Quality Driven Data Collection. Towards a System for Quality Management Based on Dashboard Information

A. Sundvoll[1], G. Haraldsen[2], B. Lagerstrøm[3], P. Næs[4], N. Bjørnstad[5]

[1] *Statistics Norway; asu@ssb.no*
[2] *Statistics Norway; gha@ssb.no*
[3] *Statistics Norway; bol@ssb.no*
[4] *Statistics Norway; pns@ssb.no*
[5] *Statistics Norway; nin@ssb.no*

Abstract
This paper discusses development actions that need to be taken in order to move towards a system for quality management of data collection, based on digital dashboard information. The paper briefly discusses quality issues in enterprise and household surveys. A total survey error perspective is used as a basic framework to understand and measure quality. We also discuss the term survey costs. We describe our point of departure by referring to today's best practice in Statistics Norway, and present thoughts of where we are heading.

Keywords: Data collection management, quality indicators, dashboard

1. Introduction

In a situation of moving towards digital communication, the demand for cost-effectiveness increases. Our basic development idea is a digital dashboard system that displays the data collection indicators for all surveys and critical sub-processes of the data collection operations, and display quality measures as red, yellow and green traffic lights. Based on traffic light information, analysis of underlying information can be performed and necessary actions can be taken. We aim at an information system which identifies where actions should be taken during ongoing data collections, and which offer tools for actions. The following features will be essential:

- The dashboard should cover a large number of ongoing data collections and data collection processes.
- Statistical process control should be used to pinpoint where or when actions are needed.
- The input to the system should be based on the Total Survey Error approach.
- Estimates of cost efficiency defined as quality over cost should include both internal and external costs.
- Continuous updating tailored for responsive design should be combined with opportunities for in depth analyses.

The process data collected today will constitute our starting point. Still, there is a need for more systematic and automatic collection of process data. Our goal is to perform direct data collection in a more adaptive, responsive manner, within the budget provided for data collection. Necessary action should be taken during the data collection process, based on online process data and quality information in order to enhance quality and cost-effectiveness in direct data collection.

An important part of Statistics Norway's current business strategy is to streamline all parts of statistical production. In order to manage the data collection process in a *cost-effective manner*, we need to display and analyze coordinated process indicators that measure data quality, *during* the data collection process. In 2016 Statistics Norway has started work on developing a dashboard system for data collection supervision and management, based on process data and quality indicators.

Statistics Norway wants to take necessary actions during the data collection phase for individual surveys based on *adaptive and responsive management principles*. To support this strategy, we need a more complete overview of information collected from the data collection production systems and to collate them with administrative data, like budget and accrued costs. Furthermore, a systematic overview of the total, current data collections is necessary in order to perform effective resource allocation. Hence, a dashboard system both needs to provide an overview and offer a direct link to survey specific information, on a detailed level.

The data collection department will use the new dashboard system in day-to-day data collection monitoring and ongoing quality management work. In addition, the data collection

dashboard will support communication between the data collector organization and internal sponsors.

2. Understanding the term cost-effective data collection

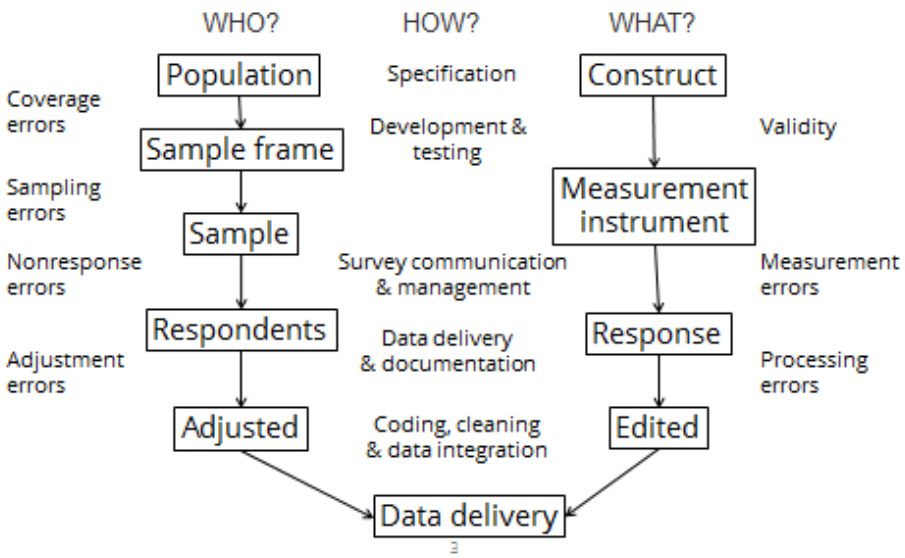
Currently, Statistics Norway is facing tougher demands concerning more cost-effective data collection and data processing. The organization is currently carrying out a structural re-organization in order to produce statistics in a more cost-effective manner, and in parallel to be able to intensify the institutional modernization program. There is an ongoing shift to digital communication in data collection. Information is distributed faster than in previous years, and feedback responses from the respondents are brought back to the survey organization more rapidly. Hence, the data collection – previously seen as phases where there were natural delays caused by shuffling of paper, are now considered as a continuous stream of communication back and forth to the respondents.

