Session 1: Quality assessment - use of checklist; user surveys; collection of quality metadata; disseminating quality information; training in quality management

IMPLEMENTATION OF QUALITY ASSURANCE ACTIVITIES
IN EUROSTAT
ANTONIO BAIGORRI, MARTINA HAHN AND HÅKAN LINDEN, EUROSTAT
I. INTRODUCTION

On its path towards the full implementation of the European Statistics Code of Practice¹, Eurostat, in 2007 adopted a Quality assurance framework (QAF) putting the requirements of the Code into concrete by providing guidelines for improvements at output and process levels. Its main features were presented to the CCSA meeting of September 2007 [Baigorri and Linden, 2007]. This paper reviews the progress with regard to the implementation of the QAF and outlines Eurostat plans on future quality assurance activities.

So far, Eurostat implementation of the QAF comprises the following five milestones:

- Adoption by the Eurostat Directors' Meeting of the methodology for quality assessments (October 2007)

- Elaboration of a plan for assessing the majority of statistical processes in Eurostat during the period 2008-2010 including a detailed work program for 2008 (January 2008)

- Integration of quality assessments with other horizontal activities in Eurostat such as evaluation, cost benefit analysis, management planning etc. (on going)

- Redefinition of a synthetic tool, the quality barometer (QB), to better reflect the existing quality information and users' demands (June 2008)

- Revision of the Eurostat and European Statistical System quality reporting guidelines (June 2008)

Following the discussions of the Working Party Quality in statistics held on 11 June 2008 and the favourable opinion given to some of the initiatives presented, Eurostat is planning to launch, in the coming months quality assurance activities related to the following issues:

- Evaluation of the 2008 program and elaboration of the 2009 Eurostat quality assessments work program

- Compilation of a limited quality barometer based on the information provided in the SDMX metadata structure quality concepts, once the switch from SDDS to SDMX has been accomplished in key statistical domains

- Facilitating and promoting of quality assurance activities in the European Statistical System (ESS)

2. 2007 -2008 IMPLEMENTATION OF THE EUROSTAT QUALITY ASSURANCE FRAMEWORK

2.1. Methodology for quality assessments

Quality assessments form the essential elements of evaluation within the Eurostat QAF. Their methodology has to be seen in the wider context of quality assurance contributing to the management of quality and Eurostat's compliance with the Code of Practice, and

¹ www.ec.europa.eu/eurostat/quality
consequently to enhancing the credibility of European statistics. A preliminary version of this methodology was included in the paper presented in the CCSA meeting of September 2007. The final version was adopted by the Eurostat Directors' Meeting in October 2007.

The following benefits of implementing quality assessments can be expected:

- At the level of Eurostat units, quality assessments fulfil two main goals.
  - To identify the strengths of statistical processes, which eventually leads to a better dissemination and transfer of good practices in the organisation.
  - To detect weak points in the production processes, which allows the implementation of improvement actions: by systematically considering every aspect of the production process and by comparing it to current best methods in quality assurance and requirements stemming from the European Statistics Code of Practice, quality assessments provide guidance for the prioritisation of improvement measures

- For Eurostat as a whole.
  - Providing support for resource allocation, planning and programming,
  - Helping improve internal coordination and communication,
  - Promoting standardisation of similar statistical processes

To implement quality assessments for most statistical processes over a mid-term horizon, Eurostat has developed several assessment categories and a number of criteria for allocating the processes to the assessment categories. A careful planning and preparation of all assessment categories is crucial for finding an optimal balance between the workload for Eurostat production units involved and the expected benefits for them.

- **Categories of quality assessments**

The Eurostat quality assessment methodology distinguishes the following categories of quality assessments, which are described in more detail in annex 1: self-assessment, supported self-assessment, internal and/or external peer review and rolling review. In order to compile information on processes, the Eurostat version of the Checklist for Quality Assessments of Statistics is used in every category of quality assessment. The Checklist is useful not only to assess the quality in a domain, but also to go through every aspect of the production of statistics, which helps in listing potential improvement actions. Experiences with using the checklist during two pilot quality assessments and improvements undertaken following the pilots have been presented by Eurostat to the CCSA meeting of September 2007 [Ecochard, 2007].

