

QUALISTAT Sub-Group

Quality of Norms

Final Report

July 1998

Contents

Section	Title
0	Executive Summary
1	Background and mandate of the Sub-group
2	Norms (statistical standards); their role in the production and use of statistics; different categories of norms
2.1	Definition of the term ‘norm’
2.2	The role of norms in the production and use of statistics
2.3	Categories of norms
3	The scope of the Sub-group: types of norms to be considered
3.1	Categories in scope
3.2	Examples of relevant norms
4	The concept of quality and its application to statistical norms; user groups to be considered; criteria for the assessment of quality
4.1	Definition of quality
4.2	Application of quality to norms
4.3	Users of norms
4.4	Meeting the needs
5	Measurement of the quality in practice
5.1	Ways to assess the quality of norms
5.2	Possible methodologies for assessing the quality of a norm
5.3	Method A – description
5.4	Method B – description
5.5	Relative advantages and disadvantages of the various combinations of methods and approaches
6	Conclusions and recommendations of the Sub-group
7	Bibliography
8	Members of the Sub-group
Annex 1	Worked example of method A using NACE Rev. 1
Annex 2	Detailed checklist of quality criteria for method B
Annex 3	Sample quality report using method B - NACE Rev. 1

0. Executive Summary

The Total Quality Management approach adopted by Eurostat requires that the quality of all outputs of the organisation should be assessed. As norms (or statistical standards) are an important output, a framework for measuring their quality is required. In May 1997 the Quality Assurance Group decided to create this Sub-group to draw up a suitable framework. The mandate for the Sub-group was to propose a definition of quality for norms, and the means to measure that quality. It did not extend to making proposals about quality assurance, or when and by whom quality assessments should be done.

This report defines the term norm, proposes classification systems for norms, and gives examples of relevant norms. It then takes the ISO definition of quality, and applies this to norms, identifying three main groups of users, and considering their needs.

Section five proposes two related methods to assess the quality of norms, one that is simple and quick, but rather crude and subjective, and one that is much more detailed and more objective, but also more time consuming. The relative advantages and disadvantages of each method are given, as well as worked examples of both methods (Annexes 1 and 3).

The conclusions and recommendations of the Sub-group are presented in Section six. They can be summarised as follows;

- The quality of a norm is likely to be difficult to assess objectively.
- The quality of individual norms should not be assessed in isolation, it is important to consider their role within wider frameworks of norms.
- The ability of a norm to satisfy the needs of its users is the best criterion by which to assess its quality. This is in line with the ISO 8402 definition of quality.
- The resources devoted to assessing the quality of a norm should be proportionate to the cost of developing that norm.
- In the future, quality assessment should be an integral part of the procedure of developing a norm.
- The methods described in this report should be used to assess the quality of a few key norms, the experience gained can then be used to refine the methods if necessary.
- There is a need for close co-ordination of the work on the quality of statistics and quality of norms. The role of the Eurostat Quality Manager is important in this respect.

1. Background and mandate of the Sub-group

One of the main tasks of Eurostat is to develop “a set of norms and methods which allow impartial, reliable, relevant and cost-effective statistics to be produced throughout the Community” and to “prepare, further develop and promote the adoption of Community statistical standards by Member States in order to improve the comparability of Community statistics, as well as the cost-effectiveness of their production”¹.

The Total Quality Management approach adopted by Eurostat requires that the quality of all outputs of the organisation should be assessed. If the current level of quality is known, it is easier to assess future improvements in quality. As no framework or methods for measuring the quality of norms existed it was decided to create a Sub-group of the Quality Assurance Group to tackle this issue.

The mandate for the Sub-group was to propose a definition of quality for norms, and the means to measure that quality. It did not include making proposals about quality assurance, who should carry out the task of assessing the quality of norms within Eurostat, or the timetable for such work. It should be emphasised here that the aim is to evaluate norms as products, and not to evaluate the people responsible for them.

Many norms are the product of long discussions in committees and are therefore compromises rather than ideal solutions. Making changes to such norms as a result of a quality assessment could take some time, particularly if new legislation is required, though if the quality of a norm is assessed to be poor, this can only increase pressure for improvements.

2. Norms (statistical standards) their role in the production and use of statistics; different categories of norms

2.1 Definition of the term 'norm'

Dictionary definitions;

Norm - (from Latin *norma* meaning a carpenter's square) A standard, pattern or type

Standard - A measure established as a criterion, any type, fact, thing, etc. serving as a criterion

For the purposes of the work of the Sub-group we agreed to use the words 'norm' and 'standard' as synonyms, and define a norm as 'an agreed or recognised convention which contributes towards a framework for collecting, processing or disseminating coherent statistical data'.

2.2 The role of norms in the production and use of statistics

Norms are vital tools for the production and use of statistics. All statistics depend on underlying concepts and definitions, and many make use of classification systems in order to categorise data and define homogenous groups. Established and recognised methods are normally used to process raw data into statistics. Harmonised norms greatly enhance the usefulness of statistics by making them easier to understand, and allowing meaningful comparisons to be made, particularly at the international level².

¹ Commission Decision (97/281/EC) of 21 April 1997 on the role of Eurostat as regards the production of Community statistics.

