Implementing Quality Control Procedures at NASS's National Operations Center

Jeffrey M. Boone

National Agricultural Statistics Service, United States Department of Agriculture
3251 Old Lee Hwy., Rm. 305, Fairfax, VA 22030
jeff.boone@nass.usda.gov

Introduction

Quality control and assurance are important aspects of any agency. The International Organization for Standardization defines quality as "the degree to which a set of inherent characteristics fulfill requirements" (2005). Quality assurance involves ensuring that objectives or targets determined are met. Quality control involves the monitoring or evaluating a product or process to ensure that the desired standards are met. For the purpose of this article, the term "quality control" will be used, although both quality assurance and quality control are being examined. The United States Office of Management and Budget (OMB) Circular A-11 provides the following definitions:

> "***Quality assurance*** refers to a program for the systematic monitoring and evaluation of the various aspects of a project, service, or facility to ensure that standards of quality are being met. Quality is determined by the intended users and stakeholders. A ***Quality Assurance Plan*** establishes the goals, processes, and responsibilities required to implement effective quality assurance functions necessary to ensure a consistent approach to quality assurance throughout the lifecycle" (2011).

In this paper, the implementation of quality control procedures at the National Agricultural Statistics Service's (NASS) National Operations Center (NOC) will be discussed. Current progress and future objectives will be presented, along with a literature review of articles relating to quality control and its application to survey data collection.

An effective quality control program will ensure that a process will provide quality results over the entire span of the process. The Guide to the Project Management Body of Knowledge explains that meeting quality requirements has certain benefits, such as less rework, higher productivity, lower costs, and increased customer and stakeholder satisfaction (2008). In the context of survey methods, this corresponds to less call backs, less time spent editing surveys and researching unusual data, more record touches (more interviews), lower overall costs, and increased assurance of high-quality data. For example, according to de Santis and Carbini, the Italian National Statistical Institute used control charts to track quality indicators, such as those relating to timeliness and punctuality, as they improved their systems gradually over time (2010).

Moreover, OMB mandates good quality and the application of quality control programs. The OMB's Standards And Guidelines For Statistical Surveys document contains the following guideline:

> "**Guideline 2.3.4:** Develop protocols to monitor data collection activities, with strategies to correct identified problems. The following issues are important to consider:
>
> 1. Implement quality and performance measurement and process control systems to monitor data collection activities and integrate them into the data collection process. These processes, systems, and tools will provide timely measurement and reporting of all critical components of the data collection process, on the dimensions of progress, response, quality, and cost. Thus, managers will be able to identify and resolve problems and ensure that the data collection is completed successfully. Additionally, these measurements will provide survey designers and data users with indicators of survey performance and resultant data quality.

2. Use internal reporting systems that provide timely reporting of response rates and the reasons for nonresponse throughout the data collection. These systems should be flexible enough to identify important subgroups with low response rates for more intensive follow-ups" (2011).

The opportunity for cost savings is always desired, especially as of late. An example of cost savings from implementing a quality control plan was realized at the United States Postal Service. The agency-wide program was used to "drive increases in revenue, reduce costs, improve operations, detect fraud, and increase efficiencies in mail service", realizing "over $350 million in savings/cost avoidance" (MicroStrategy, 2011).

Apart from discovering cost savings, having a quality control program may determine the parts of the process that work well and need to be emphasized. This could lead to suggestions in future training and employee incentives and rewards. The system will also detect problems in the process, also leading to suggestions in more training as well as possible disciplinary actions or other fixes.

Maintaining a quality control program can also lead to the recognition of process improvement opportunities, such as implementing Lean Six Sigma (e.g. see George, 2003). This can lead to continuous improvement possibilities throughout the life of the program. Although process improvement can be applied at any stage of a process, the existence of metrics already being measured will enable a smooth and quick application of process improvement.

It is often agreed-upon to avoid subjective decision-making as much as possible due to its many problems, such as bias, lack of standardization, and the inability to explain the reasoning for a decision. A quality control program can eliminate subjective decision making by using objective measurements of a process. Thus, each decision made to alter a process can be data-driven. The program will also enable the understanding of natural or common-cause variation in a process, ensuring that decisions are not falsely based on random variation.

