Quality work in statistics – from Q2001 to 2016

Hans Viggo Sæbø

Statistics Norway, Oslo, Norway, hvs@ssb.no

Abstract
The paper considers the main developments in quality work and their impact on statistics since the first European conference on quality in statistics in 2001. In addition to the establishment of quality frameworks, the developments comprise a change of focus from quality assessments to assurance and from product quality to processes. In data collection, focus has changed from surveys to administrative registers and new data sources. National statistical institutions are not the sole producers of official statistics, and there is more attention on the whole statistical systems than before.
Together with technological improvements and more use of administrative data sources, this has resulted in more statistics produced more efficiently with improved timeliness. New data sources provide new possibilities, but also challenges for quality assurance. Reflecting on these developments may guide the way forward, on creating and maintaining a culture for continuous improvement in European and national statistics.

Keywords: Quality, quality assurance, continuous improvement

1. Background

Quality has been on the agenda in National Statistical Institutes (NSIs) and Eurostat the last 20 years. The first European conference on quality in statistics, Q2001 in Stockholm, was an important milestone. The basis for this conference was the work and conclusions from the Leadership Expert Group (LEG) on Quality. Later international cooperation initiatives in this area have set the terms for the next Q-conferences, which after 2004 have taken place each second year.

The paper considers the main developments in quality work and their impact on statistics since Q2001. Seeing the developments in a longer term perspective may guide the planning of new initiatives. What can we learn that is relevant for the future of quality work in statistics?

Developments in quality work must be seen in relation with the development of society and statistics in general, characterized and driven by globalization, technological developments and political changes. Quality work and its results are difficult to distinguish from improvements due to other causes. An example illustrating this is the obvious improvements in accessibility of statistics following the development of the Internet.

2. Q2001

The first European quality conference, Q2001 in Stockholm was structured around the recommendations from the Leadership Group (LEG) on Quality, see Eurostat (2002). These recommendations were based on current thinking on quality management, typically principles from Total Quality Management (TQM), and a survey of quality practices in NSIs (Japec, 2001). According to this survey, only a few NSIs had adopted a systematic approach on quality work at that time.

The LEG on Quality formulated 22 recommendations. These dealt with the need for a systematic approach to quality improvement, measurements of product and process quality, quality reporting and other quality management tools, documentation, training and the relationships with data suppliers and users. Accomplishment of regular Q-conferences was one of the recommendations. These issues were covered in Q2001, which in addition had sessions on metadata and methodology, and some on more general issues in statistics like data collection and business statistics.

The main message from the LEG was that quality is a multi-dimensional issue, covering not only accuracy but several components. Statisticians today agree that these can be described by relevance, accuracy and reliability, timeliness and punctuality, coherence and comparability, accessibility and clarity.

The main aspects of the development of quality work since Q2001 are considered below.

3. Main developments in quality work

3.1. From TQM to a variety of frameworks

A quality framework or management system consists basically of some definitions, principles and a model linking the principles together. General quality frameworks comprise Total Quality Management (TQM), Six Sigma, European Foundation for Quality Management (EFQM), Common Assessment Framework (CAF), Balanced Scorecard, ISO and Lean or Lean Six Sigma. These systems are to a large extent based on a common set of definitions (e.g. quality as "fit for use") and principles (such as user and process orientation, improvements based on measurements and participation by all), first formulated within TQM. But they differ with respect to main focus and degree of formalisation. In EFQM and ISO emphasis is for example put on rating and certification, whereas Six Sigma focuses on quality control applying statistical methodology. Lean emphasizes improved efficiency by the reduction of waste. One of the recommendations from the LEG on Quality was that all ESS organisations should adopt a systematic approach to quality improvement and use the EFQM model as a basis, except for those already using a similar model.

However, there is a set of values and principles of official statistics which go beyond the principles of these systems. This, in particular, regards independence, impartiality and protection of data on individuals. Such requirements to official statistics were first formulated jointly in the ten UN principles of official statistics adopted in 1994 (UN, 1994).