² “L’harmonisation des concepts et des définitions comme étape indispensable de l’intégration des statistiques” Michel Poulain, GÉDAP, Université Catholique de Louvain.

2.3 Categories of norms

There are various possible ways of classifying norms, including the following;

- (a)
 - Concepts
 - Classification systems - by size, type, activity, location, etc.
 - Units (collection, processing, measurement, output) and groupings of units
 - Characteristics - the variables observed
 - Derived concepts e.g. population, Gross Domestic Product
 - Summary statistics e.g. mean, variance, inter-quartile range
 - Production
 - Methods of sampling
 - Data collection and processing
 - Calculation and presentation of results
 - Background
 - Infrastructure and legal framework
 - Confidentiality
 - Habits, traditions, etc.
- (b)
 - Compulsory (enforced by law or EU regulation)³
 - Advisory (set out in guidelines, recommendations, manuals of best practice, gentleman's agreements etc.)
 - Habitual (not written down, but traditionally used)
- (c)
 - Statistical output
 - Collection of data
 - Transmission of data
- (d)
 - Used within the European Statistical System (ESS)
 - Not used within the ESS

Norms are also present in the terminology underlying all of the above categories, and are particularly important when that terminology has to be translated into many languages.

Some legal or agreed frameworks such as the European System of Accounts (ESA)⁴ can be regarded either as a single, aggregate norm, or as an integrated set of norms. If the latter is the case it is important to bear in mind the ways in which the individual norms interact within the overall framework. The interaction of norms can be at various levels, in some cases norms may be very closely linked at the practical level, whereas in others the links may be more theoretical in nature.

3. The scope of the Sub-group: types of norms to be considered

3.1 Categories in scope

The aim of the sub-group is to provide a universal framework which can be used for assessing the quality of all statistical norms, whilst concentrating on those used within the ESS. Therefore the main categories of norms to be considered are those relating to concepts, i.e. classification systems, units, characteristics and derived concepts (but not those relating to summary statistics). Norms relating to production of statistics, for example the use of primary registers, are also considered. Both compulsory and advisory norms are included, as they are often complementary. These norms may relate to statistical output, collection or transmission of data.

³ Norms that are compulsory for one user or group of users are not necessarily compulsory for all users.

⁴ Council Regulation (EC) 2223/96 of 25 June 1996 on the European system of national and regional accounts in the Community.

Of the norms used within the ESS, very few are actually created by Eurostat in the strictest sense, some are developed from those used by other organisations, and some are adopted without modification. Often Eurostat plays a part in developing world norms alongside other international bodies. Involvement in this process at the earliest possible opportunity is clearly desirable. The degree of control that Eurostat has over a particular norm should be considered when assessing the quality of that norm.

3.2 Examples of relevant norms

3.2.1 Classification norms

- Size - SME definition (Commission recommendation)
- Type - ESA institutional sector (financial and non-financial corporations, households, etc.)
- Activity or occupation - NACE Rev.1, ISCO
- Product - CPA, CN, PRODCOM list
- Location - NUTS
- Social and demographic - ISCED, ICD, Social Indicators

3.2.2 Unit norms

- Statistical Units, e.g. units for business statistics as defined in the Council Regulation on Statistical Units (696/93), and those used in social statistics such as “household”
- Measurement units e.g. purchasing power parities

3.2.3 Characteristics norms

- Definitions of turnover, employment, etc. in the SNA/ESA.
- Harmonised core variables for population and housing census, and household surveys
- ILO definitions for labour force surveys

3.2.4 Production norms

- Council Regulation on Community co-ordination in drawing up business registers for statistical purposes

4. The concept of quality and its application to statistical norms; user groups to be considered; criteria for the assessment of quality

4.1 Definition of quality

The ISO 8402 definition of quality is 'The totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs'. Seven criteria are generally used within Eurostat for determining the quality of statistics⁵;

- Relevance
- Accuracy
- Timeliness and punctuality
- Accessibility and clarity of the information
- Comparability (over space and time)
- Coherence
- Completeness

Cost, including response burden, is considered to be an external constraint.

⁵ Defined by the Eurostat Quality Assurance Group.

All of these criteria for the quality of statistics are affected by the quality of norms, and a modified version of them can be used when determining the quality of norms, thus there is a two way linkage.

4.2 Application of quality to norms

The proposed criteria for assessing the quality of norms are listed below. They are expanded further in Annex 2. Some terms are combined where this is logical, e.g. clarity, specificity and accessibility are combined because they all concern how the norm is conveyed to users.