Quality control and related concepts have been discussed in many works in survey methodology. One such work discusses active management. Active management as a set of plans and tools used to effectively and seamlessly collect information on data collection processes to better improve survey administration practices. Active management refers to "monitoring progress, conducting timely analysis of indicators, identifying problems, implementing and communication corrective actions, and evaluating successes" (LaFlamme, Mayden, and Miller, 2008).

The term "paradata" is often used to describe information about the survey data collection process. This paradata may be used to compute survey data collection quality control metrics. Bates, Dalhammer, Phipps, Safir, and Tan (2010) provide some examples of paradata, including call records, observations of interviewers and respondents, audio recordings for interviewer and respondent interactions, items generated by computer-assisted instruments, such as response times and key strokes, etc. Paradata is used to measure survey quality in a production environment and to manage production with the goal of optimizing quality and minimizing costs. Paradata can be used for fieldwork monitoring, non-response analysis, responsive designs, aid in assessing measurement error, non-response adjustment, and to improve editing and coding.

Lepkowski, Axinn, Kirgis, Kruger, Mosher, and Groves (2010) describe four paradata paradigms: effort, active sample, productivity, and data set balance. For the National Survey of Family Growth, in reference to effort, several sources of data are measured: interviewers working, hours, percent productive, calls per day, calls per hour, percent peak calls, and screener versus main call. The authors also looked at percent occupied, percent eligible, percent nonworking, noncontacts, mean number of calls, percent of 8+ calls, percent locked buildings, percent resistant, percent hard appointment, and propensity. In productivity, the number of interviews, cumulative interviews, hours per interview, and calls per interview were measured. Many of these examples of paradata can be adapted to monitor the quality of the data and process. In the effort and productivity paradigms, these paradata may be used to create quality control metrics relating to the productivity of the interviewers but may also provide some insight on the quality of data. If data is being collected in a substandard manner, the data quality will suffer.

Lyberg (2009) explains that paradata can provide continuous updates of progress and stability checks, monitoring, input to long-run process improvement of product quality, and analysis for special and common cause variation, and input to methodological changes by finding and eliminating root cause problems. Paradata is also critical to

responsive survey designs and for providing input to organizational change. It can be used to understand variation and measure cost of poor quality and waste. He also expresses that paradata are multivariate in nature and may need to be combined to be relevant. Creating paradata archives allow reanalysis so that understanding of what is key can grow or change.

The concept of quality goes far beyond just the immediate data collection process. Lyberg describes three areas of quality: product quality, process quality, and organizational quality. Product quality involves product characteristics, data, controlled via quality framework, and other specifications set together with the user. Process quality involves process characteristics, paradata or key process variables, controlled via control charts or other tools for analysis of process variability. Organizational quality involves organizational characteristics, description of measures in place and performance, controlled via business excellence models or paradata for support processes.

Other types of paradata have also been studied. Stieger and Reips (2010) examined keystroke-type paradata, which may include the exact position at certain times; clicks with the mouse; double-clicks with the mouse; clicks on checkboxes, radio buttons, and list boxes; choices in drop-down menus; inserted text in text boxes; clicks on submit buttons; keys pressed on the keyboards; and the position of the mouse pointer every half second.

As can be seen, quality control, often under the guise of paradata, is a highly-discussed topic in recent survey literature. It is also evident that effective quality control is very important in providing high quality data in the survey data collection process. In the following sections, current methods of quality control as well as the progress towards creating a new quality control system at NASS are discussed.

NASS and the National Operations Center

The National Agricultural Statistics Service (NASS) provides timely, accurate, and useful statistics in service to U.S. agriculture. NASS conducts hundreds of surveys a year, conducts the Census of Agriculture every 5 years, provides data relating to America's agricultural products, and provides data used to determine commodity prices. Some examples of the data collected are that related to production and supplies of food and fiber, prices paid and received by farmers, farm labor and wages, farm finances, chemical use, and changes in the demographics of U.S. agricultural producers.