The European Statistics Code of Practice (CoP) which was established in 2005, is a milestone in European work on quality in statistics. It was revised in 2011 and followed up by peer reviews in 2006 – 2008 and 2013 – 2015 (Eurostat, 2011).

The revised version of CoP contains a new quality declaration for the European Statistical System (ESS) as a preamble, stating that the members of the ESS strive for joint cooperation and continuous interaction with users according to the principles of the European Statistics Code of Practice and general quality management principles including commitment of leadership, partnership, staff satisfaction and continuous improvement, in addition to integration and harmonization. Hence, some principles from general quality management

frameworks like TQM have been added to CoP as a common quality framework for the ESS. In addition, a "Quality Assurance Framework (QAF)" with good statistical practices is linked to the CoP principles and indicators, see Eurostat (2015). Compared to the first version of CoP, the 2011 revision also strengthened the Quality commitment principle and its indicators, and added a few more indicators, in particular to reinforce professional independence, better cover statistics based on administrative data, and to promote standardisation and communication of quality.

During the last years a series of other similar principles and frameworks for statistics have been developed by international organisations, for example the UN National Quality Assurance Framework (NQAF) (UN, 2013a). IMF and OECD also have their quality frameworks, and OECD has recently developed a recommendation on good statistical practice (OECD, 2015).

During the last 10 years, risk analysis has been implemented in the public sector in many countries, also in statistics and in international organisations. In some NSIs this has been integrated in their quality work.

Almost all the 32 countries exposed to European peer reviews in 2013 – 2015 followed some sort of systematic approach to quality management. In addition to CoP, TQM is used by many NSIs and EFQM and Lean by some. Some NSIs are certified according to ISO. In spite of this, most countries got recommendations on improvements in this area, often to enhance quality management and elaborate and publicise quality guidelines on their web sites.

3.2. From assessment to assurance

Almost all 22 recommendations from the LEG on Quality addressed measurements and analyses (of quality dimensions, processes, user and staff satisfaction) and documentation. It was followed by a LEG on Quality Implementation Group. This group initiated work on a series of handbooks with descriptions of good practices in the use of quality assessment tools in statistics.

The main results were summed up in a Handbook on Data Quality Assessment Methods and Tools (DatQAM) (Eurostat, 2007) covering quality reports and indicators. Sæbø (2014) has carried out an assessment of later experiences with these tools.

There has been a development from quality assessment to quality assurance during the later years. As pointed out by Sæbø (2006) stand-alone self-assessments do not necessarily provide a correct picture. But they constitute a convenient tool as a part of preparations for reviews and audits. They have been used prior to the European peer reviews, and in preparation for internal quality reviews carried out in several countries.

The extent of quality reviews has increased. At Q2001 reviews and audits were almost absent from the agenda, not to speak about experiences in his area. At Q2012 and Q2014 several examples of experiences with such reviews or audits were given. The last European peer review identified such reviews in several countries, like in Austria, France, Greece, Italy, UK and Norway. Some of these were judged as innovative practices by the peer review teams. In some cases audits are carried out with the purpose of certification. However, what is crucial is that reviews and audits result in improvement actions that are followed up regularly.

On the level of statistical systems or institutions there has been a large increase in assessments outside Europe, for example light peer reviews or adapted global assessments.

3.3. From products to processes

To obtain improved product quality it is necessary to work on improving the production processes. There is no doubt that there is more focus on processes in statistics today than 15 years ago. This development has been supported by the introduction of the Generic Statistical Business Process Model (GSBPM), see UNECE (2013). Such a model is a reference for the NSI's work on improving the processes, in particular by implementing standardized systems and methodology. In the last European peer review more than half of the countries stated that they had implemented a statistical business process model, and most of the remaining countries had done it partly (Eurostat, 2016a).

In Statistics Norway we have used the business process model both for structuring documentation, standardization, risk analyses and as a basis for our improvement work with internal quality reviews and within a Lean framework.