- **Relevance** - It is easier for suppliers to give data if the norms used are relevant and meaningful to them. Norms must also be appropriate and relevant for the producers of statistics and the users of the statistics produced.
- **Clarity, specificity and accessibility** - Norms should be easily understood by all parties, and particularly in the case of terminology, should be easy to translate into other languages. Sometimes it is better to have several sentences defining a concept, as the meaning of a single word can be distorted by translation. Norms must also be easily accessible to users, in a suitable format.
- **Conformity** - There are two aspects to conformity, whether users apply norms correctly and consistently, and whether the norm conforms with external requirements, e.g. legal, data protection and ISO requirements (i.e. that it should be possible to implement norms, and also to check that they are applied correctly).
- **Collectability and flexibility** - The ease (and cost) of collection of data relating to a particular norm should be considered. Flexibility in the way that a norm can be applied e.g. the possibility of further national breakdowns of an international classification system, or the flexibility to move progressively towards a common target, can improve the quality of a norm from the point of view of those applying it. Flexibility to adapt over time can also help the norm maintain its relevance. However, there is also a danger that flexibility could have adverse effect on universality and stability, thus reducing the ability of the norm to meet the needs of those wanting to make such comparisons.
- **Universality and stability** - Universality refers to comparability between countries, e.g. for European purposes norms must apply to all Member States. It is useful if they can also be used in a global context to allow international comparisons. Stability refers to consistency over time, changes to a norm can cause discontinuities in statistical time series, though the quality of a norm may change over time, e.g. a classification system may become outdated.
- **Coherence** – Harmonised norms that can be used for many purposes across different statistical domains, whilst still meeting key needs, are preferable. It is important to be aware that as such norms have a wide range of users it is more difficult for them to meet all needs. Norms designed for very specific purposes may at first seem to be of higher quality because they meet all needs of their very limited range of users. Complementarity of norms in related domains is also important, particularly in large systems such as national accounts. Compatibility with non-statistical (e.g. accounting) norms may often be desirable.
- **Exhaustiveness** - There are two aspects to exhaustiveness. Exhaustiveness within a norm, i.e. whether it covers all necessary features of what is being described or measured, and exhaustiveness of a system of norms, i.e. whether any further norms are needed to produce certain statistics. Exhaustiveness is of particular importance for classification systems, where all possibilities should

be covered, and uniqueness, i.e. the principle that it should only be possible to classify something to one and only one category, is also important.

4.3 Users of Norms

The quality of a norm is determined by its ability to satisfy stated or implied needs. The producer of statistical norms should consider the needs of three main groups below (there may be some overlap between these groups).

- Data suppliers
- Producers of statistics - at regional, national, and international levels
- Users of statistics (intermediate and final) - as far as Eurostat is concerned these can be further split into six sub-groups;⁶
 - European Union institutions e.g. Commission, Parliament, Council
 - Government institutions within Member States, local, regional and national, including National Statistical Institutions
 - International organisations e.g. United Nations, OECD
 - Bodies representing others e.g. trade federations, trade unions, the media
 - Business users e.g. enterprises
 - Other users e.g. academics and the general public

To get an overall measure of quality it is necessary to identify the users in each group, determine their needs, and see how well they are met. For norms with many users it might be worth considering which users are most relevant or important to the statistical process, and concentrating on their needs. It is important to remember that statistical norms may fully meet the needs of the original intended users, but may not meet the needs of subsequent, perhaps unforeseen, users. Therefore the stage a norm has reached within its life-cycle is relevant when determining its quality, and it may be necessary to re-evaluate the quality of the norm when new users become known.

In addition it is sometimes necessary to consider a fourth group, those who use statistical norms for non-statistical purposes, for example, those who use statistical classification systems for administrative purposes. It is quite likely that many of these users of norms will be unknown, and therefore it will be difficult or impossible to assess their needs. As the purpose of this exercise is to provide a means for the evaluation of statistical norms, the main focus will be on the needs of users of norms in the field of statistics, but also including some consideration, where relevant, of the needs of other users, especially those of administrative bodies that supply data to be used for statistical purposes.

4.4 Meeting the needs

Users and producers of statistics often want more detail than suppliers of data are willing to give. There is always pressure to minimise the statistical (and administrative) burden, and therefore financial cost, to suppliers. Good norms can help here. For data suppliers the key factors will therefore be relevance, collectability and clarity.

Producers of statistics will also consider relevance, collectability and clarity to be important, particularly when they have to apply norms to their data. In order to check their data they may also consider universality, coherence, and stability over time to be important. If they are to publish their data they will also have to consider legal limitations.

⁶ Defined by the Eurostat Survey of Users Group.

For users of statistics, the key factors in the quality of norms will be more varied. Specialist users, with experience of dealing with statistics are likely to want universality, stability and coherence, whereas generalist, casual users of statistics would probably place more emphasis on clarity and relevance in order to be able to easily understand what the statistics are telling them.

The term 'needs' will always include elements of cost and time for all three groups of users of norms. A norm that meets most requirements may be considered to be of higher quality than one that meets all requirements but would make the statistics considerably more expensive or time-consuming to produce.

5. Measurement of the quality in practice

5.1 Ways to assess the quality of norms

There are various possible ways to assess the quality of a norm;

- Self-evaluation - this could include empirical methods where possible - e.g. comparing variances within and between the groupings of classification systems.
- Surveys of users – asking users of the norm to what extent it meets their needs. Ideas as to how the norm might be improved may also be useful when considering future revisions to the norm.
- Peer reviews - i.e. inviting competent external bodies to assess quality. This could also include an element of benchmarking - i.e. making comparisons with other organisations for mutual benefit.

It is important that the approach chosen is made as objective as possible. Empirical methods are likely to be the most objective, but may be very difficult to apply to many norms.

It might be possible to assess the quality of voluntary norms by looking at their take-up rate. The more they satisfy the needs of the relevant parties, the more likely they are to be used.

