

## **Dimensions of Statistical Quality**

### **A discussion note about the quality initiatives of some international organisations**

*Willem de Vries, UNSD<sup>1</sup>*

***Summary :** Eurostat, IMF and OECD have developed or are currently developing comprehensive quality frameworks for statistics. Other international agencies, such as UNSD and FAO, have taken quality initiatives of a more limited scope. There are similarities, but also differences in the various approaches, partly due to differences in objectives. This discussion note raises some harmonization issues.*

#### **1. Introduction**

International organizations have long been active in the promotion of quality management of statistical data by developing handbooks and guidelines. The examples are numerous: over the years UNSD has issued a wide range of manuals and technical notes on e.g. census organization and management, international trade statistics, national accounting (jointly with others), environment statistics and environmental accounting; the IMF has done the same for Balance of Payment Statistics and Government Finance Statistics, etc. All these initiatives have in common that they aim at the improvement of official statistics by offering (technical, methodological, organizational) frameworks for certain areas of statistics.

More recent is the attention for quality issues in a general, managerial sense: the promotion of statistical quality by means of comprehensive quality management systems, quality programs and quality frameworks. This form of quality thinking has been prominently on the agenda over the last five years or so. Over the past decade, many national statistical offices have embarked on explicit quality programs. Such programs may have widely different scopes and objectives. The broadest approaches start with formulating mission statements, developing multi-year corporate development plans or strategic plans etc<sup>2</sup>. Narrower approaches focus more specifically on data processing procedures or on assessment of user satisfaction with specific statistical products. For an overview and a wealth of country-specific details, the Good Practices website of the United Nations Statistics Division is a useful source<sup>3</sup>.

---

<sup>1</sup> This discussion note represents the personal views of the author and does not necessarily reflect UN views or policies.

<sup>2</sup> For example: Eurostat Corporate Plan 2001; Statistics Denmark Strategy 2005; Statistics Netherlands: Statistics that count, 2001.

<sup>3</sup> <http://unstats.un.org/unsd/goodprac/default.asp>

In addition to national initiatives, several *international* organizations have taken quality initiatives as well<sup>4</sup>. While there are strong similarities between these initiatives, there are also differences. To a large extent these differences have to do with the purpose and scope of the initiatives. Moreover, the terminology that is used to describe various aspects of quality is to a large extent identical, but there are also several –sometimes subtle- differences in the definition and interpretation of concepts. This discussion note gives a brief overview of the comprehensive approaches of Eurostat, IMF and OECD<sup>5</sup>. In addition, it discusses the more narrow approach of UNSD. The question is raised whether or not there is a case for harmonization of these efforts.

## 2. *Quality elements*

There is general agreement that, ultimately, quality has to do with user needs and satisfaction. The International Organization for Standardization (ISO 8402) defines quality as follows: “Quality is the totality of features or characteristics of a product or service that bear on its ability to satisfy stated or implied needs of customers”.

In statistics, quality used to be primarily associated with accuracy. It is now recognised that there are other important dimensions. Even if data are accurate, they do not have sufficient quality if they are produced too late to be useful, or cannot be easily accessed, or conflict with other credible data. Therefore, quality is increasingly approached as a multi-dimensional concept.

Statistics Canada, one of the first national statistical offices to adopt a comprehensive quality program<sup>6</sup>, distinguishes the following characteristics of quality. (Most of these are found – in some form or other- in other national frameworks as well, and also in the international initiatives that are the primary topic of this discussion note)

The *relevance* of data or of statistical information is a qualitative assessment of the value contributed by these data. Value is characterised by the degree to which the data or information serve to address the purposes for which they are produced and sought by users. Value is further characterized by the merit of those purposes, in terms of the mandate of the agency, legislated requirements and the opportunity cost to produce the data or information.

*Accuracy* of data or statistical information is the degree to which those data correctly estimate or describe the quantities or characteristics that the statistical activity was designed to measure. Accuracy has many attributes, and in practical terms there is no single aggregate or overall measure of it. Of necessity these attributes are typically measured or described in terms of the error, or the potential significance of error,

---

<sup>4</sup> In addition, several international quality conferences have been held over the last couple of years. See e.g. <http://www.nso.go.kr/eindex.html>, [www.statcan.ca/english/conferences/symposium2001](http://www.statcan.ca/english/conferences/symposium2001), [www.q2001.scb.se](http://www.q2001.scb.se).