NASS currently performs its data collection in a decentralized environment from 6 regional Data Collection Centers and 46 Field Offices. The decentralized nature of our data collection and frame maintenance activities creates the possibility for increased non-sampling errors, contributing to an increase in total survey error. Over the last several years, NASS has attempted efficiency gains through regionalization of activities related to data collection and print/mail activities.

NASS began centralizing its data collection procedures with the opening of a National Operations Center (NOC) in St. Louis, MO, in August, 2011. The main objectives of this initiative are to reduce the source of error inherent to data collection activities, improve data quality, and reduce operational costs. The center will have various functions, including telephone interviewing, the processing of paper questionnaires, maintenance of NASS's list frame, training, survey development, Blaise programming, and web survey system programming.

Current Procedures

NASS currently has some quality control procedures in place. For telephone enumeration, these include the daily reports of response rates, post-survey reports of enumerator performance, post-survey cost reports, incentives for contract interviewers, enumerator monitoring, call backs, and enumerator performance evaluations. The enumerator monitoring, performed by supervisors, involves a paper form that is filled out by hand and filed in a local location. The information gathered from the call backs, such as verifying information obtained by the initial contact, the ensuring of proper enumerator protocol, and professionalism of the enumerator, is also recorded by hand on paper. The call center quality control procedures, as well as other procedures throughout NASS, that are currently in place are not well coordinated throughout all data collection facilities and other locations, nor is information needed for these control procedures in one place or available in a manner to observe the information across all data collection facilities.

For the processing of paper forms, many forms are hand edited. NASS does not currently track the number of edits or the time it takes for forms to be edited. Some forms, such as the Census of Agriculture and the Cash Rents Survey, are processed at the Bureau of the Census's National Processing Center (NPC) in Jeffersonville, IN. The NPC utilizes an electronic tracking system that provides numerous reports, including document tracking information such as location of a form and the number of forms at different steps of the process. NASS is currently in the process of developing a Tracking and Control System for forms processing activities at the NOC.

For all methods of data collection, NASS has an edit system that detects unusual reported values. These unusual data values are either automatically edited or brought to the attention of a staff member, who investigates and often modifies the reported values based on an investigation. NASS is also moving towards using other automated editing systems performing statistical edits of unusual values that do not necessarily require human intervention.

Quality Control at the National Operations Center

A central point of the creation of the NOC is for the improvement of data quality. As discussed above, an effective quality control program will ensure that the data collection process will provide quality results. An organized effort is necessary to make certain that the processes at the NOC are successful. With successful implementation of the program, NASS can continue to evolve from information learned in the data collection processes at the NOC.

Although NASS has some quality control procedures, as mentioned above, these may be ineffective in the NOC's environment. The call center at the NOC is very large compared to our existing data collection sites, presenting some potential problems. Monitoring evaluations are currently done on paper. With the large number of enumerators at the NOC, many monitoring reports will need to be created, yielding an inefficient monitoring process when done on paper forms. In other words, these informal methods NASS currently uses will not simply scale up to the amount of work at the NOC. In addition to the evaluations being recorded on paper, the information collected from call backs is also recorded on paper, which yields the same problem as the hand-filled monitoring reports. Also, a standardized system will assist in providing high quality data by ensuring consistency in the data collection process. Currently, NASS has different Data Collection Centers and Field Offices which perform many tasks slightly differently. Thus, data quality may vary based on the location of the calling. The area of the processing of paper forms also uses a tracking and control system that may surface problems with processing, such as bottlenecks in the process, similar to that currently used by the NPC. The large number of forms could not efficiently be hand-edited, as NASS does now for many of its forms. Effectively tracking all documents also allows more control over the process. With better control, errors and non-optimal processing areas will be more easily identified and used to improve data quality. To ensure the standardization of NASS's frames, the frames maintenance procedure will also need to be monitored for quality. With all maintenance activities being performed at the NOC, the quality control program can ensure that the frames are consistent and of high quality. With all the above examples, it can be seen that a quality control program at the NOC is necessary for successful implementation and operation.