5.2 Possible methodologies for assessing the quality of a norm

Two separate methodologies are considered here. The first, (method A), is intended as a relatively quick, easy and therefore low-cost approach. The second, (method B), makes use of a checklist of previously determined criteria (see Annex 2) to allow a much more thorough evaluation of all aspects of the quality of the norm. Both methods can be adapted for use with the self-evaluation, survey of users and peer review approaches.

These methods differ slightly from those used to measure the quality of statistics in that they involve the production of a report combining a numerical score with a description of the development of the norm. The second element is necessary because it provides useful background information and helps to explain why a norm was developed in a certain way. It should be written in a neutral tone, and should not attempt to judge the process used to develop the norm because this would detract from the aim of measuring the quality of the product. This part of the report may also include an assessment of whether revisions to the norm should be considered.

It is important that the numerical scores should not be taken as absolute measures of quality, but rather as a guide. They are most relevant when considering norms with a high degree of similarity of purpose and use. This is because norms developed and used only for very specific purposes and a limited number of users, are likely to better satisfy user needs, and therefore score more highly, than norms with many users and widespread applications, without necessarily being “better” norms.

5.3 Method A – description

- Determine the category of the norm, including its origins, evolution, purpose and the degree of control Eurostat has over it.
- Determine the users of the norm, group them into data suppliers, producers of statistics, and users of statistics.
- Assess how well the norm meets the needs of each of these groups of users, taking into account related norms and larger frameworks of which the norm may be part. Consider also the constraints of cost and time.
- Determine the overall quality of the norm, again in the context of a wider framework of norms where appropriate. This should be done by giving a score for each of the three main groups of users of norms as follows;

For each group:	Norm meets all needs	Score = 5
	Norm meets between 80% and 99% of needs	Score = 4
	Norm meets between 60% and 79% of needs	Score = 3
	Norm meets between 40% and 59% of needs	Score = 2
	Norm meets between 20% and 39% of needs	Score = 1
	Norm meets less than 20% of needs	Score = 0

The three scores can then be presented separately or added together to give the norm an overall score in the range zero to fifteen. There is, however, a danger that in extreme cases (e.g. scores of 5, 5 and 0) a single total may give a misleading impression of the quality of the norm.

- Produce a report detailing the steps above and including a brief description of the development of the norm, considering constraints such as the need for consistency with related norms, and including recommendations for revision if appropriate. Please note, however, that the consideration of constraints should not be used solely as a means to justify a poor quality score.

A worked example of the assessment of quality of a norm using this method is given in annex 1.

5.4 Method B – description

This method uses the first two steps of method A (i.e. determining the category and users of the norm), followed by a description of which users of the norm have been considered, and how their needs have been taken into account. The next step involves looking in detail at the different components of quality for a particular norm. Annex 2 contains a checklist of the criteria to be considered.

In method A the three groups of users of norms are assumed to have the same weighting, this assumption may not always be valid. In method B they are each assigned a relative weight (w_g) in the range 0 to 15, such that the sum of these weights ($\sum w_g$) is equal to 15. For example if a norm is of little importance to data suppliers or producers of statistics, but is vital to users of statistics the weights could be 2; 2; 11 respectively.

The list of criteria in Annex 2 is then considered separately for each user group. It is, however, recognised that the seven criteria will not all be equally relevant to all users of norms. Almost all should be relevant to users of statistics, and most to producers of statistics. The needs of data suppliers are rather different, therefore some of the criteria will have little relevance for this group.

To compensate for this a weight is assigned to each of the seven criteria (w_c) such that the sum of these weights ($\sum w_c$) is equal to 100. Criteria that are of little relevance to a particular group can therefore be given a low (or even zero) weight for that group. A score for each criterion (s_c) is determined making reference to the points listed under that criterion, such that $0 \leq s_c \leq w_c$. The total score (S_g) is then calculated (i.e. $S_g = \sum s_c$) and multiplied by the weight for that user group to give the weighted score ($S_g w_g$). When this process has been completed for all three main user groups the three weighted scores should be added to give the overall score for the norm ($\sum S_g w_g$). This score should be between 0 and 1500, and if divided by 100 can be compared with scores obtained from method A.

The wide variety of statistical norms means that some of the detailed points within each criterion will be more relevant for certain types of norms than others. These points are not always directly linked to the quality of norms, (e.g. 3.1.2 Do all users apply the norm consistently?), but may indicate possible quality defects. In certain cases it may be appropriate to consider other points as well as, or instead of, those listed in Annex 2, as long as they fall within the scope of the criterion being considered.

As well as the numerical score, the report on the quality of the norm should include;

- A detailed description of the procedure followed in order to create the norm, including the policy background, compromises due to negotiations and lobbying, the methodological framework, and constraints such as the need for consistency with related norms.
- Relevant information on the implementation and subsequent development of the norm.
- How the norm interacts with other norms, and how it fits into larger frameworks of norms.
- References to any similar norms that might be of higher quality.

Please note, however, that the consideration of constraints should not be used solely as a means to justify a poor quality score.

A worked example of the assessment of quality of a norm using this method is given in annex 3.