A Summary Assessment Report together with a Quality Diagram is established after each quality assessment. The report contains a list of actions that could be undertaken to improve the quality of the assessed statistical process. For each action, three elements are specified: who is in a position to achieve this action, what is the time frame for its implementation and the extent to which implementation already began. The exercise also attaches importance to identifying of good practices that could serve other Eurostat production processes.
- Criteria for linking the category of assessment to the statistical process to be assessed

For the purpose of the quality assessments statistical processes (and outputs) to be assessed are defined as the collection, processing and compilation and dissemination of statistics for the same area. In general, it is assumed that processes of a domain which lead to outputs with different periodicities (e.g. monthly and quarterly data) are so different from each other that they would benefit from an individual quality assessment each.

Thus, statistical processes as defined here provide a relatively homogenous entity that is stable over a given period of time and allow for equal treatment in assessing the various entities, comparing assessments and focusing on the statistical production process and its output. In addition they can serve other activities ongoing in Eurostat such as the review of priorities or defining costs (done normally within a rolling review), burden and benefits (done within almost every kind of quality assessment with the review of the relevance of the statistics). Item 2.3 covers this aspect in more detail.

A preliminary inventory of all the statistical processes was compiled from Eurostat units. It describes processes in terms of the following characteristics to be used for allocating the category of quality assessment to be applied:

- Relevance and visibility of output
- Resources of the process based on the number of full-time-equivalent staff
- Degree of involvement of Eurostat in the data validation and production chain
- Degree of NSIs or NCBs involvement (can be understood as a proxy for the degree to which data is based on official sources)
- Justification for data collection (legal basis, gentlemen's agreement, etc)
- Periodicity (to identify irregular statistical processes or processes with a less than annual periodicity)

Further characteristics serve the decision about the timing of the quality assessment:

- Availability of quality report considered a pre-requisite for the quality assessment
- Quality assessment already carried out and when

2.2 Quality assessments plan for 2008-2010

In early 2008 Eurostat established an office-wide assessment plan for reviewing most of the 128 statistical processes during 2008-2010 [Ecochard and Szczesna, 2008] keeping in mind the need to minimize the burden for production units and to apply as much flexibility as possible.

In allocating the category of assessment to the individual processes, the following principles were applied:
• Processes intensive in the use of human resources and producing regularly highly relevant output on a legal basis justify a rolling review. This concerns 11 processes.

• Processes involving one FTE staff or less, with either a periodicity less than annual, or where data are loaded in New Cronos only, would benefit from a self-assessment. There are 19 processes in this case.

• Processes for which Eurostat involvement and thus control of the process is rather minimal, are proposed to be assessed on explicit request of the domain manager only. There are five such processes.

• For other processes—constituting the majority— the type of assessment should be determined with flexibility, taking into account the characteristics of each specific process and the needs of the process owners.

The 2008 round of quality assessments was started with one hour seminars organized with each of the Eurostat Directorates concerned. During these seminars, the staff responsible for statistical processes to be assessed in 2008 received more in-depth information about how this exercise integrates in the overall quality activities of Eurostat, about the tools to be used and practical details of the assessments including on the categories for the assessment. From these seminars the following 2008 work program emerged: 16 self assessments (SA), 14 supported self assessments (SSA), 1 peer review (PR) and 4 rolling review (RR).

2.3 Integrating quality assurance with Eurostat activities

Quality assurance in Eurostat is an activity by itself but it is conceived to integrate at most existing demands on dissemination, reporting and compliance monitoring from the Commission and other stakeholders, by providing input that should avoid repetitive work, contribute to minimizing burden for production units and allow to profit from synergies of other horizontal activities in Eurostat. Fig. 1 includes a detailed version of the Eurostat QAF showing the links of quality assurance with other horizontal activities.

*Fig.1. QAF Integration with other applications*
• **Input for the evaluation function**

The Commission evaluation function requires the performance of evaluation activities related to the work program of each Directorate General. In relation to Eurostat, some of these evaluation activities are integrated in the QAF, in the form of rolling reviews, providing a consistent approach for quality assessments. Rolling reviews are quality assessment of high intensity where the external intervention is essential. As described in annex 1 external contractors monitored by Eurostat are selected to perform the review and the existing information on quality is complemented with user's surveys and contacts with producers. Evaluation activities in Eurostat also include reports to the European Parliament/Council that also serve as input for the quality assessments.