There is also a potential for better use of analyses of process variables. Sæbø (2014) points out that use of such variables other than resource inputs is still limited in official statistics, often confined to analysing response rates and managing interviewers. Editing and the effect of this is one area where this methodology should be suitable. This process normally counts for a relatively high share of resources used for the production of statistics.

3.4. From surveys to big data

The first Q-conferences were characterized by survey methodology to a larger extent than the latter ones. This was reflected in the name of the conferences in 2004 and 2006, including methodology and survey statistics explicitly in the title which for the other conferences have been "European conference on Quality in Official Statistics".

There has been a trend from quality issues in traditional surveys (sampling, questionnaires, non-response etc.) to administrative registers, and during the latest conferences new data sources and big data.

While Q2001 treated quality of administrative registers only in a session on business registers and macroeconomics, there were 5 sessions devoted to this in Q2014. In addition there were 2 sessions on big data and one on multi-source data production. In Q2016 there are 3 sessions both on administrative and big data and one on multi-source statistics. In addition, both administrative and big data are treated in other sessions.

Nordbotten (2012) gave a keynote speech at Q2012 in Athens on the developments in statistics before Q2001 and after 2012. He mentions major developments in the 20th century such as the introduction of punched cards even before the century started, the scientific development of sampling theory after the World War II, electronic processing in the 1950ties and the use of administrative registers for statistics from the 1960ties. The latter was boosted by the introduction of unique identification numbers, first for the population in the Nordic countries.

After that registers have become increasingly important and used as a source for statistics, also supplemented or linked to by data from other sources including survey data.

Big data are frequently discussed as a new source for statistics, even if they are not much used for official statistics so far. Big data are characterized by the size and complexity of the data sets, but also by being real-time or very timely data. Examples are data from social media on Internet, mobile telephone data, data from global positioning systems (GPS) and traffic data from automatic registrations. They often originate from an unknown population and represent methodological challenges. But they are not new, even if the notation came around 2005. Just think about data from satellite images to some extent used for statistics 30 years ago.

Statistics Norway today collects some data from Internet (web-scraping) for the consumer price index (e.g. flight prices). Another example of use of new and in this case big data will be data from automatic measurements of electricity consumption based on meters which will be installed at all consumers in 2017.

New data sources have challenged quality work and the tools that have been applied for sample surveys. The quality frameworks such as CoP/QAF have the end users of statistics as the point of departure. Users demand statistics with quality components such as accuracy and timeliness. The same components are valid for input data, but here the scope is the fitness for use by the statistics producer. A quality framework for input data, from registers or other sources, will therefore be different from CoP or others mentioned above, and should not be confused with these. Work has been done on developing indicators for both administrative registers and big data, see for example BLUE-ETS (2011) and UNECE (2014). In addition to the same or similar indicators as those used for quality of statistics, additional indicators comprise completeness and integrability (possibility to link data).

3.5. From NSIs to statistical systems

NSIs are not sole producers of official statistics, even if the NSIs normally have a coordination role that comprises quality assurance. The European Statistical system currently consists of Eurostat, 32 NSIs and 335 other producers who submit data to Eurostat (according to Eurostat web site). Representatives of the latter have participated in the Q-conferences, but these have

been dominated by participants and contributions from the NSIs, academics and international statistical organisations. At the European peer reviews completed last year coordination was an issue, and up to 3 other producers of European statistics or so called Other National Authorities (ONAs) filled out self-assessments and were interviewed. Out of 707 improvement actions for the NSIs, 121 were linked to their coordinating role (Eurostat, 2016b). Some of these dealt with quality issues, for example on reviewing the quality of statistics produced by other institutions.

The peer reviews also covered cooperation and integration in Europe, thus dealing with the quality of European initiatives. In the Vision 2020 quality work is focused (Eurostat, 2016c).

International quality work and cooperation have played an important role in supporting work on quality within the NSIs. Now quality of both the National and European Statistical Systems have come into focus, and this represents a challenge for our quality work that was not the same 15 years ago, or at least not thought about in the same way.

