European Conference on Quality in Official Statistics (Q2016)
Madrid, 31 May-3 June 2016

The Second Round of Peer Reviews – Experience of (Some Experts at) Statistics Finland[1]

Kari Djerf[1], Sirkku Mertanen[2]

[1] *Statistics Finland, Helsinki, Finland; Kari.Djerf@stat.fi*
[2] *Statistics Finland, Helsinki, Finland; Sirkku.Mertanen@stat.fi*

Abstract
The second Peer Review exercise was conducted in 2014-215. We will evaluate the process, the self-assessment forms as well as the organising it both at national and European level.
Finally, some further ideas will be given on how to make the whole self-assessment and peer review process more consolidated before the next round.

Keywords: ESS Code of Practice, Self-Assessment, Peer Review, Auditing

1. Introduction

The self-assessment is the key element of making the whole process work well as it should lay the basis for evaluation. This time it was much better organized than in the first round, 2006-2007. Still there were parts in the main questionnaire not completely thought and evaluated in advance which caused some confusion. Much of the confusing elements were born from the Quality Assurance Framework. We will comment some of the points. The actual filling process was known to be time consuming. The process will be described and some evaluation

[1] The authors were active participants of the whole Code of Practice peer review process 2013-2015. The opinions presented are their own and do not necessarily reflect those of Statistics Finland.

of its success and problems will be presented. There will also be an assessment how well the results matched the effort.

Statistics Finland was the first organization to be evaluated in the peer review phase after the pilot reviews and, thus, some practicalities may have not been fully operational at the time. Based on our experience we will comment the role of the consulting organisation, selection of the pool of peers, the actual peer review process and reporting. Similarly the domestic experience will be provided, especially what was learnt, how the improvement actions look like and how the organisation can and will use those in the future work.

We also discuss how to improve the whole self-assessment and peer review process more consolidated before the next round. Finally, some thoughts of the future development of the Code of Practice will be presented.

2. Organising for the second peer review round at Statistics Finland

A specific preparatory committee was set up in June 2013. The aim was that the group prepares and coordinates the whole process starting from the self-assessment and finishing to final evaluation. Originally the group consisted of eight members but due to various reasons six members participated in it thru the whole time span. The final meeting took place in March, 2015.

Main tasks of the preparatory working committee were

1. Prepare and complete the self-assessment of Statistics Finland
2. Provide other national statistical authorities (ONAs) with support and co-operate with them during the ONA self-assessment filling process and collect self-assessments from the ONAs
3. Prepare timetable for the peer review and carry out technical arrangements
4. Collect the requested documentation, take care of the translations when necessary, and mail those in due time.
5. Agree with suitable experts to participate in the peer review sessions

6. Agree with representatives of various stake-holders and ONAs to participate in the peer review sessions
7. Prepare the information for the staff, stakeholders and ONAs on the peer review and take care on its publishing
8. Prepare a final report from the whole process.

The committee was chaired by the PR National Coordinator Ms. Sirkku Mertanen.

2.1. Self-assessment form (NSI)

The preparatory committee began its work in the summer to have proper time for self-assessment the deadline of which was originally in March 2014. A lot of time was devoted to fill in the self-assessment form as well as finding the correct references for each of the 300 questions. For the first we went back to check the background material used in the 2007 peer review and evaluated which of those are still valid, usable with updates, and those which should be completely replaced by new ones. The next effort was to compile a basic list of the experts who could help the committee to make the first version of SA. About 30 persons participated in the kick-off event where the full peer review process was described and the actual assessment work was discussed.

The actual filling process started in the autumn. The whole questionnaire was split into bunches of topics relating to each other, sometimes comprising with the principles but mostly the topics were combined from various principles. Small working groups consisting of 2-4 persons were formed and they made the first proposals for the preparatory committee. They also used other experts' knowledge when necessary.

The committee checked the whole first version and made some changes and corrections. However, it was felt that the reply was not good enough and the quality of replies varied too much. Therefore a new process round was set up, and the working groups were reformulated in order to give new insight to it. Some topics required even more filling-in rounds. Finally the replies with necessary documentation were sufficient enough, and the whole questionnaire was read thru by six experts who also checked the English language. Also the references were investigated and new versions were written where necessary.

After the filling phase the SWOT analysis was carried out for each principle. It was carried out in common seminars of preparatory committee, some experts and directors, two times. Finally the committee made the final checks and some editorial work before the proper language planning. The references were collected to the Statistics Finland's website to make accessing them easier. Almost all references were given as links, only about 10 documents were sent with the SA form.

About 70 experts in total were involved in this phase of the process. Total work load was more than 200 equivalent person days, i.e. almost one full-time person year. The self-assessment ended with filling the web template.

2.2. Self-assessment of the Other National Authorities

The Advisory Board of the Official Statistics of Finland was informed of the exercise in advance. They also agreed that all OSF organizations fill in the short ONA form (in Finnish) even though some of those are actually not ESS providers. Those forms were collected and evaluated. The results were given to the OSF.

