# Distilling the Statistical Sea

*International Association of Official Statistics conference, Kiev,
12-14 September 2012*

## Richard Alldritt and Richard Laux, UK Statistics Authority

**Abstract**
This paper argues that statisticians need to do much more to cater for the diversity of user needs, and in particular to communicate directly with decision makers. It proposes a series of relatively simple steps that would help to ensure that clear messages are distilled from the sea of data, metadata and other detail that characterise most statistical releases, while retaining the impartiality and objectivity for which they are rightly valued.

Official statistics make the world a better place; that is why we produce them. This fundamental belief deserves to be stated more often and with more confidence. The reason it is not often said is perhaps because whilst some statisticians regard it as self-evident, others are not quite sure that they can demonstrate its truth. Whatever inhibitions there may be about making the claim, the job of official statisticians is self-evidently to *make* it true and to guide others in maximising the social and economic benefit that flow from the use of official statistics.

Statistics are only worth producing at public expense to the extent that they will be used, either now or in the future, in ways that do some good. The first of the UN Fundamental Principles of Official Statistics talks about 'statistics that meet the test of practical utility'. That is an elegant way of saying 'statistics that do some good'. The quantity of 'good' done by the use of official statistics is not itself susceptible to much in the way of systematic analysis. One set of statistics may support the development of a new taxation regime; another may support the electorate in voting in a national Parliamentary election. It would be hard to say whether one 'good' is of more importance than the other. Statisticians tend to worry about things that cannot be quantified but we must just accept that sometimes a broad subjective judgement that a set of statistics supports socially and economically important decisions is enough on which to plan our work.

The statistician's instinctive discomfort with such intractable concepts may explain their wariness when it comes to making public statements about the use that is made of statistics or the benefit to society that flows from that use. It is notable that, at least in the UK, the publication of official statistics is still only rarely accompanied by an indication of the uses that are expected to be made of the statistics, or the relevance and utility of the statistics in relation to those uses. This is despite the fact that including such advice has been a requirement of the UK Code of Practice for Official Statistics since the beginning of 2009. And the need to consider these questions of utility has been implicit in the UN Fundamental Principles for much longer.

The UK Statistics Authority has legal powers to enforce compliance with the Code of Practice and takes the view that only by exploring questions of use, relevance and utility can the quality of the statistics be properly judged, the statistical service be shown to be of real benefit to society, and the published products be steered towards maximising that benefit. The Statistics Authority is just one of many 'statistical council' bodies around the world, but we hope that others would share that basic tenet.

To make the case for continued, or increased, public funding for statistical work we must not only demonstrate convincingly that official statistics do good; we must argue that producing those statistics is a better way to spend public money than on other public services. If we do not do that, we should not be surprised if democratically elected governments think that perhaps the money for statistics might be better spent on some extra hospital beds or policemen.

Of course, there is no definitive analysis that will prove the case one way or the other. But that does not diminish the importance of taking the argument as far as we can. Government statisticians must seek to understand the use that is, and can be, made of official statistics and the nature of the associated public good. Fortunately, a comprehensive analysis need not be the goal. Often all that is needed is some convincing illustrations which show how the statistics make the world a better place. The more such illustrations we can produce, the better.

And it is a gratifying fact that the same steps that help to make the case for public funding also help the statistical service to reach out to users of the statistics and increase the beneficial impact of the work. To maximise the public good, official statisticians must promote and facilitate use; and doing that means first investigating the current use of statistics and taking at least a broad view on how much good follows from that use.

Thus a competent statistical office needs to have the capacity to investigate and document uses, form a defensible view on the public benefit that flows from those uses, promote awareness of the statistics and their suitability for different uses, and support the user in making decisions and taking actions. This is a distinct set of skills, different from those required for conducting surveys or managing data. But perhaps even more than a different set of skills, it implies a particular culture or attitude. It is a culture that seeks actively to pick up clues as to who is using statistics and how they are being used; carefully documents and shares that knowledge; talks to users every day; uses the information gained to change and develop its outputs. It is a service delivery culture where the current and future needs of the user drive and motivate the organisation.