4. Have statistics improved?

There is no doubt that there is a higher consciousness about quality in statistics today than 15 years ago. That quality consists of several components which have to be balanced vs. each other and costs is common sense today. Use of tools such as self-assessments and internal quality audits are widespread.

The development has been facilitated by international work on quality, in Europe in particular linked to the establishment of CoP, the peer reviews and the follow up of improvement actions. It is believed that the Q conferences have contributed as well.

But has this led to improved quality in statistics? Marker (2015) has considered this. His obvious answer is yes, due to some of the same developments in quality work discussed in this paper. However, our users and their requirements change, so do the technological possibilities, and there is always a demand for new statistics and better timeliness. Changes after Q2001 in the product quality components of statistics as given in CoP are considered briefly in the

following, and some challenges are mentioned. Changes in the quality of processes and the institutional environment are considered summarily.

4.1. Relevance

New statistics have been produced in several areas, in Europe in particular due to new EU regulations believed to improve relevance. More statistics have been adapted to the needs of the general public also due to technological developments with Internet and social media, i.e. better accessibility. New developments linked to globalisation are addressed, but represent challenges to the coverage of economic statistics and our ability to describe migration, such as statistics on refugees. A recent challenge is how to describe and eventually follow up the Sustainable Development Goals. Official statistics is by nature a bit conservative, also to secure comparability over time.

4.2. Accuracy and reliability

The relative importance of accuracy has probably decreased, for the benefit of timeliness. This is in most cases corresponding to user needs, since the question is not to provide statistics that are as accurate as possible, but as accurate as needed for the purpose. More use of administrative data systems may improve accuracy, but accuracy of statistics based on sample surveys has probably decreased due to decreasing response rates linked to respondents' unwillingness to participate and difficulties to trace them. Statistics Norway has experienced reduced response rates for important voluntary surveys such as EU-SILC over a long period, and it costs more to keep it up on a reasonable level. New data based on bar codes, electronic traces and information on web are believed to be an alternative for some of these surveys, but peoples' attitudes are difficult to measure without asking.

4.3. Timeliness and punctuality

Timeliness has improved, by using more administrative data, better technology, more preliminary figures and just by measuring and publishing timeliness at least internally. This has in particular been the case for annual statistics, to a less extent for short term statistics.

Measuring punctuality according to international standards (with release calendars) has probably led to improvements in this area as well.

4.4. Coherence and comparability

In this area there is room for improvements. The internal quality reviews in Statistics Norway (Sæbø and Byfuglien, 2013 and Sæbø, 2014) have revealed a need for better coherence in published statistics, i.e. grouping and comparison of similar statistics covering the same area, from both Statistics Norway and other statistics producers. There has been an increase of other producers which represent a challenge in our coordination and quality work. More use of administrative data systems for statistics may also represent a challenge if these systems are not based on statistical standard definitions and classifications.

4.5. Accessibility and clarity

There is no need to justify that accessibility has improved a lot since 2001, due to the development of Internet and the possibilities for communication it has created. The web has developed from representing only a supplement to traditional channels for distribution of statistics at the time of Q2001, to the main if not the only channel for this now. Output databases where users can download statistics for free are widespread (the few possibilities that existed 15 years ago were often pay services). Use of APIs (Application Programming Interfaces) for bulk downloading of tables from the data banks of statistical institutions is an example of recent developments. This enables the users to utilise the statistics in their own environment and applications. Statistics Norway has just released this type of service in full scale. Microdata for research are available to a greater extent than before, but there are challenges linked to cover an increasing demand. In general there is still a need for better and simpler metadata and explanations for the users.

4.6. Statistical processes

Improved processes are a prerequisite for improved products, and there are reasons to believe that processes have improved accordingly. Key words are more and easier use of administrative data and more standardised systems driven by or in addition to technological

improvements. While electronic dissemination basically was established before 2001 (though developed a lot since then), there has been a breakthrough for electronic data collection since that time. For example, the share of data submitted electronically from industries to Statistics Norway has increased from close to zero to more than 90 percent during this period. However, there are challenges linked to the structure of data and metadata accompanying them, a prerequisite for efficient use of data for statistics, as well as for the understanding of the disseminated results. The best solution is to extract data directly from registers and administrative systems such as accounting systems of the data providers.