• **Standardisation of production processes**

The ongoing standardisation of the statistics production processes, as undertaken within the Eurostat data-life-cycle (CVD) project activities, is important for effective quality improvements and forms in so far a cornerstone of any quality assurance framework. Conversely, quality assessments, supported by the assessment checklist, help in the standardisation of processes in Eurostat through inter alia facilitating the application of common IT tools developed in the context of the CVD. Along these lines, statistical quality assurance also promotes the development and implementation of Current Best Methods for statistics production, protocols and policy documents.

• **Links between quality assessments and Euro-SDMX metadata structure**

The Euro-SDMX Metadata structure, which will replace the SDDS templates, fully integrates the ESS quality dimensions. When operational, the integration of quality (indicator) information into the Euro-SDMX template will, as a first step, facilitate the task of filling in the template from the existing information provided in quality reports. Once this information is available the Quality Barometer may be compiled based on the template without any further intervention of the production units in the data collection,

• **Input for the Eurostat Annual management plan**

Quality assessments help to identify weaknesses in some production processes and to formulate improvement actions needed to improve standards in some production processes. In so far these issues recur in several processes, the improvement actions should be incorporated into the Annual management plan for enhancing office-wide improvement actions. By having an integrated and office-wide approach, quality improvements can be achieved in a systematic way assuring comparability across different statistical processes and over times.

2.4. **Redesign of the Eurostat Quality barometer**

The Eurostat Quality barometer (QB) is a tool for providing early warnings in case there are structural weaknesses or in case of deteriorating quality in the ESS. At the same time, it could be used for monitoring the effects of quality improvements for the statistics production, and as well to ensure users that high quality statistics are produced in the ESS.

A pilot version of the QB was developed in 2007. The purpose of the pilot QB was twofold: (i) to identify the availability of quantitative quality information in the different statistical domains, and (ii) to provide an example on the structure for presenting
information on quality. The pilot QB was based only on available Eurostat summary quality reports/ quality monitoring reports in order to avoid additional burden for collecting the information. The focus was on quality dimensions rather than statistical domains and no analyses of results were provided. After this experience and following recommendations from the DM, at the end of 2007 an internal reflection was conducted in Eurostat, enlarged to an expert group of quality managers from some NSIs on how to better adapt the QB to management needs taking into consideration the sometimes limited quality information for some statistical domains.

Based on this reflection it was proposed to start with a restricted QB in 2008 and to establish a more general and integrated approach, in line with the original scope of the QB, when more information becomes available.

For the time being, it is not possible to have standardised quality information covering all main statistical outputs since:

- Quality reports do not exist for all the statistical outputs in the ESS.
- The quality reporting requirements are not fully harmonised.
- The existing quality reports are not always updated on a regular basis.

It has therefore been proposed to the Working Party Quality in Statistics [Eurostat, 2008a] to begin to produce a restricted QB. Such a version should be based on a limited set of quality indicators that are:

- applicable to most statistical processes,
- covering all the main quality dimensions,
- well defined and easy to provide, and
- easy to interpret and understand.

The Working Party has endorsed this proposal based on a restricted set of Standard Quality Indicators prioritised for the QB which is available in annex 2. Obviously the quality information is limited. This is the case for accuracy, as an example. The coefficient of variation, the unit response rates and the revision size show only parts of the inaccuracy. While this is clear for the producer of statistics it is important for the users to emphasise what has been measured and what not.

2.5 Revision of the European Statistical System quality reporting guidelines

The basic guidelines for reporting on the quality of statistics are given in the European Statistical System Standard Quality Report (SQR). The Standard Quality Report provides explanatory text on what to report with respect to the six quality components defined in the Definition of Quality. In addition, the Standard Quality Report is accompanied by a comprehensive explanatory handbook, How to make a Quality Report (HiMaQR), containing good practices for quality reporting.

The objective of the Standard Quality Report is to be a tool for the harmonisation of quality reporting across the ESS. Its purposes are to provide a common source for the definition of quality requirements formulated in Commission Regulations etc., to enable
cross-country quality comparisons of national surveys in the same statistical domain, and
to provide a basis for work towards cross domain quality comparisons.