It is also important to note that with many field offices, the opportunity for local knowledge of the agricultural products and practices is present. Different areas have different agricultural products, so local knowledge helps NASS understand possible errors in the reported data as it is being reported. However, once the data collection activities are centralized, the ability to apply local knowledge during data collection will be dramatically reduced. Thus, a quality control system must be in place to ensure that NASS provides the same level of data quality in the changing environment.

Progress

The steps to create a quality control system are straightforward. First, the metrics used to monitor the quality of the data collection process must be identified. Both productivity and data quality must be considered. Next, a method for capturing these metrics must be developed. This includes both developing an interface or system to record the necessary information for the calculation of these metrics and identifying a location for this information to be recorded. Finally, a method for displaying these metrics must be developed. This sort of "dashboard" will provide the users of the system with the information required to make data-driven decisions in a timely manner. This process seems simple, however, in some business environments, many difficulties arise. These issues will be discussed later in this article.

Metrics

The first step towards creating a quality control system is to identify the metrics that will be used to monitor the quality of the data collection process. Metrics should have the ability to answer a question to result in an objective, data-driven decision. When performing objective assignments, it is suggested that companies should "follow a set of principles to develop metrics specific to their needs" (Pipino, Lee, and Wang, 2002). The metrics should measure both the production of staff, such as telephone interviewers, as well as the quality of the data. These measures will provide a way to detect issues in the data collection process and with the quality of the collected data. If issues are detected, actions may then be taken to correct or minimize these problems.

When selecting productivity indicators, it is important to consider the following characteristics. The indicators must be easily understood, measureable and comparable at any point in data collection, consistently updated throughout data collection, and relevant, interpretable, and comparable at different levels of aggregation (LaFlamme, 2009).

Some call center metrics discussed in the literature include interview duration, interviews by time of day, interview result, and outliers, as well as the interview's language, cooperativeness of respondent, interview mode, reasons for interview break-offs, type of non-interview cases, and time of interview and module/section times (O'Reilly, 2010). Other metrics discussed are how interviewers chose to enter data (keyboard v. mouse), when and how they navigated in a complex instrument, how interviewers handled edit messages when they were triggered, and how more complex sections of the instrument (e.g. tables) were handled (Mockovak and Powers, 2008).

Numerous metrics have been identified for use at the NOC, and a list of those currently chosen for use by the operational group (call center, forms processing, and frames maintenance) and type of metric/variable (production, data quality, and global information) is seen below in Table 1. Note that some metrics can be considered both measures of productivity and data quality, such as the average length of call. For example, the length of the call gives information on how much work is being done and is thus a productivity measure. However, if the average length of calls yielding completes is unusually low, the quality of the data collected may be compromised. Also note that many of the items in the table are simply variables output by a system as opposed to tradition metrics. More comprehensive metrics may be derived by combining multiple items together, creating indices. Some indices for monitoring information are discussed later in this section.

Table 1: Identified Metrics.

	Call Center	**Forms Processing**	**Frames Maintenance**
Metrics Relating to Productivity	Samples Completed [B]	Forms Processed	Number of Frame Changes
	Usable Completes Per Hour Worked [B]	Time Between Events	Number of Change Requests from FOs
	Refusals (%Refusal) [B]	Processing Time	Time Until Change is Made
	Inaccessible (%Inaccessible) [B]	Keying Time	
	Total Cost (Based on Wage)	Cost Per Sample	
	Cost Per Sample		
	Average Length of Call [B]		
	Number of Dials Until Interview [B]		
	Type of Respondent		
	Punctuality and Availability of Enumerators		
	Information from Employee Skills and Availability Database		
Metrics Relating to Data Quality	Monitoring Information	Keying Errors	Coverage Measures
	Average Length of Call [B]	Number of Edits Required	Inaccessible (%Inaccessible)
	Number of Edits Required		Time to Make Change
	Callback Information		

Global Information Variables	Hours Worked	Hours Worked	Time Change Request is Submitted
	Shift (Day or Evening)		Time Change Request is Completed
	Day of Week		
	Date		
	Survey Name and ID		
	Enumerator Name and ID		
	Number of Disruptions in Data Collection		
	Reasons for Leaving		

[B] Call Center measures currently available in Blaise or can be calculated by Blaise variables alone.

