonetheless give it their best shot. Some statistics have a very long shelf-life: they will still be useful and relevant to decisions that are going to be made in 20 or even 100 years' time. Examples include statistics on public health, the environment, agriculture and so on. So statistical

offices must speculate about what those long distance future decisions might be. And then claim credit for informing those future decisions. Part of the quality of official statistics is their potential to still have utility, to still be relevant, long into the future. Claiming credit for this is not difficult or particularly controversial, but is hardly ever done in an open and confident way.

In investigating use, the nature of the user is of secondary importance to the nature of the *use*. At present there is too much emphasis on classification, or stratification, of *users*. That is a classic statistical approach but it does not lead to great insights. The range of users is of course relevant to the way the statistical product is packaged and presented. But we do not need to know much about who exactly the users are to understand the value associated with the use. Statistics on public health are potentially used by a vast range of organisations and individuals to inform many kinds of decisions that have some public value. We can get lost in trying to identify groups of users. That is not as important as deciding what value the statistics have in relation to the decisions those users might make. It is not the 'who' of decision-making that needs to be categorised, it is the 'what'.

4. Implications for quality frameworks

If, at the fundamental level, quality is to be seen as fitness for use (or 'utility'), then the steps required to establish the fitness of statistical outputs should be part of the quality management framework. The first of these steps is to identify the different uses made of the statistics and to illustrate those uses with specific examples. It isn't a matter of counting them, or of trying to count the users associated with each use. What is needed is to set out the ways in which each 'use' offers some benefit, or value, to society or the economy.

There are already some circumstances where the need for a formal business case obliges a statistical office to make explicit assumptions about the future use and value of proposed new statistics - and this is powerful where it is done. Population censuses are an example where, because of their huge and one-off cost, a detailed business case is often required. As part of this, statistical offices may have to go in to detail in formal documents about the assumed use and value of the census products. If they can do it for census data, they can do it for any statistics. Incorporating this obligation in to formal quality management systems in respect of all official statistics, whether existing or planned, should be seen as a prerequisite for effective quality monitoring.

To do that consistently may require a framework of its own – a standard approach to presenting the different uses of statistics and a standard approach to determining (but not quantifying) the public value of each use. This is where international discussion should be focusing. We can identify that official crime statistics are used, among other things, to inform decisions with the aim of preventing crime. We can make the observation that this use is widespread and give examples to illustrate the use. But what value shall we associate with that use? We can identify that consumer price indices are used, among other things, to inform decisions about uprating social benefits and pensions; and give examples of how this is done. But what value should the statistician claim in respect of this use?

The answer is not to search for a deeply scientific method; not to look for some spurious quantitative index of public value to score each use. Instead statistical offices need to be bolder and play by the same rules as others seeking public funding. Just as auditors evaluate risks as low, medium or high on the basis of very little actual evidence, perhaps statisticians should develop an agreed (if arbitrary) set of labels for the public value of a use of statistics. This might have a sense of scale as in the case of describing a risk as 'high'. Or it might be purely qualitative: the use of crime statistics in relation to crime prevention might be labelled as 'valuable to all citizens'; the use of the consumer price index to inform changes in benefits might be labelled as 'fundamental to economic policy and equity'. Statistical offices should not shy away from such judgements. No-one is better placed than they are to make them. And if such assessments should be challenged, then that at least initiates a debate about the value of statistics. That is a debate that, in the end, statistical offices cannot lose.

The identification and description of the uses –and related value - of official statistics may be all that is needed to make the case at a political level; and will do so better than any attempt at quantification. There is no exact science here but an internationally standardised approach may nonetheless stop quality management becoming purely about process; and may serve to focus the minds of both statistical offices and those who fund them on what really matters.