<sup>5</sup> For a more detailed comparison, a paper by Alberto Signora and Michael Colledge of the OECD is worth reading (see References). Parts of this paper draw on the Signora/Colledge paper.

<sup>6</sup> Statistics Canada (1998) *Quality Guidelines*, third edition, October 1998, available at [www.statcan.ca](http://www.statcan.ca).

introduced through individual major sources of error - e.g., coverage, sampling, non response, response, processing and dissemination.

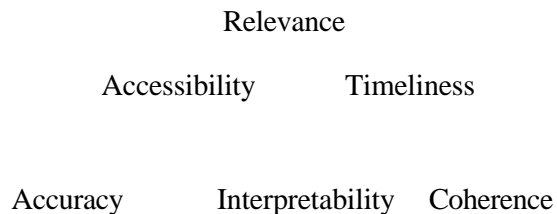
*Timeliness* of information reflects the length of time between its availability and the event or phenomenon it describes, but considered in the context of the time period that permits the information to be of value and still acted upon. It is typically involved in a trade-off with accuracy.

*Accessibility* reflects the availability of information from the holdings of the agency, also taking into account the suitability of the form in which the information is available, the media of dissemination, the availability of meta-data, and whether the user has reasonable opportunity to know it is available and how to access it. The affordability of that information to users in relation to its value to them is also an aspect of this characteristic.

*Interpretability* of data and information reflects the ease with which the user may understand and properly use and analyse the data or information. The adequacy of the definitions of concepts, target populations, variables and terminology underlying the data and information on any limitations of the data largely determines their degree of interpretability.

*Coherence* of data and information reflects the degree to which the data and information from a single statistical program, and data brought together across data sets or statistical programs, are logically connected and complete. Fully coherent data are logically consistent - internally, over time, and across products and programs. Where applicable, the concepts and target populations used or presented are logically distinguishable from similar, but not identical, concepts and target populations of other statistical programs, or from commonly used notions or terminology.

These elements of quality are overlapping and interrelated, often in a confounding manner. The current thinking at Statistics Canada, however, is that the elements may be viewed as the hierarchy below.



### 3. Approaches to quality

#### A. Eurostat

The starting point of Eurostat's definition of quality of statistics is the ISO norm 8402 that was mentioned before. This puts the user and his needs at the centre of all considerations. Looking at the term quality from the point of view of a user of international statistics has led Eurostat to a definition of (statistical) quality that is composed of seven dimensions:

- *Relevance*: Statistics are relevant when they meet users' needs. Relevance requires the identification of users and their expectations;
- *Accuracy*: is defined as the closeness between the estimated value and the (unknown) true value;
- *Timeliness and punctuality*: Statistics are only useful when the figures are up-to-date and published on time at pre-established dates;
- *Accessibility and clarity*: Data have most value when they are easily accessible by users, are available in the form users desire, and are adequately documented ("metadata" according to the type of user). Assistance in using and interpreting figures should be part of the providers' tasks;
- *Comparability*: Data are most useful when they enable reliable comparisons across space like countries or regions and over time.
- *Coherence*: When originating from a single source, statistics are coherent in as much as elementary concepts can be combined reliably in more complex ways. When originating from different sources, e.g. from different surveys with differing frequencies, statistics are coherent insofar as they are based on common definitions, classifications and methodological standards.
- *Completeness*: Areas for which statistics are available should reflect the needs and priorities expressed by the users. Completeness is an extension to relevance, for completeness does not only mean that statistics should serve user needs but also they should serve them as completely as possible, taking restricted resources into account.

#### B. IMF

The International Monetary Fund's Data Quality Assessment Framework (DQAF) was developed as an assessment methodology that aims to provide structure and a common language for the assessment of data quality. It is hierarchical and has five levels. The top level contains five *dimensions* of quality as well as a set of *prerequisites* for quality. The coverage of these dimensions recognizes that data quality encompasses characteristics related to the institution or system behind the production of the data as well as characteristics of the individual data product. The second and third levels of the framework contain *elements* and *indicators* of quality, respectively. These first three

levels are referred to as the generic framework as they are intended for use with all types of data.

*Prerequisites of Quality* refers to overarching institutional conditions for the pursuit of data quality. These conditions are identified in recognition of the idea that data users, who often cannot replicate or otherwise verify data, must place their trust in the institutions that produce statistics and the people who staff them. Further, these prerequisites typically influence more than one of the five dimensions in the DQAF. The DQAF groups the indicators of this kind into three elements: *legal and institutional environment, resources, and quality awareness*.