Each variable/metric will be measured to answer a specific question in regards to quality. For each of these, a description will be given involving the question to which it is responding, how it is to be recorded or from where it will be drawn, how often it should be monitored, and who would need to track it.

Metrics relating to productivity, such as number of calls, length of calls, and number of refusals, as well as cost, are relatively simple to determine; however, those relating to data quality are more difficult to ascertain. More research is underway to discover these measures.

To record data regarding information collected from monitoring sessions, an Enumerator Evaluation Form has been developed. This instrument, constructed using Blaise, was created to replace the paper evaluation forms that are currently in use. This provides a method of capturing the data electronically, allowing users to quickly identify any recurring problems. This is a great improvement over the former method, where data was captured on paper, which was stored in a local location. That method was acceptable when there were only a few interviewers in one location, but with the creation of the NOC, the large number of monitoring sessions would create too many paper forms to adequately manage. The Form also allows higher-level users to view the data collection process across multiple data collection locations to understand performance beyond just that of a single location. The stored data also provides a means for statistical analysis of possible reasons for poor or exceptional performance. To view the information collected, a simple dashboard was created in Excel. Numerous data views and graphs enable the coach or supervisor to view information from individual monitoring sessions as well as across groups of interviewers (by coach or supervisor) and over time. This not only allows easy evaluation of an interviewer; the dashboard also allows higher-level users to examine any differences in rating methods by various coaches and supervisors, leading to a notification when calibration sessions should be held to ensure consistency across raters.

A common issue with the determination of metrics involves the number of metrics to monitor, as well as their importance relative to the question posed. Monitoring too many metrics may not only consume too much time and effort but may signal a problem when there actually is no problem. The more areas one looks for problems, the more false alarms one will find. Also, if the order of importance of these variables is not maintained, the monitor may focus on unnecessary measures while missing some important ones. A common method for alleviating this problem is to create an index combining information from multiple variables. This would enable the monitor to observe fewer variables without losing any information. For example, three indices have been created to summarize the information gathered from the Enumerator Evaluation Form. Each provides a single number to show overall performance in the introduction, body, and closing sections of the interview. For ease of use, the metrics are calculated as a percentage of the maximum obtainable score, leading to a simple 0%-100% scale.

Some metrics, such as the average length of a call may vary greatly by the survey being conducted since the expected completion times vary amongst surveys. Thus, it is erroneous to observe the metrics such as these over time across surveys without controlling for the expected value. If this can be assumed known, the metrics may be standardized so that they may be viewed across surveys. However, adequate data must be collected to determine the expected values for a specific survey. For very short survey periods, the data may not be available for the first run of the survey. Another problem is present when a very large number of surveys are completed, such as in NASS's case. It is very difficult to maintain a large number of these expected values to flow into the system and can be problematic to continue to update these values. These issues are currently being researched.

Another problem in the case of call center metrics is inherent in the difficulty of some cases. For example, some geographic regions have a naturally lower response rate than others. In this case, it would be erroneous to consider just the response rates of enumerators across all regions, since those with the more difficult cases will probably have lower response rates. Thus, studies have been done on creating a "difficulty" or "response propensity" index (see, e.g. LaFlamme and St-Jean, 2011). This would enable the coach or supervisor to note the difficulty of the cases when examining enumerator productivity. NASS is currently performing research to develop a response propensity index to classify cases that may have a high probability of non-response.

Storage

An important aspect of a quality control system is data storage. The physical space may sometimes be a problem, but beyond just the concept of size is the business problem of database maintenance. Database maintenance includes ensuring that items allowed to be stored in the database are chosen in a manner such that no information is replicated and every variable may be referenced in a standardized manner.

NASS is currently developing a new centralized database for the storage of work-in-progress data. Statistics Canada also uses a centralized database as a paradata warehouse to store all necessary information (O'Reilly, 2010). The new NASS database is being created in a very organized manner. Although it is a very good practice to create a database such as that described above, it creates a bottleneck for system development that relies on the database for storage and data access. This has been a very difficult obstacle experienced in the creation of the quality control system. Approval for every variable must be obtained, including the decision on a variable name and the table in which it will be located.