Two big ONAs, the Information Centre of the Ministry of Agriculture and Forestry[2] and Finnish Customs were selected to be included in the peer review and their forms (in English) were sent via Eurostat's IPM system.

[2] After the Peer Review the institutions dealing with basic production were combined, including their statistical functions. The new organization is Natural Resources Institute Finland.

2.3. Integration and co-operation forms

Those specific forms were discussed and filled in at Statistics Finland. The forms were sent at the same time as the SA.

2.4. Peer review

Because the self-assessment phase was a rather heavy process the preparatory committee recommended the Statistics Finland directors that the actual peer review should be arranged as soon as possible, i.e. among the first countries. Main motivation was that all the experts involved in preparations would remember all issues discussed better the sooner the review mission took place. It was realized and Finland was the first country when the peer reviews were really started.

The peer review team requested some additional material and links before the visit, and the timetable was agreed in advance following the PR guidelines. The peer review was arranged in August 2014. The peers were knowledgeable and well acquainted with the self-assessments and Statistics Finland web contents. Also a Eurostat observer took part in the first days. The atmosphere was good and relaxed even though the work days were very long indeed including some ad hoc meetings and presentations. Altogether 58 different experts from Statistics and 33 stakeholder representatives (incl. ONAs) took part in the sessions.

Finally the PR team presented their main findings and put some improvement issues forward tentatively. They also requested some additional written explanation on certain details before the final report was drafted.

2.5. Final report

The draft final report was due in two weeks after the visit. However, Eurostat postponed the timing to have time to check the contents of 3-4 first reports. Obviously the reason was to make the reports as comparable as possible. The delay affected the timing of all subsequent tasks.

The preparatory committee checked the report and made proposals for changes. The improvement actions proposed by the PR team were discussed with the appropriate departments and a timetable was made for their subtasks and benchmarks. Also a divergent opinion was made for two improvement actions.

2.6. The use of resources

The peer review preparations, realisation and other tasks took about 300 equivalent person days. Altogether the time use from 2013 till early 2015 was 541 work days, about 2.5 person years.

2.7. Experiences from the whole process

The idea of involving a lot of expert as early as possible was agreed when the preparations were started. One of the reasons was the experience from the first peer review exercise. At that time everything had to be passed very quickly and, thus, less than ten persons took actually part on filling in self-assessments and preparation of the peer review visit. It was considered necessary then but it was felt that more people should be participating.

This time it was thought that the process can be used to broaden the knowledge of the CoP and quality management issues in general, to improve the information change from different statistical units, and possibly give some building blocks for improving coherence among the working processes. Some of the ideas were met, some other were not.

The experts were selected quite carefully from the "best scores" which also meant they were quite busy for most of the time. For the next time the selection process should be lighter, and perhaps the number of experts invited could be smaller than in this exercise.

The process to filling in the self-assessment took too much time and effort. Even though the selected experts and directors were in central role as knowledgeable professionals there were too many phases of checking, cross-checking, commenting and rewriting. That led ultimately to the situation that a handful of core experts and committee members made the final versions – and used substantial time in doing that.

However, the similar problems were not present during the peer review preparations. The task was more concrete, of course, but perhaps all participants learned something. At least the preparatory committee made clear plans and invited less persons, and the national coordinator had a tight grip on everything. Almost all the persons invited were involved in the self-assessment exercise so they already knew the issues. After the draft report was received the work was mostly normal "business as usual".

Generally we can state that the whole process at Statistics Finland was fruitful despite the heavy use of resources.

3. Organisation of the Peer Review exercise – European level

3.1. Overview

ESS Committee decided in early 2012 to go on with preparations of the second Peer Review wave. The initial planning was carried out by the Expert Group on Peer Reviews and after its plan was accepted in the ESS meeting later in 2012 a specific ESSC Task Force to Develop the Methodology of the Reviews. The TF acted from December 2012 until last autumn.

From the quality work point of view this type of multilayer organisation sounds slightly strange. The ESS Committee will naturally keep a close eye on the realization of such an important and burdensome process. However, the WG on quality contains professionals on quality work (mainly quality managers) from all member states, EEA countries and associate countries. In addition, there are often representatives from other important institutions, like the ECB and UNECE. The WG was informed on the peer review plans on very general terms in their meeting late 2012. In 2013 there was no WGQ meeting and in 2014 the exercise was already going on. Thus, the members were actually not able to express any comments (except those who happened to be members of the specific TF). We believe that one should use synergy, co-operate in these type of important issues and collect all kind of helpful comments and information. Therefore we believe that closer co-operation between various actors would help: provide update information, give valuable opinions and help all parties in realizing such a heavy process.

Technical arrangements were outsourced. Probably it was a necessity because those tasks require a substantial effort. From the statistical agency point of view the requirement of circulating mails and papers via the consulting agency was sometimes frustrating: there were technical problems, inaccuracy of details and arrangements etc. Ultimately many things were directly agreed with Eurostat or peer reviewers.