4.7. Institutional environment

As for the institutional environment, there is little doubt that the introduction and use of frameworks such as the CoP, supported by the peer reviews and following up by the international statistical organisations, have led to improved compliance with the principles on professional independence, objectivity and impartiality in addition to quality commitment in the NSIs. Equal access to statistical releases for all users at the same time are focused and supported by advance release calendars (first promoted by the IMF in 1996, and later emphasised in the CoP). Dissemination policy documents, which to an increasing extent seem to be available on the NSI web sites, will often describe this, as well as for instance principles on treatment of errors and of revisions. The UN (2013b) carried out a global assessment on the compliance with its Fundamental Principles of Official Statistics comparing 2003 and 2011. It was found that full compliance with the principles linked to independence had increased significantly in this period (for example from 44 to 60 per cent for the principle on relevance, impartiality and equal access).

5. Lessons learnt and the way forward

The message from Q2001 that quality has several dimensions is now common knowledge in statistics, well established in CoP and other quality frameworks. Balancing quality components and costs is the issue when it comes to quality work.

However, the explosion of quality frameworks and systems in general and for statistics might create some confusion, even if these are based on much of the same principles. Outside statistical institutions this development is driven by consultants, since it is a built-in feature of their business to promote new initiatives. Different international statistical organisations have also developed separate quality frameworks, though with many common aspects. There are differences in scope and target groups, but establishing new frameworks and indicators may imply double work. For a statistical institution that needs continuity, is it important to keep values and principles over time, also their wording. Deming's and TQM's principle constancy of purpose is just as valid today as when it was formulated more than 30 years ago (Deming, 1982). Future quality work must build on what has been successful in the past. Better coordination between initiatives taken by international statistical organisations is needed.

Exploiting new data sources is not new to statistics, but is more on the agenda than ever, also because traditional sources may be more difficult to use and due to competition from other actors. New sources challenge quality work, since quality of input data largely is uncontrolled or even unknown. There is a need to develop or rather agree on quality criteria for input data used for statistics, like it has been done for output or product quality. This does not presuppose completely new quality frameworks, and should not be confused with quality frameworks for a statistical institution or for output quality.

Quality assurance presupposes improvement actions that are followed up. Quality reviews and audits resulting in recommendations and actions lists which are followed up really make a difference and should be pursued.

There is room for more and better use of the business process model and analyses of process variables as a basis for improvements.

Things take time. Even if there have been developments in quality work and improvements in quality of statistics since Q2001, many of the same challenges persist, and there is still room for doing more following the recommendations from the LEG on quality. This for example regards improving relationship with both data suppliers and users, and developing,

recommending and in particular implementing good or best methods and practices linked to the most common processes.

Quality constitutes a key factor in the Vision 2020 for European statistics. An exercise to map quality elements of the Vision 2020 projects to CoP provides input to the next round of modifications of the code, together with the experiences from the last peer review round. It is already foreseen that the CoP may need to be extended to cover coordination and possibly innovation (better). The last area comprises big data.

Quality assessment and assurance in their nature are conservative, dealing with controlling and improving existing processes. However, user orientation and new data and technology require and render possible new solutions. Quality work is a continuous task and not only about doing things right, but also about doing the right things.

International cooperation will still play an important role supporting national quality work in statistics, and there will be no lack of issues for future Q-conferences.

References

BLUE-ETS (2011): *Blue- Enterprise and Trade Statistics – Deliverable 4: List of quality groups and indicators identified for administrative data Sources.*
Available at http://www.blue-ets.istat.it/fileadmin/deliverables/Deliverable4.1.pdf

Deming, W. E. (1982): *Out of the crisis.* Cambridge, Mass. MIT Center for Advanced Engineering Study, 1982.