In short, one could explain the term cost-effectiveness by stating the equation quality over cost. The Total Survey Error (TES) model combines Groves et al's Survey Cycle Model (2004) with Biemer and Lybergs (2003) distinction between variable and systematic errors. Quality is defined as absence of errors. The TSE model is a conceptual model; the total survey errors have not yet been calculated. What is therefore important is to focus on, and measure major quality aspects. In household surveys bias because of nonresponse and measurement errors caused by difficult cognitive tasks are the most important errors. Business surveys are often mandatory and consequently have a high response rate. In these surveys, however, mismatch between questions and available information cause severe measurement challenges.

We use an expanded Survey Cycle Model suggested by Haraldsen et al (2015) which adds project planning and management steps to the traditional sample and measurement process (figure 1). This is useful because it links different kinds of errors to different phases in the data collection processes.

When it comes to the cost component in the cost efficiency formula, in data collections this component should include external cost. Actual response burden is used as an estimate of external cost.

Figure 1: A total error approach in surveys, including survey planning and management



How does Statistics Norway document and describe cost-effectiveness in production of statistics? One point of departure would be that the quality information is partly hidden within the production systems, and not always easy to detect and collect. A specification of key process indicators, and a system to display such information would improve the quality documentation. If we can start out by a conclusion that cost-effective data collection means enhancing quality in crucial process steps within the known framework conditions of money, time, available methods and technology, we have defined a point of departure for the work on visualization, analyzing and monitoring the data collection process.

3. Today's practice – use of indicators and action taken

The management system of Statistics Norway's Department of data collection and methods has over the last years been strongly influenced by lean management. This implies a change in how the organization reports problems and errors that occur in the data collection process, and how deviations are monitored and followed up. Lean management consists of lean operation

management, coordinated meeting agendas and actions, new tools for problem solving and a systematic feedback dialogue. A lot of data are collected and discussed on different levels in the organization. What is missing, however, is clearer priorities on what data to look at. A dashboard system should make cost efficiency actions more efficient.

A new data collection strategy was launched in November 2015, covering 2016-2018. The strategic initiatives are divided into annual action plans. The main message is cost-effectiveness - to gather and assemble data collection by modernization of methods and technology and combine data sources.

A number of indicators are reported annually and half yearly to the Ministry of Finance. These indicators concern survey response, fraction of electronic reporting, response burden in man-years and the distribution of response burden with respect to enterprise size. These results are documented in an annual and half annual report. Statistics Norway's board is quarterly updated on indicators regarding survey response and electronic reporting. On the department level Statistics Norway produces an annual activity plan containing specific expectations and goals concerning operating business for data collection and goals for development. The goals are broken down into tertiary actions, and visualized on a physical board used for lean-management. Along with tertiary actions, each unit report weekly challenges and errors in data collection production and the focus areas are pin-pointed for targeted follow up. Every second week, findings in an error reporting list is discussed and necessary actions are taken. On the department level, strategic issues are regularly discussed along with analysis of crucial process indicators.