The guidelines for quality reporting were last updated in 2003 and given the recent
developments in the areas of quality assurance frameworks, quality assessment tools and
methods, and experiences of process quality assessments in the ESS, the guidelines are
being revised. In particular, the following aspects will be elaborated and incorporated into
the *ESS Standard for Quality Reports* [Eurostat, 2008c] replacing the SQR:

- The context of quality management framework, quality assurance and quality
  assessment.
- Producer and user oriented quality information.
- Process quality and output quality.
- Reporting requirements (such as inclusion of standard quality indicators, quality
  assessment of administrative data and overall quality assessments).

The revised *ESS Guidelines for Quality Reports* [Eurostat, 2008b] replacing the HtMaQR
will be fully aligned with the ESS Standard for Quality Reports and, in addition, the
following aspects will be elaborated and incorporated:

- Integration of quality assessment tools (how to implement systematic quality
  assessment and how to work on continuous improvement of quality etc.).
- Practical guidance on how to calculate quality measures and references to the quality
  assessment sources used.
- Experiences and good practices for quality reporting in Eurostat and the Member
  States.
- Office-wide management tools for monitoring of data quality.
- Labelling and certification of statistics.

The revised versions of the quality reporting guidelines, e.g. the ESS Standard for
Quality reports, and the ESS Guidelines for Quality Reports are to be ready by
September 2008.

### 3. Future Eurostat Quality Assurance Activities

The remit of the European Statistical System Working Group on Quality in Statistics
encompasses both the internal Eurostat quality work and European Statistical System
quality activities. In the last meeting of 11 June 2008 Eurostat informed Member States
on its quality assurance framework and related activities, in particular on the three years
plan for the implementation of quality assessments in Eurostat. A proposal for
redesigning the quality barometer based on a restricted set of quality indicators, to be
compiled in the future on the basis of the Euro SDMX metadata structure was discussed
and agreed. Finally the plan prepared by Eurostat to enhance quality assurance activities
in the ESS based in the sharing of good practices and interchange of experiences was
welcomed by Member States and plans for implementation are on going. In the following
sections these future projects are described more in detail.

#### 3.1. The implementation of the Eurostat quality assessments: follow up
In 2008, 36 quality assessments are being implemented in Eurostat. By the end of the year the evaluation of the exercise will take place and a report is planned for the Directors' Meeting. This evaluation report should cover several aspects of the exercise conducted during 2008 including a possible need for adjusting the instruments used: Assessment Checklist, quality diagram, template for strengths/weaknesses and improvement action. The report should also indentify recurrent weaknesses in statistical processes in order to feed improvement actions into the Eurostat Annual Management Plan (AMP) and to identify initiatives that need to be implemented by horizontal units in order to facilitate quality improvements by production units.

In parallel to this report Eurostat will need to draw the Quality assessment program for 2009, which needs to be ready before the end of the year. While in 2008 the main focus has been laid on self-assessment and supported self-assessment, it would be desirable for 2009 to move forward in the process and to concentrate on more objective assessments as peer and rolling reviews.

3.2 Compilation of a restricted quality barometer based on the Euro SDMX Metadata Structure quality concepts

Following the endorsement of the proposal for a restricted QB by the Working Party Quality in statistics, Eurostat is planning to implement this approach by the end of the year. This more horizontal QB (e.g. covering only a few quality indicators that are available for many statistical processes) should allow drawing solid and elaborated conclusions although limited to those quality indicators covered. For statistical processes included in the foreseen successive implementation of the new Euro SDMX Metadata Structure across Eurostat after summer 2008, this template should be the main source of information for the restricted QB.

3.3. Plan to enhance quality assurance activities in the ESS

To foster the European Statistical System implementation of the European Statistics Code of Practice [Eidukynaitė and Hahn, 2008], in particular with regard to quality management and quality assurance activities related to Principle 4 of the Code "Quality commitment", a program steered towards the implementation of existing standards, guidelines and tools through mutual learning and support was agreed in the ESS. To this end a project on so called quality coaching has been discussed and supported by the Working Party Quality in statistics in its meeting of 11/6, bringing together quality managers from National Statistical Institutes. The terms of reference, to be defined with the support of Eurostat, will focus on the improvement actions agreed during the ESS peer reviews in the area of quality commitment. The following principles apply:

- **Voluntariness**: NSIs' participation (either as a coach or coachee) will be on a voluntary basis taking into account individual preferences and availability.