3.2. Self-assessment questionnaire(s)

The self-assessment questionnaire for the NSIs was completely new, and very comprehensive indeed: some 280 questions, kind of SWOT analysis in each principle, good practices etc. If all sub-questions were calculated the number would increase much over 500 questions. The questions were by and large based on Quality Assessment Framework contents. The evaluation was made in ordinal scale with three alternatives: not implemented, partly implemented, fully implemented. All answers must be accompanied with justification and links to relevant documentation.

Technically the new questionnaire was really much better than the one used in the first exercise. However, the choice between “not”, “partly” and “fully” is mostly a matter of choice. It is almost impossible to make a general statement based on 200 statistics or so. Their contents, methods, processes etc. may differ very much from each other so that the final decision is most often based on the opinion of the last person looking at the answers. Therefore suggestion for the next time: reduce the number of questions and sub-questions, and take examples of statistics which comply fully and those which do not. Then it might be easier to check the situation during the peer review. Also the reliance on the QAF could be thought anew: at least issues existing in many principles (e.g. use of administrative data) should be combined as much as possible.

A specific note must be made on the electronic form. It did not function as it should, and at least one lengthy copy-paste filling was in vain: the data was destroyed. The second trail was partly successful but even then the final version had to be restructured in Eurostat and verified by the National Coordinator.

Other questionnaires were much shorter and easier to answer. Therefore the need for changes in their contents should be based on the analysis of results from the round 2 and emerging needs when the next exercise will be planned.

3.3. Peer reviews

The reviewer team were sought by the consultation agency but everyone knows that other networks were used to find suitable candidates. The final composition of about 30 peers was accepted by Eurostat. They were trained and the composition of teams were made using some experimental design so that the teams did not stay the same. There was a general agenda of the peer review structure and also a list of questions to be asked. Some guideline was obviously given that the team should find about 20 improvement actions to be listed. One ex-director general was appointed to oversee the PR reports and check their contents and quality.

Obviously the way of managing the peers was much better than in the first exercise: the quality of the reports are better now when it comes to their contents. Still there are issues relating to cultural differences which can be seen from the reports. Also the request of having about equal number of improvement actions leads to somewhat funny situations: in some cases the list of actions contain really difficult to realize changes in processes while there are lists containing many small scale actions.

A specific issue concerns those improvement actions which were felt to be either incorrect or not appropriate, e.g. asking for such changes in legislation which cannot be realized. The NSIs could give their divergent opinion for such actions. Some NSIs used the opportunity, altogether 39 out of 707 recommendations. Nearly half of the cases relate to professional independence. Without having the guidance for peers it is impossible to say whether they were given too strict orders on those issues.

A possible way to improve the peer review process even more is to compare the strategy with the professional external auditing, like the one of the ISO. It would require a common check-list to be made, obviously based on the self-assessment questionnaire. Then the basic contents would always be the same to make those issues comparable. But there can also be specific issues for each NSI because they are different.

4. The ESS Code of Practice

The Code of Practice was created in very special circumstances more than 10 years ago: the first clear misconducts in publishing economic and financial statistics were proved, and Eurostat faced accusations of wrongdoing as well as “serious irregularities” in higher administration. However, it is fair to say now that part of the crisis was politically motivated. As we know, the two-sided crisis was taken seriously and led to rapid creation of the rules which we now know as the CoP. Its aim is to foster independence, integrity and accountability of the national and Community statistical authorities.

The economic crisis has continued ever since and also new cases of serious problems with economic and financial statistics have emerged. Thus the need for the code is still valid. The quality assurance framework (QAF, 2009) tries to make it more concrete. The CoP was slightly edited in 2011 but the main contents was kept the same. However, some of the contents could be reviewed and the contents could be tied more clearly with some quality framework. It could be a good idea to create a new high-level working group to reconsider all aspects of quality, and work in close co-operation with the WG Quality.

5. Conclusions

The experience from the second Peer Review round was positive in general. The overall organisation was much better planned and informed in advance than in the first round. However, some parts could be improved both at European and national level before the next round.

The overall planning should be made in close co-operation with the WG Quality in order to keep all NSIs informed and involved in the process. The self-assessment questionnaire should be much shorter. It means its dependence on QAF might be cut, or the contents of the QAF should be made more concrete. The peer reviews could be tied more closely with real type of auditing and one of the peers should be trained auditor. Eurostat could reconsider the idea of outsourcing technical assistance. It might be much easier if all exchange of information and messages was centralized to one and only point which also knows and can decide on various technical issues. And finally, the technical instruments should be chosen carefully. The technical problems of the PR round no 2 show that it could be much easier to work with very common tools, just like MS Word documents. Finally, the contents and role of the Code might be reconsidered in the coming years.

At the national level a lot can be improved. The preparations were made too heavy: filling in the self-assessment forms need not involve as many people or as many checking rounds as we had. Thus some core team should be put in place and give them necessary time and power to collect the information needed. The experts called for the reviews should be chosen so that they represent all types of staff. They should also have proper possibilities to discuss with each other and prepare.