On the unit level, there are meetings corresponding to the weekly meeting on department level. For business surveys process indicators on distribution of information and questionnaires to survey units are supervised, progress indicators are monitored, standardization issues are discussed, action to address queues in critical production systems are discussed, and so forth. The general system for sample unit administration contains process indicators on sample distribution and size, reminders, responses, number of fines and the handling of complaints regarding fines. The system for sample unit administration,

together with the system of administration of interview surveys forms a rich data source for survey monitoring and analysis. Data on costs are provided from other systems. The division for respondent services logs a lot of information concerning each data collection: Number of inquiries, how inquiries are managed and followed up, time used per case, distribution of different categories, applications for exemptions regarding fines, postponing of closing date for data collection etc.

Regarding household surveys, we partly follow the same pattern of management, meetings and main structure of monitoring as for business surveys. However, the information is drawn from different production systems than those for business surveys and attention is drawn more strongly to control of costs. Hence, one main challenge is harmonizing data collection management between household- and enterprise surveys. Actions that seem to be working well across all surveys are for instance the error report list, concerning process errors detected during data collection. Over the past two years, careful follow up and root cause analysis have been effective. The number of critical errors in data collection has decreased. An important issue is the discussion concerning root cause and shared solutions.

Another success factor is quarterly reports on population registers that are made available for users. The report lists suspicious units and unlikely/illogical data. Tailor made reports are made and distributed to municipalities and regional tax authorities, based on the Eurostat Blue-ETS work. Register information constitute an important basis for data collection.

Over the past years we have streamlined the helpdesk services for respondents. The respondent service unit produces timely and relevant factual analyses, based on logged and shared information. Feedback reporting to data collection and subject matter units is an important part of the ongoing quality improvement work. The unit has intensified contact with enterprises that pay fines instead of returning questionnaires. Also, major language improvements have been made in letters, material e.g. used in regular surveys. Nevertheless, a transition to digital communication with respondents requires new ways of communicating and thereby changes of procedures concerning data collection.

The division for household surveys has over the past years achieved major improvements regarding efficiency. Costs are reduced by making use of the national register of contact information. Also, the quality of contact information has increased, which has impact on the data collection. Administrative process data are linked with survey data, in order to control costs and optimize use of resources. In addition, the interviewers are provided with more register information than earlier, which enables them to monitor and tailor their contact attempts. Coordinated activity plans yield effective data collection processes, with elements of adaptive design.

4. Dashboard information – process data and quality indicators

The paper is titled "Towards a system for quality management based on dashboard information". This means that we should form a system focusing on enhancing quality – by minimizing errors and extra costs within general institutional framework conditions and survey specific constraints. In order to form a tool for data collection monitoring – we asked ourselves five research questions:

- What kind of data are we interested in?
- Who is the data about?
- Where can the information be obtained?
- When is the data available?
- How should we design the data collection?

Regarding the first question, we are interested in process data that can be collected automatically and are fairly easy to collect. The data are either related to the survey units or the survey variables. Generally, measurement errors are both the most important and the most difficult to detect.

The data should be obtained from internal and external production systems connected to a data collection chain. Furthermore, the data should be available in real time, in order to act in a responsive and adaptive way during a specific data collection.

The data collection is designed as generic data program script that makes it possible to view and cascade the data visually as the data collection proceeds.

Paradata, or survey process data are essential for managing the data collection process. An active fieldwork management approach is dependent upon monitoring core processes using extracted data. An active fieldwork approach consists of several steps; planning, monitoring, identifying problems and finding solutions, communicating and taking action, evaluation and documentation.

Currently, paradata are attracting attention. Snijkers et al (2013) gives an overview on possible paradata to collect. Based on this overview we made the following groups displayed in table 1 beneath, based on simple priority criteria such as whether the data were easy to extract from underlying production systems and most importantly, in order to address quality issues.