5.5 Relative advantages and disadvantages of the various combinations of methods and approaches

The main advantages and disadvantages of each method are as follows;

- Method A** Advantages;
- Simple and easy to understand
 - Can be used by non-specialists and with little training
 - Fairly quick, and therefore cheap

Disadvantages;

- Very subjective
- Crude scoring method

Method B

Advantages;

- Thorough assessment of all aspects of quality
- More accurate scoring method for purposes of comparison
- More structured approach
- Provides a suitable structure for quality reports
- More objective than method A, but still at least partly subjective

Disadvantages;

- More time-consuming therefore more expensive

Each combination of method and approach also has potential advantages and disadvantages. The table below shows the main problems that might be encountered for each combination.

	Method A	Method B
Self-evaluation	Self-evaluation must not become self-justification.	
	There is a danger that reviews may be too subjective, and that most norms will be given a score of around 10.	Again there is a risk of subjectivity, but the more structured nature of this method should reduce this.
Survey of users	This approach assumes that users can be identified. This will often be the case for the main users of the norm, but there is a risk that the views of some, particularly unforeseen and secondary users, will not be taken into account.	
	Given the time and expense involved in contacting users of norms, the more structured, criteria based approach of method B would probably be more appropriate.	To assess all the criteria fully would require a very detailed questionnaire, therefore there will probably be a need to be selective, but taking care not to introduce too much
Peer review	It may be difficult to find suitably qualified peers who have not been involved in either creating the norm, or creating a “rival” norm, therefore total objectivity can not always be guaranteed.	
	This method is probably rather simplistic for a meaningful peer review.	

6. Conclusions and recommendations of the Sub-group

There are many different aspects to be considered when determining the quality of norms. Very little previous work has been done in this area, though certain parallels can be drawn with aspects of work on the quality of statistics. This report discusses definitions, roles and categories of norms, then goes on to define quality and explore how that definition can be applied to norms. Key features of this are the identification and grouping of the users of norms in order to be better able to determine their needs.

It is likely to be difficult in practice to give a definitive assessment of the quality of a norm. There is often little or no data available to allow a fully objective empirical assessment to be made. It is therefore necessary to base the assessment on opinions as to how well the norm satisfies the needs of its users. This makes any such assessment essentially subjective, but the use of a clear framework or template to form the assessment can introduce at least an element of objectivity. Information on the background and development of the norm can help to make the assessment more meaningful.

Two methods are proposed for assessing the quality of norms. It is intended that these methods should be seen as complementary. Method A is far from ideal, but has the advantages of being quick, cheap, and easy for non-specialists to understand. Method B is more thorough, and therefore more time consuming. By following the checklist of quality criteria it is possible to produce detailed reports on individual norms.

Some norms are clearly more important than others, and resources should be concentrated on assessing those that have the largest impact on users. In general, it can be expected that the cost of developing a norm will be roughly proportionate to its impact, (the inevitable exceptions should be treated as special cases). Therefore, a general rule that x% of the development costs of a norm are allocated for quality assessment could be used to allocate resources. (x should probably be in the range 2-5). This will help to determine the combination of method and approach to be used. There is considerable knowledge and experience amongst Eurostat staff concerning the development of norms, this should be fully exploited when assessing the quality of norms.

There is a need to pay attention to the two-way links between the quality of statistics and quality of norms. When preparing a report on the quality of a norm, existing quality reports covering statistics created using that norm should be taken into account, and vice-versa.

In summary, the main conclusions and recommendations are;

- The quality of a norm is likely to be difficult to assess objectively.
- The quality of individual norms should not be assessed in isolation, it is important to consider their role within wider frameworks of norms.
- The ability of a norm to satisfy the needs of its users is the best criterion by which to assess its quality. This is in line with the ISO 8402 definition of quality.
- The resources devoted to assessing the quality of a norm should be proportionate to the cost of developing that norm.
- In the future, quality assessment should be an integral part of the procedure of developing a norm.
- The methods described in this report should be used to assess the quality of a few key norms, the experience gained can then be used to refine the methods if necessary.
- There is a need for close co-ordination of the work on the quality of statistics and quality of norms. The role of the Eurostat Quality Manager is important in this respect.

7. Bibliography

There are few works directly concerning the quality of norms, but in addition to those referred to in the footnotes, the following contain at least some material of relevance.

Abrahamse, A. (1996) Harmony by regulation: on the impact of EU Regulations on international harmonization of business statistics. Proceedings of the 82nd DGINS Conference – Statistical Business Registers: Problems and Opportunities, Vienna, Austria, 1996. Eurostat.

Beekman, M. and Struijs, P. (1993) The Quality of Economic Concepts and Definitions. Statistical Journal of the United Nations ECE 10, 1-15. IOS Press.

Daney de Marcillac, L. (1994) Lessons learned in implementing NACE Rev.1 in the European Union. Proceedings of the Eighth International Roundtable on Business Survey Frames, Heerlen, The Netherlands, May 1994.

Defays, D. (1995) Is harmonisation possible? Paper prepared for the conference on methodological issues in official statistics in Stockholm, Sweden, June 1995.

Depoutot, R. Quality definition and evaluation. Paper presented at the Fourth Seminar on the Future of European Social Statistics / Harmonisation on Social Statistics and Quality, 26-27 March 1998. This paper contains many further references that may be of interest.