In terms of the skills and knowledge required, one key one is a thorough understanding of the subject to which the statistics relate. If a statistician is responsible for statistics on health and health services, he or she must understand not just the processes by which the relevant statistics are produced but also the institutions and concepts to which they relate and the conventions and cultures of those to whom the statistics refer. As just one example of this, self-reported ill-health can be influenced by the availability of social security benefits for people with long term illness or disability. Not understanding that could result in too much confidence being placed in changes over time.

A second important skill is the ability to communicate directly with journalists and with the sort of individuals and institutions whose decisions will be influenced by the statistics. Again, this involves understanding their perspectives and their agendas. Empathy with the user of statistics may be one of the most valuable skills in a statistics office.

A third essential area is the writing and editing skills to give statistical advice an authoritative voice. It is not enough that the advice, if understood, is technically correct. A confusing description of the trend over time, or geographical patterns will not support the user in finding the relevant message in the statistics.

The need for more user-facing skills and culture has increased as the flood of statistical data from administrative sources has increased. 'Administrative sources' here include all records

kept by public and other bodies where these can be accessed for statistical purposes. Their volume and quality have increased dramatically over the last 30 years and so has the flow of aggregate statistical data derived from them. We see users confronted by an incoming tide of statistical material, when what they want – in this metaphor - is a reliable flow of clean drinking water. Statisticians tend to show the user the sea; not distil it down into something more useful. The job of official statisticians now and in the future will be to turn the sea of material into something useful by understanding the needs of users and the relevance of the available statistical material to the user; and guiding the user towards the messages in the statistics.

It is important to recognise here that statistics do not normally guide decisions and actions directly. Few decisions are directly dependent simply on the latest figure in a series. Mostly the user needs to extract a message from the figures first. The user of unemployment statistics might look at the trend over the last two years and say 'our programme to tackle youth unemployment seems to be working but we need to do more of it'. That is the message that drives decisions and actions. The statistician should not try to steer the actions of the user but must make it as easy as possible for the users to find the message for themselves.

It is also important to understand what the word 'user' really means. Often the big decisions and actions (that lead to the most 'good') are taken not by individual people but by organisations and institutions. The big users will often be committees, boards and teams. However, some decisions *are* taken by individual members of the public – where to live, what job to do, where to send children to school, who to vote for in the next election – and their use of official statistics is a fundamental part of the democratic process. Their understanding, confidence in, and use of official figures will often be shaped by journalists and the news media. Arguably the news media will also have a big influence on the people who make up the committees and boards of the big users too. Thus we have a complex user structure with the news media playing a central and influential role in drawing attention to statistics; commenting on their worth; and placing a first interpretation on new data. Official statisticians must not only do all they can to make sure that journalists understand the statistics; they must seek to reach beyond the news media to speak directly to the institutions and individuals who can ultimately do some good through their decisions and actions.

The essential steps in guiding beneficial use are these:

- Investigate, understand, document and regularly report the types of use made of the statistics. It is not vital to get a comprehensive description. It is more a matter of illustrating the role the particular statistics play.
- Explain the strengths and weaknesses of the statistics when they are published in relation to the broad types of uses. This serves both to encourage use and warn the over-enthusiastic user not to place too much reliance on statistics that have weaknesses. The goal is not to maximise use but to maximise beneficial use. Sometimes the statistician must discourage use by warning the user against drawing invalid messages.

To keep the second step above manageable, it is sufficient to address the types of use made of statistics in generic terms. For example, crime statistics can be used for several distinct purposes: managing police resources; guiding crime prevention policies and decisions; as a geographical deprivation indicator; as a source of knowledge about personal risk. These and perhaps a few other uses can be identified. They are socially very valuable but the important point to note is that there are not that many of them. And it does not matter greatly *who* is using

the statistics for the assessment of personal risk.  The statistical advice is going to be much the same to all users in that group.

The statistical sea is rightly regarded with some concern by the user of statistics.  It is deep and increasingly polluted with poor quality data.  In the UK, the Code of Practice for Official Statistics guides official statisticians in what to do to help the user but the most fundamental requirement is that statisticians and statistical offices must first *want* to make the world a better place; and then take on the task of understanding and explaining how official statistics can achieve that.