- **Openness**: NSIs participating in one activity as a coach can also participate in another as to be coached. While in principle targeted to the NSI's quality manager, quality coaching could also involve other members of the organisation.

- **Corroboration**: Coaching activities would supplement rather than substitute respective activities already organised bilaterally by some NSIs, thus enabling more NSIs to participate in an exchange in a more systematic way. Coaching will support NSIs improvement actions already planned, reinforcing the implementation path.
• **One-to-one:** While embedded in a common framework and terms of reference, the coaching process is based on a personal, informal and private one-to-one exchange. It does not involve lecturing or training of a broader audience. Coaching is thus a structured dialogue that enables to shadow own experiences, to follow-up planned activities and to address problems and obstacles where necessary and when they occur.

• **Framing:** The coaching exercise will be limited in terms of its scope and its timing. The focus will be on quality management and quality assurance activities only and the exercise will be carried out during a well-defined timeframe. All coaching projects will be completed before the 2009 meeting of the Working Group on Quality in Statistics so that the approach can be evaluated.

Following a contracting session organised by Eurostat, the quality coaching will basically consist of e-mail exchanges and structured telephone meetings. A concluding evaluation dialogue organised by Eurostat will serve to sum-up the experience and to draw conclusions.

The following benefits are expected:

• **At the level of the NSI to be coached**
  - Bespoke approach based on hands-on experience of the coach with the implementation of ESS standards and tools.
  - Personal support based on the immediate needs ranging from a confidential sounding board; a devil's advocate to test thinking and plans and/or simply a confidant to initiate quality activities.
  - More effective and possibly faster implementation of improvement actions.

• **At the level of the coaching NSI**
  - Rewarding and stimulating personal experience in the tradition of the peer review exercise, building personal contacts, gaining in-depth knowledge of NSI partners and possibly fresh look on own work.
  - Limited investment based on e-mail exchange and telephone meetings during a well-defined period of time, following a kick-off meeting organised by Eurostat.

• **At the level of the ESS**
  - Exploitation of synergies across the system, lowering overall implementation costs and frictions.
  - Intensified dialogue and relationships between ESS quality managers.
  - System-wide higher implementation path of an important area of the Code of Practice in line with ESS standards and good practices.
  - A boost to knowledge transfer and sharing of good practices with regard to quality management and quality assurance activities across the ESS.
REFERENCES


Annex 1. Categories of quality assessments

1. Self assessment

In a self-assessment, the Checklist is filled in by the person (or team) responsible for the statistical process. The role of the quality team is to assist the domain manager during this process and to ensure, to the extent possible, the coherence of assessments across Eurostat.

- The process manager fills in the assessment Checklist. In this process, the manager uses her or his knowledge of the process under assessment. Available information on quality, such as quality reports, process descriptions, reports to the Council/EP, are also used.

- The process manager compiles the Assessment Diagram and the Summary Assessment Report listing the strengths and weaknesses of the statistical process and the improvement actions.

- The quality team provides support on request. This will help:
  - To ease the process of filing the Checklist;
  - To assure a correct interpretation of the different elements of the Checklist;
  - In using the tools coming with the Checklist (the Assessment Diagram and the Summary Assessment Report).

2. Supported self-assessment

In a supported self-assessment, the Checklist is filled in under the responsibility of the domain manager with extended support from the quality team. Thus, the burden placed on the production unit is reduced and a high degree of coherence of assessments across statistical processes (and over time) is ensured.

- The quality team pre-fills the Checklist to the extent possible using the information available. Assessment questions, i.e. questions that are used to build the Assessment Diagram, are left unanswered at this stage.

- A first meeting between the process manager and another person from the production team on the one hand and the quality team on the other hand, takes place to review the answers proposed by the quality team. Assessment questions are filled in during this meeting.

- The quality team compiles a list of strengths and weaknesses and of improvement actions based on the information provided by the production unit and drafts a first version of the Summary Assessment Report and of the Assessment Diagram.

- In a second meeting, which foresees the participation of the Head of unit responsible for the process, the outcomes from the assessment are further discussed in order to come up with a version of the Summary Assessment Report and of the Assessment Diagram, on which the process owner takes on the responsibility: this procedure remains a self-assessment.