Rainer, N. (1995) The Revised System of International Classifications. Eurostat.

8. Members of the Sub-group

Mr Tapio LEPPÖ	(Chairman)
Mr Alfonso ARPAIA	(Replaced by Mr Paolo CARIDI)
Mr Fausto CARDOSO	
Mr Louis DANÉY DE MARCILLAC	
Mr Raoul DEPOUTOT	
Mr Marcel ERNENS	(Replaced by Mr Paolo CARIDI)
Mr Svein GAASEMYR	
Mr Peter STRUIJS	
Mr Steven VALE	(Report compiler)

Annex 1 - Worked example of method A using NACE Rev.1⁷

Please note that this section should be seen only as an illustration of a method, and is not meant to be a definitive evaluation of the quality of this norm.

Background details - context and degree of control

NACE Rev.1 is a classification system for economic activity. Its use is compulsory within the European Union for business statistics. It is used both for collection of data and for statistical output. It has been adapted from ISIC Rev.3, a classification system drawn up by the United Nations, and recommended for use throughout the world. NACE Rev.1 contains more detail on the areas of the economy particularly relevant to Member States, though can be easily aggregated to ISIC Rev.3. Therefore Eurostat can be said to have partial control over NACE Rev.1.

Users of the norm

- Data suppliers - in a national context these often tend to be enterprises, who usually have little direct contact with NACE Rev.1 in this role.
- Producers of statistics - this group use NACE Rev.1 to classify businesses, and group those with similar activities. This is vital for the production of aggregate data. This group includes statisticians at national, Eurostat and world levels.
- Users of statistics - this group tends to require aggregate data, though often in some detail, for certain parts of the economy. Their needs include all of the categories mentioned in section 4.2.

How well are needs met?

- Data Suppliers - the needs of this group are for a logical, practical, and simple system which will minimise the burden on them. It could be said that NACE Rev.1 meets perhaps 85% of the needs of this group.
- Producers of statistics - the needs of this group are for coherence between domains, universality, stability over time, and, more at the national level clarity, collectability, flexibility and relevance, as codes often need to be derived from descriptions of activities supplied by businesses. NACE Rev.1 provides some coherence with other economic data in that it is linked to other classification systems e.g. CPC and CPA for products, HS and CN for international trade. It is used throughout the EU, and is coherent with ISIC Rev.3, which is used worldwide. It is an important and well-integrated part of the wider system of business statistics, though is also used in other domains. It is reasonably stable over time in the short and medium term, though economic development means that further revisions are likely to be needed in the longer term. It is a comprehensive system covering all areas of the economy, and at a national level the extra flexibility of further breakdowns are allowed as long as they aggregate to NACE Rev.1 classes. It is reasonably clear and easily accessible to users who need to allocate codes based on descriptions. Various computer packages are now available which can do this with a high level of accuracy. Therefore it seems that NACE Rev.1 meets about 90% of the needs of this group.

⁷ For more information about this norm, please see "NACE Rev.1: Statistical classification of economic activities in the European Community", Eurostat 1996.

- Users of statistics - because of the wide range of possible uses it is difficult to generalise but relevance, coherence, universality and stability are likely to be important. NACE Rev.1 seems to meet these needs reasonably well, though sometimes there is a demand, particularly from the private sector, for more detailed or differently structured breakdowns of data than those provided for by this system⁸. Therefore, perhaps about 65% of needs are met.

Overall Score

Using the scoring system in section 5.3, NACE Rev.1 would score 4, 4 and 3, giving an overall score of 11.

Past and Future Development

NACE Rev.1 was developed as part of an integrated system of economic classifications. It is also consistent with other international activity classifications such as ISIC Rev.3. It was created after long and detailed negotiations between experts from all Member States of the EU, and is therefore to some extent a compromise. It has many and widespread applications both within and outside the realm of statistics. It may not always be the optimum classification system for these applications, but is often used simply because it has already been developed, and is widely available and recognised.

Economic and technological developments mean that any activity classification will need revising from time to time. This has to be balanced by the need for consistency over time in order to create meaningful time series statistics. Due to the wide range of users of statistics there will always be pressure for different definitions and groupings of activities, but it is impossible to satisfy everyone. There does not seem to be a pressing need to revise NACE Rev.1 at present, though it would be appropriate to start planning now for a revision perhaps in five to ten years time.

⁸ "Alternative Aggregations and the Standard Industrial Classification" Shaila Nijhowne, Statistics Canada for the 12th meeting of the Voorburg Group on Service Statistics, Copenhagen, 15-19 September 1997. See also the Commission Recommendation of 8 February 1996 on a common aggregation for the purpose of making economic analyses, OJ No L38, 16.2.1996.

Annex 2 Detailed checklist of quality criteria for method B

1. RELEVANCE

- 1.1 Who are the users of the norm?
- 1.2 What are the needs of the users of the norm?
- 1.3 Is the norm relevant to these needs?
- 1.4 Is the norm generally accepted?
- 1.5 Is the norm used for non-statistical purposes?
- 1.6 If so, is the norm suitable for those purposes?