Table 1: Dashbord - data of interest

Priority Group 1: Response indicators
Survey response indicator
Non eligible indicator
Non response indicator
Non contacts
Refusals
Inability to cooperate
Other non-response causes
Mode distribution rate (different web-portals, interview, paper) Reminding rate Recontact rate Number/share of fines/fine rate (gross and net) Number/share of complaints

Priority Group 2: Sample distribution data
Representativeness indicator (sample in comparison to population distribution)
Mode response rate
Mode return rate
Mode change rate
Distribution of initial sample (divided into strata of interest)
Non eligible units pro strata
Distribution of net sample (progress/completion indicator)
Non-response pro strata

Priority Group 3: Process data/administrative data
Data reception problems Number of observations of problems regarding transfer of data/submit
Helpdesk data: Number of contacts, relative to sample size Distribution of questions in different categories
Internal costs (survey specific)
Internal costs (general)
Queue-data
Delays
Process errors

Priority Group 4: Survey quality indicators
Percieved response burden
Actual response burden
Response analysis: Item nonresponse Invalid response formats Logical errors Mathematical errors
Tailored paradata: Timestamps Correction made

In order to proceed further and to handle the transition to digital communication, we need a more sophisticated and digital dashboard for monitoring data collection and a foundation to obtain a responsive approach in data collection management. More paradata is required. We need to conduct the data collection from enterprises in a more cost-effective manner. Complete digital solutions are required. Progress indicators for data collection and response data need to be linked with accumulated costs. We need additional quality measures, and we need to match survey communication processes to when (time of year, month etc) respondents can provide available information.

5. Responsive design in household and business surveys

As already mentioned, the monitoring of business surveys and household surveys in Statistics Norway are carried out in different ways. Because household surveys normally are voluntary and carried out with interviewers, main focus areas in managing are traditionally on survey response and interview costs. Presently each survey is set up by an a pre-established best practice (adaptive), so basically not much are altered during the data collection process (responsive) other than prioritizing of interview resources and pinpointing subsets of sample units that need special attention in order to avoid bias. To some degree action plans also describes tolerance levels and preferred actions (responsive) (Lagerstrøm and Thomassen 2012).

Business surveys are different in that they normally are mandatory and run with self administrated web questionnaires. Survey management is pretty much governed by pre-determined, fixed dates (adaptive). Large enterprises are undertaken specific treatment, in order to avoid coverage problems (responsive).

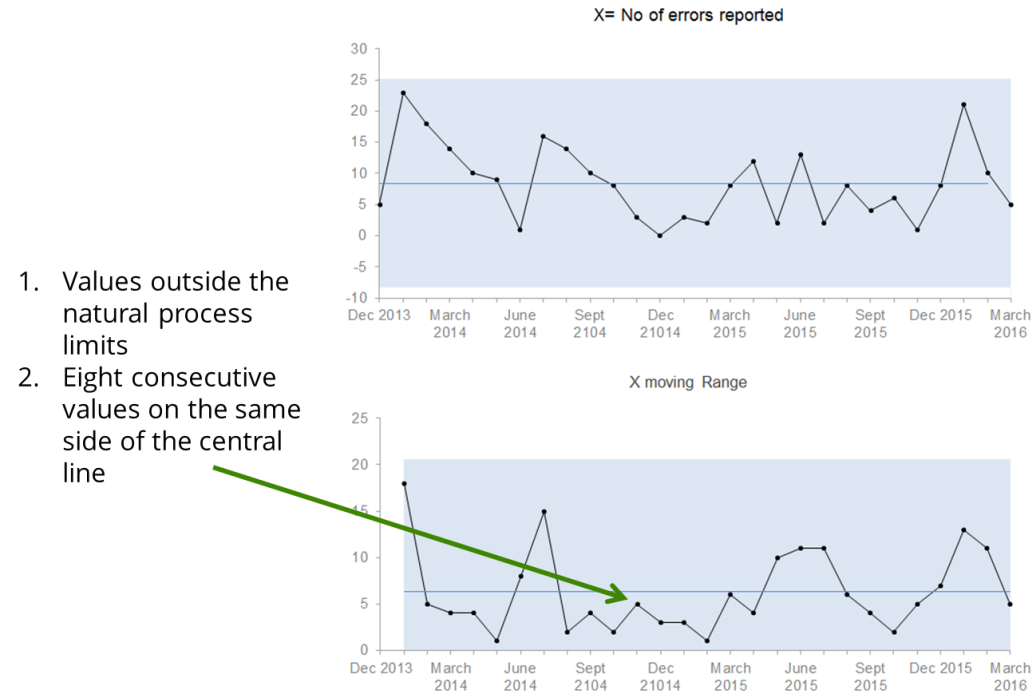
6. Towards a system for statistical data process control in quality management

Statistical process control dates back to the 1920s (Shehwart 1931). It is a method for quality control based on statistics and graphical process charts, and is commonly used to supervise manufacturing processes. Most variation is routine variation which does not call for any particular action. Statistical Process Control is about identifying results that are not routine.