3. Peer review
In a peer review, the procedure is similar to that used for a supported self-assessment, except that an expert, not belonging to the production unit, is invited to take part in the assessment. The reviewer brings in technical expertise in the domain being assessed and increased objectivity, making for greater credibility of the assessment.

The reviewer can be an expert from within Eurostat (internal peer review), in which case he or she will possess an extensive knowledge of the domain. The reviewer can be a senior statistician for instance. He or she can also be chosen outside of Eurostat, for example in NSIs (external peer review). In that case, the objectivity of the review is strengthened; this can be useful for the review of statistics where this element is important.

4. Rolling Review

In a Rolling Review a more complex assessment of the statistical process is implemented by reviewing the statistical data, the process to produce them, the interactions with data providers and the user satisfaction. An external contractor will implement the rolling reviews and be supported by the evaluation function of Eurostat.

A Rolling Review involves:

- the implementation of a partner survey addressed to the providers of the statistical data, namely the National Statistical Institutes of the 33 Member States, EFTA countries and candidate countries (see Annex 1).

- the implementation of a user survey that is addressed either to a known cycle of users or a wide cycle of users. In case the response rate is very low, interviews with a selected number of users may be used to get a better overview on users’ views (see Annex 2).

- an assessment of statistical products and outputs addressing the unit inside Eurostat producing the statistical data. The assessment of statistical products and outputs should be done with the help of the production unit in Eurostat responsible for the statistical process under review.

- partner and user surveys should be implemented in a majority via electronic means (e-mail, web questionnaire or similar) but it could also be envisaged to interview a selected number of partners and users face-to-face to receive better feedback on the questions.

- at the end a report will be prepared containing the summary and analysis of all the surveys and assessments, a review of the production process itself as well as recommendations for improvement.

- the recommendations for improvement should be discussed thoroughly with the unit concerned and those recommendations accepted should then be timed and followed-up.
## Annex 2: Restricted set of Standard Quality Indicators prioritised for the QB

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Methodology</th>
<th>Level of calculation</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R2) Rate of available statistics</td>
<td>(number of values provided in a dataset) / (total number of fields for which data has to be provided)</td>
<td>NSI, EU</td>
<td>- Number of values to be provided/ applicable is often according to requirements in regulations.</td>
</tr>
</tbody>
</table>
| (A1) Coefficient of variations | (the square root of the estimate of the sampling variance) / (the estimated value) | NSI, EU | - To be calculated if sample surveys.  
- Aspects to be included in variance, like non-response, coverage adjustment etc. has to be defined. |
| (A2) Unit response rates | - Un-weighted unit response rate: (number of respondent units used in estimation) / (number of in-scope units + number of eligibility unknown units)  
- Weighted unit response rate | NSI | - To be included when the applied definitions are harmonised. |
| (A7) Average size of revisions | - Mean absolute revision  
\[ MAR = \frac{1}{T} \sum |R_t| = \frac{1}{T} \sum \left| \hat{\theta}_t - \theta_t \right| \] | NSI, EU | - Useful measure to gauge the size of revisions avoiding offsetting effects due to neg. and pos. revisions. |
| (T2) Time lag between the end of the reference period and the date of first results | (release date of provisional results) – (date of reference for the data) | EU | - The time unit should be number of days. |
| (AC2) Number of accesses to databases | - number of (unique) hits  
- number of extracts | EU | - Has to be agreed to measure hits and/or extracts.  
- Not always possible to measure the number unique hits. |
| (AC3) Rate of completeness of metadata for released statistics | sum of availability (for example: 2=fully; 1=partly; 0=not available) of metadata for the different aspects included in the metadata template divided by the total for the relevant aspects for which metadata should be provided | NSI, EU | - Will be possible to calculate when the Euro-SDMX is implemented as standard template in the ESS |
| (C1) Length of comparable time series | number of reference periods from last break in time-series | NSI, EU |

(i) In general all indicators but in particular the indicators A1, A2, A7 and C1 requires that the survey methodologies are to some extent harmonised in order to be able to define calculation formulae that allow meaningful comparisons across countries.

(ii) The calculation method of the quality indicators at EU level require that the weighting procedures are specified as well.