2. CLARITY, SPECIFICITY AND ACCESSIBILITY

2.1 Clarity

- 2.1.1 Is the norm specified in a way that is clear and easily understandable for all users?
- 2.1.2 Is technical jargon kept to the minimum level necessary?
- 2.1.3 Is the reason for the norm explained clearly?
- 2.1.4 Is the logic behind the norm explained clearly?
- 2.1.5 Are guidance or methodological notes provided to users of the norm?
- 2.1.6 Are these notes clear and comprehensive?

2.2 Specificity

- 2.2.1 Has accurate and recognised terminology been used in the specification of the norm?
- 2.2.2 Does the specification of the norm follow existing conventions?
- 2.2.3 Are references to other norms or legislation factually correct?
- 2.2.4 Is the norm translated into all official languages?
- 2.2.5 Are translations of the norm technically correct?

2.3 Accessibility

- 2.3.1 How is the norm disseminated?
- 2.3.2 Is the norm readily available to users in a suitable format?
- 2.3.3 Are potential users made aware of the existence of the norm?

3. CONFORMITY

3.1 Conformity with the norm

- 3.1.1 Do all users apply the norm correctly?
- 3.1.2 Do all users apply the norm consistently?
- 3.1.3 If the norm is not applied correctly / consistently, what are the reasons for this?
- 3.1.4 If the norm is not applied correctly / consistently, what are the consequences of this?

3.2 Conformity of the norm with external requirements

- 3.2.1 Does the norm take into account national and international legal requirements?
- 3.2.2 Does the norm take into account data confidentiality requirements?
- 3.2.3 Does the norm meet ISO requirements? i.e.
 - Is it possible to implement the norm in practice?
 - Is it possible to check that the norm is applied correctly?

4. COLLECTABILITY AND FLEXIBILITY

4.1 Collectability

- 4.1.1 How easy is it to collect the data necessary to apply the norm?
- 4.1.2 What level of response burden does the norm impose on data suppliers?
- 4.1.3 Are the financial costs associated with using the norm considered to be reasonable?
- 4.1.4 Does the norm allow statistics to be produced within a reasonable time scale?

4.2 Flexibility

- 4.2.1 Is there any flexibility in the way in which the norm can be implemented?
- 4.2.2 If so, are there clearly defined limits to this flexibility?
- 4.2.3 Have any derogations been granted that postpone implementation of all or part of the norm?
- 4.2.4 If so, are these for a clearly specified time period?

5. UNIVERSALITY AND STABILITY

5.1 Universality

- 5.1.1 Is the norm used throughout;
 - a) the European Union?
 - b) the European Economic Area?
 - c) other groupings of countries?
 - d) the world?
- 5.1.2 Was the norm applied concurrently in the above areas?

5.2 Stability over time

- 5.2.1 When was the norm devised?
- 5.2.2 Has the usefulness of the norm changed over time?
- 5.2.3 Has the range of application of the norm changed over time?
- 5.2.4 Has the norm been revised since its introduction?
- 5.2.5 If so, is it possible to create links to previous versions?
- 5.2.6 What effects have revisions had on the production of time-series statistics?

6. COHERENCE

- 6.1 In what domains of statistics are the norm used?
- 6.2 Is the norm part of a larger system of norms (e.g. ESA)?
- 6.3 Is the norm compatible with those used in other statistical domains?
- 6.4 Is the norm compatible with those used for statistical purposes in other countries or international organisations?
- 6.5 Is the norm compatible with non-statistical norms (e.g. accounting standards)?

7. EXHAUSTIVENESS

7.1 Exhaustiveness of a norm

- 7.1.1 All Norms - Does the norm cover all necessary features of what is being described or measured?
- 7.1.2 Classification systems - Does the classification system cover all possibilities, with a single unique category for each item to be classified?

7.2 Exhaustiveness of systems of norms

- 7.2.1 Are there any gaps in the systems of norms of which this norm is part?

Annex 3 - Sample quality report using method B - NACE Rev.1

Please note that this section should be seen only as an illustration of the type of report that could be produced using this method. It is not meant to be a definitive evaluation of the quality of this norm. It focuses mainly on the mathematical part of the quality assessment, and is therefore not as comprehensive as a full report compiled under this method should be.

Background details - context and degree of control

NACE Rev.1 is a classification system for economic activity. Its use is compulsory within the European Union for business statistics. It is used both for collection of data and for statistical output. It has been adapted from ISIC Rev.3, a classification system drawn up by the United Nations, and recommended for use throughout the world. NACE Rev.1 contains more detail on the areas of the economy particularly relevant to Member States, though can be easily aggregated to ISIC Rev.3. Therefore Eurostat can be said to have partial control over NACE Rev.1.

Users of NACE Rev.1

- Data suppliers - in a national context these often tend to be enterprises, who usually have little direct contact with NACE Rev.1 in this role. Often they are asked merely to provide a description of their activities. They are still users of the norm however, because the descriptions they provide have to be coded, either manually or electronically. The ease with which the classification system copes with the descriptions and terminology used by enterprises is therefore proportional to the accuracy of the coding. If a description does not readily fit into a single category then either the enterprise has to be re-contacted, increasing its response burden, or there is a risk it will be wrongly coded, and therefore may receive inappropriate surveys in future. Clarity, specificity and collectability are therefore likely to be important factors for enterprises.