For the purposes of this discussion note, we deal with the first and parts of the second level of the framework only. The five IMF dimensions of quality are *integrity; methodological soundness; accuracy and reliability; serviceability; and accessibility*. They are defined as follows.

*Integrity* refers to values and related practices that maintain confidence in the eyes of users in the agency producing statistics and ultimately in the statistical product. Confidence by users is built over time. One important aspect is the trust in the objectivity of statistics. It implies that professionalism should guide policies and practices and it is supported by ethical standards and by transparency of policies and practices.

*Methodological soundness* refers to constructs and principles of accounting that are basic building blocks of macroeconomic data. Concepts and definitions, scope, classifications of flows/stocks, sector allocation of units, and valuation and time of recording are key aspects of macroeconomic datasets. These are typically covered in standards and guidelines that are agreed on in international consultative processes and in recognized sets of good practices. Thus, following the standards, guidelines, and practices promotes soundness of the data in these respects. International comparability is enhanced to the extent that countries follow international standards.

*Accuracy and reliability* refer to the soundness of the input data and statistical techniques and the degree to which output data portray reality. Accuracy and reliability are more specifically defined as follows.

*Accuracy* refers to the closeness between the estimated value and the true value that the statistics were intended to measure. Assessing the accuracy of an estimate involves evaluating the error associated with an estimate. In practical terms, there is no single aggregate or overall measure of accuracy. For sample survey-based estimates, the major sources of error include coverage, sampling, non-response, response, processing, and problems in dissemination. For derived estimates, such as for national accounts or balance of payments, sources of error arise from the surveys and censuses that provide source data; from the fact that source data do not fully meet the requirements of the accounts in terms of coverage, timing, and valuation and that the techniques used to compensate can only partially succeed; from seasonal adjustment; and from separation of price and quantity in the preparation of volume measures.

*Reliability* refers to the closeness of the initial estimated value to the subsequent estimated value. Assessing reliability involves comparing estimates over time. In other words, assessing reliability refers to revisions. In light of the policy and media attention given to first estimates, often a key point of interest is in how close a “flash” or preliminary estimate is to the subsequent estimates. In evaluating the closeness of the initial estimates to the subsequent estimates, it useful to consider the sources of revision, which include (1) replacement of preliminary source data with later data, (2) replacement of judgmental projections with source data, (3) changes in definitions or estimating procedures, and (4) updating of the base year for constant-price estimates. As a general rule, the smaller and fewer the revisions, the better.

*Serviceability* refers to the relevance, timeliness, and consistency of the statistics. It is broken down into *relevance, timeliness and consistency*.

*Relevance* refers to the extent to which data address the purpose for which users seek them. Their purpose is usually defined in terms of multiple uses and users. In assessing relevance, one approach is to gauge relevance directly, by polling users about the data. Indirect evidence of relevance may be found by ascertaining where there are processes in place to determine the uses of data and the views of their users or to use the data in-house for research and other analysis. The uses and users of a given dataset may change over time, and new needs may arise that require new data; thus, the best processes have a dynamic nature.

*Timeliness* refers to agreed IMF dissemination standards.

*Consistency* refers to logical and numerical coherence. It may refer to consistency over time, within datasets, and across datasets (often referred to as inter-sector consistency). In each, consistency in a looser sense carries the notion of "at least reconcilable." For example, if two series purporting to cover the same phenomena differ, the differences in time of recording, valuation, and coverage should be identified so that the series can be reconciled. Inconsistency over time refers to revisions that lead to breaks in series stemming from, for example, changes in concepts, definitions, and methodology. Inconsistency within datasets may exist, for example, when two sides of an implied balancing statement—assets and liabilities or inflows and outflows—do not balance. Inconsistency across datasets may exist when, for example, exports and imports in the national accounts do not reconcile with exports and imports within the balance or payments.

*Accessibility* refers to the availability of statistical information to the user. For this definition, both "availability" and "statistical information" are interpreted broadly, and the reference to "users" implies a range of users—public and private, sophisticated and casual, new and experienced. Availability refers to both the physical ability to obtain the information and the interpretability of that information. Statistical information refers to data and metadata (that is, descriptions of the data) as well as information about how to use and find data and metadata. The range of users leads to such considerations as multiple formats for dissemination and different levels of detail and technical content for metadata.

### ***C. OECD***

The OECD Statistics Directorate is currently developing a quality framework. A Task Force has drawn on the work of several (national and international) statistical organizations and has adapted this to the OECD context.