Eurostat (2002): *Quality in the European statistical system – The way forward.*
Available from: http://ec.europa.eu/eurostat/en/web/products-eurostat-news/-/KS-45-02-814

Eurostat (2007): *Handbook on Data Quality Assessment Methods and Tools (DatQAM).*
Available from: http://ec.europa.eu/eurostat/web/quality/quality-reporting.

Eurostat (2011): *European Statistics Code of Practices.*
Available from: http://ec.europa.eu/eurostat/web/quality/overview.

Eurostat (2015): *The Quality Assurance Framework of the European Statistical System (ESS QAF).*
Available from http://ec.europa.eu/eurostat/web/quality/overview.

Eurostat (2016a): *Report to the ESSC on cooperation and integration within the European Statistical System.*
ESSC document 2016/28/7/EN

Eurostat (2016b): *Report from the Commission to the European Parliament and the Council on the implementation of the European Statistics Code of Practice and coordination within the European Statistical System.*
https://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-114-EN-F1-1.PDF

Eurostat (2016c): *The ESS Vision 2020.*
http://ec.europa.eu/eurostat/web/ess/about-us/ess-vision-2020

Japec, L. (2001): *Survey of Quality Practices in National Statistical Institutes.* Paper presented at Q2001, Stockholm, Sweden, May 14-15, 2001. Available from:
http://www.scb.se/sv_/Vara-tjanster/Kurser-och-seminarier/Dokumentationer-fran-tidigare-kurser-och-seminarier/Quality-Conference-May-1415-2001/Conference/

Marker, D. (2015): *How Much Has Quality Improved at National Statistical Institutes in the Last 25 Years?* Presentation at the 30th Anniversary of the Journal of Official Statistics Conference. Stockholm, Sweden, June 2015. Available from: http://www.scb.se/en_/Services/Courses-and-seminars/JOS-Anniversary-2015/JOS-Anniversary-2015/Presentation-Material/

Nordbotten, S. (2012): *Before the Q2001 and after the Q2012.* Paper presented at Q2012, Athens, Greece, June 2012. http://nordbotten.com/articles/Q2012.pdf

OECD (2015). *Recommendation of the Council on good statistical practice.* Available from http://acts.oecd.org/Instruments/ListByInstrumentDateView.aspx.

Sæbø, H.V. (2006): *Systematic Quality Work in official Statistics – Theory and Practice.* Paper for the European Conference on Quality in Survey Statistics, Cardiff, 24–26 April 2006. https://www.ssb.no/a/english/publikasjoner/pdf/doc_200606_en/doc_200606_en.pdf

Sæbø, H. V. and Byfuglien, J. (2013): *Quality reviews in Statistics Norway – Setup and experiences.* Paper presented at Nordic Statisticians' Meeting, Bergen August 2013. http://www.ssb.no/omssb/tjenester-og-verktoy/kurs-og-seminarer/nordisk-statistikermote-vedlegg/_attachment/130427?_ts=1402e6500d0.

Sæbø, H.V. (2014): *Quality Assessment and Improvement Methods in Statistics – What Works?* Statistika: Statistics and Economy Journal - No. 4/2014. Available at https://www.czso.cz/csu/czso/statistika-statistics-and-economy-journal-no-42014-4s5bpz2fko

UN (1994): *Fundamental Principles of Official Statistics.* Revised version with new preface in 2014 available from: http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx

UN (2013a): *National Quality Assurance Framework.* Available from: http://unstats.un.org/unsd/dnss//QualityNQAF/nqaf.aspx

UN (2013b): *Implementation surveys.* http://unstats.un.org/unsd/dnss/gp/globreview.aspx

UNECE (2013). *GSBPM (Generic Statistical Business Process Model).* http://www1.unece.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model.

UNECE (2014): *A Suggested Framework for the Quality of Big Data.* Deliverables of the UNECE Big Data Quality Task Team, December 2014. Available from:
http://www.unece.org/unece/search?q=A+Suggested+Framework+for+the+Quality+of+Big+Data.+&op=Search