In a multi-national context (e.g. within the ESS) data suppliers often tend to be National Statistical Institutes or other national or international organisations. In the role of data suppliers, these users of NACE Rev.1, are likely to place most emphasis on relevance and universality, because they need a standard framework, used and understood by all, in which to supply their data.

- Producers of statistics - this group includes statisticians at national, Eurostat and world levels, who use NACE Rev.1 to classify businesses, and group those with similar activities. This is vital for the production of aggregate data, and requires that the classification system and guidance notes are clear, accurate and easily accessible. Their job is made easier if the data necessary to code enterprises are easy to collect and verify. Coherence between domains and stability over time can both help, as they reduce the number of different systems that have to be learnt.

- Users of statistics - this group tends to require aggregate data, though often in some detail, for certain parts of the economy. The six groups of users of statistics identified by the Eurostat Survey of Users Group, are all to some extent users of the norm NACE Rev.1. These groups are;

- (i) European Union institutions e.g. Commission, Parliament, Council
- (ii) Government institutions within Member States, local, regional and national, including National Statistical Institutions
- (iii) International organisations e.g. United Nations, OECD

The key needs for these users are universality, stability, and coherence.

- (iv) Bodies representing others e.g. trade federations, trade unions, the media
- (v) Business users e.g. enterprises
- (vi) Other users e.g. academics and the general public

Relevance, clarity and accessibility will be more important for these groups.

User group weights

NACE Rev.1 is of considerable importance for producers and users of statistics, but less important for data suppliers. Therefore the user group weights allocated are;

Data Suppliers	2,5
Producers of Statistics	6,5
Users of Statistics	6,0
Σw_g	15,0

Assessment of quality for each user group

This should be based, wherever possible, on information supplied by users of the norm, e.g. through a survey of users.

1 Data Suppliers

Relevance – This is very important to data suppliers who are asked to code economic activities. The nomenclature used must be logical, and reflect the natural sub-divisions in economic activity as perceived by the data suppliers. If this is not the case, then the burden on data suppliers increases as they need longer to understand the nomenclature and correctly classify an activity. NACE Rev.1 generally satisfies these needs, though in common with other nomenclatures, can present problems when classifying enterprises with mixed activities. Therefore weight = 25 and score = 21.

Clarity, specificity and accessibility -

The weights and scores for each group of users should be summarised in tabular form, e.g.;

	Weight	Score
Relevance	25	21
Clarity, specificity and accessibility	5	3
Conformity	10	8
Collectability and flexibility	30	25
Universality and stability	10	7
Coherence	5	4
Exhaustiveness	15	12
	100	80

Overall Score

If we assume that the total scores for producers and users of statistics are 85 and 70 respectively, the overall score is;

$$(2,5*80) + (6,5*85) + (6*70) = 1172.5$$

Background Information

1. Description of the procedure followed in order to create NACE Rev.1, including the policy background, methodological framework, and constraints such as the need for consistency with related norms.

For NACE Rev.1, the desire for a certain degree of consistency with other international activity classifications such as ISIC Rev.3, and also the development of a system of harmonised European classifications are important factors to note. The advantages in terms of enhanced coherence and comparability, can be compared to the disadvantages the classification being less than optimal due to these constraints.

Work on the creation of NACE Rev.1 started in 1986. A committee of experts from Eurostat and the Member States of the European Union, was established. This committee used ISIC Rev.3 (which Eurostat was also heavily involved in developing) as a starting point, and adapted it to European needs. As a result, NACE Rev.1 is consistent with ISIC Rev.3 at higher levels, and can be aggregated to maintain this consistency at lower levels, whilst allowing a more detailed breakdown for European purposes where required. Council Regulation 3037/90 of 9 October 1990, made it compulsory for Member States of the EU to use NACE Rev.1 from 1st January 1993, when compiling and transmitting data relating to economic activity.

2. Relevant information on the implementation and subsequent development of NACE Rev.1.

NACE Rev.1 has now been implemented in all Member States, and several other European countries. This was made easier by allowing Member States the flexibility to add an optional 5th digit to the standard 4 digit code, where this was considered necessary for national purposes. Publications containing information about NACE Rev.1, and guidance notes for its use have been published by Eurostat following the agreement of representatives of the Member States. These representatives meet regularly in the form of a sub-committee of the Statistical Programming Committee (SPC) referred to as SPC (NACE). Issues relating to the development of NACE Rev.1, including rulings on case law are agreed at these meetings.

3. How NACE Rev.1 interacts with other norms, and how it fits into larger frameworks of norms.

NACE Rev.1 is part of a coherent system of economic classifications. It has close links to product classifications such as the CPA and the Prodcom List, and also the Combined Nomenclature, which is used to classify goods for trade purposes. NACE Rev.1 is also an important part of larger frameworks of norms such as the ESA, and is used in many statistical domains, e.g. business, social and regional.

4. References to any similar norms that might be of higher quality.

Comparisons can be made between NACE Rev.1 and NAICS, the economic activity classification system used in North America. The main difference between the two systems is that NAICS was developed using methodology created specifically for that purpose, and without any reference to existing international classifications, and may therefore be a more technically optimal system, but it lacks the level of international comparability that NACE Rev.1 allows.