Our idea is to combine this tool with quality indicators derived from the Total Survey Error, approach, and which can be continuously updated during data collection periods and from survey to survey. In this way the system should help us to focus on what is important and spare us from using time on noise and routine variation.

The key tool in Statistical Process Control is process charts. We use a XmR plot described by Stephen Few (Few 2015). The illustration given in figure 2 uses the number of errors reported to our help desk in the period from December 2013 to March 2016.

Figure 2: XmR Sensor Plot: Errors reported to help desk December 2013 to March 2016



XmR sensor plots are a combination of two charts; one chart showing how the quality indicators used change over time (or between data collections) and one which compare each measurement with the previous (moving Range) and in this way measure if what we observe is a part of a lasting trend or not. The shaded part of the graphs indicates the upper and lower limit for what should be considered routine variation (natural process limits). The horizontal central line shows the present average value. Values outside routine variation or eight consecutive values on the same side of the central line indicate incidents which call for attention.

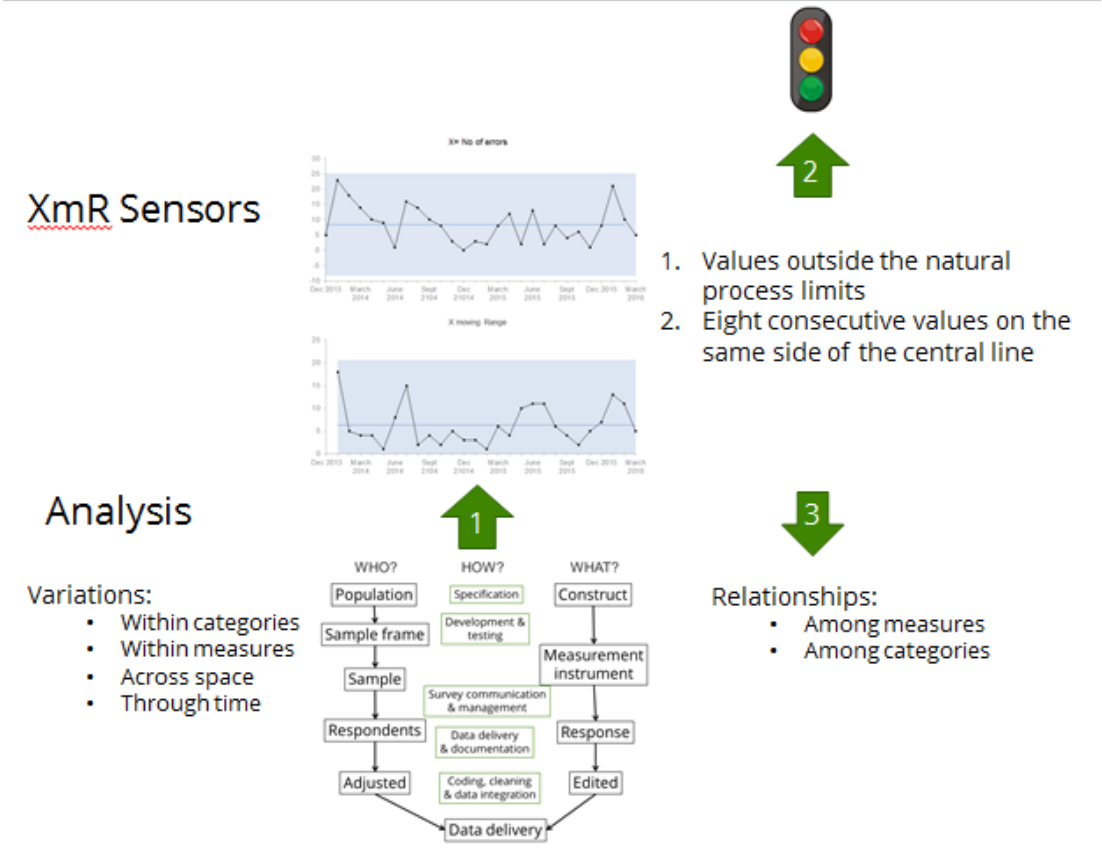
Incidents that are not routine my either indicate a problem or a success (e.g. either that the response rate is lower than what should be expected or higher). The x-graph gives us an answer to that question. The causes of non-routine incidences, however, are normally found in the details. Hence we need to drill down in detailed information to accomplish the necessary analysis of interest and to look for appropriate actions. We often have general knowledge from literature and previous practise about common causes of survey errors. Furthermore, as long as we know where we have a problem, the most common action is to spend more money. What