Quality is viewed in terms of eight dimensions:

- *Relevance* of data products for internal and external users; the degree to which the data serves to address the purposes for which they are sought by users.
- *Accuracy*: the degree to which the data correctly estimate or describe the quantities or characteristics that they are designed to measure.
- *Credibility*: refers to confidence that users place in those products based simply on their image of the data producer, i.e. the brand image. Credibility is in part determined by the integrity of the production process.
- *Timeliness*: reflects the length of time between the availability of data products and the event or phenomenon they describe.
- *Punctuality*: implies the existence of a publication schedule and reflects the degree to which the data are released in accordance with it.
- *Accessibility*: reflects how readily the data can be located and accessed from within the OECD data holdings. Internal and external users might have quite different perceptions of accessibility because of the differences of access methods.
- *Coherence*: reflects the degree to which data are logically connected and mutually consistent. Distinction can be made between coherence within a data set, coherence across datasets, coherence over time and coherence across countries. Achieving coherence across countries is one of the major sources of value added provided by the OECD.
- *Cost-efficiency*: is a measure of the costs and provider burden relative to the output.

### ***D. UNSD; a restricted approach***

While the initiatives of Eurostat, IMF and OECD are comprehensive in the sense that they aim at covering all aspects of statistical quality, some other international organizations<sup>7</sup> have been taking quality initiatives that focus more specifically on a limited set of quality elements. Two activities of UNSD are worth mentioning in this respect.

---

<sup>7</sup> A recent example is the questionnaire that FAO uses to assess the users' satisfaction with their statistical products. Other examples are the focus groups that UNSD sets up from time to time to assess publications or websites.

The first is the release of a Common Code of Statistical Practice in the United Nations System (1996). The 1996 Common Code has two parts. Part I gives guidelines for the presentation of statistical data in publications with an emphasis on clarity, completeness of meta-data, description of various sources of error, presentation of the limitations of results and offering assistance to users. Part II pertains to the planning stage of surveys to be undertaken by UN agencies and their representatives (e.g. consultants). Among other things, it stresses the need for: clarity of purpose, specification of objectives, target population and variables, explanation of feasibility of the survey, assessment of non-sampling error, description of data editing and quality control procedures etc<sup>8</sup>.

The second UNSD initiative relates to an internal review of data collection, data editing and data dissemination practices (a report will be finished by the end of 2002). UNSD collects a broad range of statistical data (from demography to international trade, and from environment to national accounts) from a wide variety of sources (national statistical offices, ministries, printed and electronic publications, other international organizations). The purpose of the review is to identify best practices in data capture, data editing and valuation, presentation of meta-data etc.

#### ***4. Approaches to quality; the case for harmonization***

In principle, harmonization of (statistical) ‘standards’ that international organizations develop is always desirable. This is particularly true when

- Standards apply to countries and country data, data collection methods, data frameworks, common definitions and classifications, in order to enhance international comparability of statistics.
- Different approaches, definitions, periodicity etc. would put an additional reporting burden on countries.

The question is to what extent the quality initiatives that were mentioned before affect countries in either of these two manners. It would seem that, at least currently, this is not necessarily the case. Whether or not the European Union considers imposing a standard quality model for statistics on Member States, is at this stage unclear. However, the EU Member States have already agreed on a Quality Declaration for the European Statistical System and Eurostat has already developed a Standard quality report/ reporting system with which countries may have to comply in the future. Likewise, the IMF is also imposing quality (dissemination) standards on countries that have signed up for the General Data Dissemination System) of the SDDS (Special Data Dissemination Standard). Several countries may have to deal with both Eurostat and IMF requirements. When OECD joins with its own quality standards and/or –reports, the burden for countries could increase yet.

---

<sup>8</sup> Chapter 2 of the Statistics Canada Quality Guidelines equally provides a description and quality checklist for each of the steps of the design and execution of a survey or collection.

Clearly, whatever the desirability for harmonization may be, achieving a unified quality system for official statistics is not going to be easy, considering how the different quality approaches and –systems concerned are embedded in agreements reached by the various organizations.

Nevertheless, it would seem useful to get the quality managers of the various organizations together to discuss the possibilities for convergence. Perhaps it is possible, as a beginning, to agree on a common ‘superstructure’ and on some issues of basic terminology. Looking at the various approaches, it seems to me that the following five common Dimensions and Aspects of quality may be distinguished.