we should want, however, are more accurate and cost efficient actions. For this purpose analyses of former incidents should be fed into a knowledge bank that later incidents can profit from.

In figure 3 we have included the XmR chart in a model for quality driven data collection. Input to the XmR chart are quality indicators derived from the TSE approach (arrow 1). Status is given to the management in traffic lights which indicate if none (green), one (yellow) or two (red) of the criteria for actions are met (arrow 2). A manager dashboard would typically consist of several traffic lights reporting different kinds of quality indicators or indicators from different ongoing surveys. While the quality indicators chosen stem from descriptive statistics (measurements of variation), the basis for actions will be analyses of correlations (arrow 3).

In the example from errors reported to the help desk in figure 1, the period from August 2014 to March 2015 calls for attention (pointed at in figure 2). A closer look reveal that the majority of errors reported generally are about sample management, and that the reduction in errors reported we focus on coincides with a reduction in sample management problems. To better understand why this was so and if we can pick up ideas from this period which could lead to lasting improvements, we need to go back to the underlying reports and analyse in even more detail what happened.

Figure 3: Basic principles for data collection dashboard design



7. Where do we go from here?

From here we will follow two main paths. We will test the system for some ongoing surveys, probably some short time business and household surveys. In this test we will use well known and established quality indicators like nonresponse or the one used in the previous example. What we first of all want with this is to set up and test the technical and functional part of such an information system. The success factor will be to what extent the information flow can be automated or run by minimum human resources. A comparison between business and household surveys is interesting because what are the most important quality issues differ between these two kinds of data collections (see chapter 1 and 3 in Snijkers et al 2013).

Next we will use other surveys to exploit new sources of quality information. One of these surveys is a revised and modernized version of the questionnaire for the Structural Business Survey (SBS). This survey is interesting because it tries out a proactive approach to error prevention (Haraldsen, Hendriks and Holt 2016). It also includes some tailored paradata and improved questions about response burdens which has not yet been translated to quality indicators.

What ultimately should be our success criteria is that the quality of our surveys improves while what it costs to do so should go down.

References:

Biemer, P. P., and Lyberg, L. E. (2003), Introduction to Survey Quality, Wiley, Hoboken, NJ.Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., and Tourangeau, R. (2004), Survey Methodology, Wiley, Hoboken, NJ.

Haraldsen, G., Snijkers, G and Zhang, L.C. (2015) A Total Survey Error Approach to Business Surveys. International Total Survey Error Conference Baltimore, Maryland September 19-22 2015

Snijkers, G., Haraldsen, G., Jones, J. and Willimack, D.(2013) Designing and Conducting Business Surveys, Wiley, Hoboken

Shewhart, W A (1931) Economic Control of Quality of Manufactured Product. New York: D. Van Nostrand Company..

Few, S. (20115) Signal. Understanding what Matters in a World of Noise. Calefornia: Analytic Press

Lagerstrøm, B. O. and Thomassen, B. (2012) A systematic and strategic system for data collection at Statistics Norway. Geneva, Switzerland 31. October – 2. November 2012. United Nations and Economic Commission for Europe Conference of European Statisticians. Seminar of New Frontiers for Statistical Data Collection.