- Dimension 1 are *User related elements*:

*Aspects*: Relevance, Timeliness and Punctuality

- Dimension 2 are *Presentation related elements*:

*Aspects* Accessibility, Clarity and Metadata

- Dimension 3 are *Data related elements*:

*Aspects*: Accuracy, Completeness and Comparability

- Dimension 4 are *Statistical process related elements*:

*Aspects*: Methodological soundness, Efficiency and Coherence

- Dimension 5 are *Institutional elements*:

*Aspects*: Legal environment, Planning mechanisms, Resources

I would furthermore suggest to work towards a set of basic, conceptual definitions for the Aspects<sup>9</sup>:

- Relevance: degree to which statistics meet users’ needs
- Timeliness: time elapsed between release of data and reference period
- Punctuality: degree to which pre-announced releases dates are met
- Accessibility: ease with which statistical data can be obtained by users
- Clarity: degree to which statistics are understandable for non-expert users
- Metadata: availability of information describing sources, definitions and methods
- Accuracy: distance between the estimated value and the (unknown) true value
- Completeness: degree to which statistics fully cover the phenomenon they are supposed to describe

---

<sup>9</sup> Clearly, these are just summary definitions that may have to be worked out and refined.

- Comparability: degree to which statistics are comparable over space (between countries) and time (between different time periods)
- Methodological soundness: adherence to professional methods and (internationally) agreed standards
- Efficiency: degree to which statistics are compiled in such a way that the reporting burden is minimized
- Coherence: degree to which data from a single statistical program, and data brought together across statistical programs, are logically connected
- Legal environment: degree to which statistical legislation is enacted, in conformity with the Fundamental Principles of Official Statistics.
- Planning mechanisms: the degree to which countries have instituted procedures for systematic, long-term planning of statistical operations.
- Resources: the degree to which statistical systems are properly funded and staffed.

## 5. *Conclusions*

- Some international organizations have developed quality assessment frameworks that look similar, but are different in scope, terminology and level of detail.
- When countries have to report to these international organizations about quality aspects of their statistics and statistical systems, they may have to do so more than once and they will inevitably be bothered by the differences that exist.
- Therefore, there is a case for some form of harmonization, both in terms of substance and perhaps also in terms of the collection of information.
- In this context, it might be useful if the quality managers of the various organizations got together to look at the possibilities for harmonization<sup>10</sup>.

---

<sup>10</sup> UNSD would agree to facilitate such a harmonization process.

## *References*

Eurostat (2000) Assessment of Quality in Statistics, April 2000

Eurostat (1998), 'Proposal for a quality report in statistics', Document A4/98

Franchet Y., (1998), Improving the quality of the European Statistical System, discussion paper for *DGINS Seminar*

Franchet Y., and W. Grünewald (2001), Eurostat's Approach to Quality, presented at the 51<sup>st</sup> Session of the International Statistical Institute, September 2001, Seoul

IMF (2001) Data Quality Assessment Framework, available at <http://dsbb.imf.org/glossary.pdf>

Grünewald, W. and H. Linden, Quality measurement –Eurostat experiences, Symposium on Achieving Data Quality in a Statistical Agency, Statistics Canada, Ottawa, 2001

Handbook on the Operation and Organization of a Statistical Agency, Revised Edition, United Nations Statistics Division; manuscript December 2001

Preparation of Sample Survey Reports; technical document of UN Statistics Division, 1992

Quality Framework for OECD Statistics; internal OECD paper, Paris, 2001

Quality Declaration of the European Statistical System, in: Summary Report from the Leadership Group (LEG) on Quality, Eurostat, Luxembourg, 31 July 2001.

Signora, A. and M. Colledge, Quality Framework Task Force, Notes on the definition and dimensions of quality; OECD, 2002

Statistics Canada (1998) *Quality Guidelines*, third edition, October 1998, available at [www.statcan.ca](http://www.statcan.ca).

De Vries, W., Co-ordination as an instrument to improve the quality of statistics (Statistical Journal UN/ECE, 1982);

De Vries, W., Performance, quality, management; how statistical offices try to do better, paper for UNSD Kuala Lumpur seminar; September 2001

De Vries, W., Are we measuring up? Questions on the performance of national statistical systems (International Statistical Review, 1998)

United Nations Statistics Division: Common Code of Statistical Practice in the United Nations System - Module 1. Part I. Dissemination and Publication of Statistics; Part II. Data Collection Practices: Planning Statistical Surveys, New York, April 1996.

United Nations Statistics Division, Good Practices in Official Statistics, <http://unstats.un.org/unsd/goodprac/default.asp>