Beyond the issue of obtaining the actual storage space, a transfer program must be written or modified to move the data from the system from which it was gathered to the database where the data will be stored. This process is generally performed by an extract, transform, and load (ETL) process. The ETL process extracts data from some source (say, a CATI system), transforms the data to match the protocol of the destination database, and loads the data into the destination database. In the case of the quality control system, this has been written to extract data from the necessary origination systems. However, many of the variables required for the quality control system are not currently being transferred. Thus, the ETL must be modified to account for the transfer of the necessary data.

Another important aspect in using a newly-deployed centralized database is ensuring the data originating in different systems is in the same standard form when it is placed in the database. For example, consider the unique interviewer ID that will be used to identify a specific interviewer. This interviewer may have some data that is tracked by different systems. The data from these systems may be merged across the different systems to create the metrics that will be observed. If the systems are not using the same identifier for the interviewer, the data may not be merged properly, and the resulting metrics may not be properly computed, or may be unable to be computed. Note that it may not be the case, due to the architecture of the database, for all necessary information to move to that database. Thus, some data may need to be pulled from other databases by the quality control application.

Monitoring

Once the metrics have been determined, a monitoring method must be selected to obtain these metrics. For telephone enumeration, this includes recording the production measures, such as the length of a phone call or refusal rate, in a database. These measures can be obtained from many CATI systems, such as the Blaise system, so only a location of the data depository needs to be determined. Measures relating to enumerator performance that are not automatically captured, such as the ratings given by the monitoring staff, will need to be recorded. Once developed, the location of this data depository also needs to be determined. Other information, such as the skills, availability, attendance record, and time in job of each enumerator, will be tracked. This should also be automated in a system and the data made available in an accessible database.

The use of Computer-Assisted Recorded Interviewing (CARI) is also currently being researched. This system allows interviews or select segments of interviews to be recorded, allowing monitors to evaluate sessions after they have occurred instead of relying entirely on live interviewing. This can enable monitors to provide information on evaluations which may be very useful in explaining the results of a monitoring session to the enumerator. CARI can

be used to record pre-interview interactions so that you have a "truth" against which to compare quality indicators (Bates et al., 2010). NASS is in the early stage of adapting a CARI system for use in its data collection activities.

For the other functions of the NOC, metrics may be tracked by the computer software, using the systems developed for forms processing and the systems for frames maintenance. Data from these systems will also need to be made available in a database.

Once the data is in the database, a computer software application will be used to illustrate the data to the staff that are to use it. This dashboard will present the data as descriptive tables as well as graphs and charts. It is important to ensure that field managers have the ability to "drill down in trace files to examine details of interviewers" (O'Reilly, 2010). Many software applications have the capability to create dashboards. Currently, NASS is considering Microsoft SharePoint due to its present availability. Something as simple as an interactive web page may also be sufficient. The choice of the software will be determined by the availability of resources available, including current licenses, cost of new software, employee knowledge, and employee time resources. For effective live monitoring of the data collection process, the data must be available on an ongoing up-to-date basis. Many problems that may arise in a process must be detected quickly, so a batch process that runs only once or few times a day may not be sufficient.

An alternative method of quality control reporting involves management reports. O'Reilly provides some examples of management reports, including summary-level reports on Regional Office Totals, Cumulative/Weekly Report, and Average Cases Per Interviewer; and interviewer reports on Highest Non-response Rate for Salary of Reference Person; Highest Number of Regular Occupied Interviews Completed in Less than 20 Minutes; Highest Number of Cases Completed 12AM-7:59AM; Highest Vacancy Rates, and more (O'Reilly, 2010). However, for active monitoring of the data collection process, the data must be provided in a quick and ongoing manner, so a dashboard is highly recommended during the data collection period. Some reports, such as those mentioned above, may be useful for higher-level management and post-survey analysis.

Another portion of the quality control program will involve calibrating the monitors and other staff. This ensures that the staff is performing the same tasks in the same way. For example, a calibration session for call center monitors may be held to standardize ratings by having all monitors rate the same calls and discuss the results. A calibration session for frames maintenance may involve having all frames maintenance staff perform the same task and discussing the results. This is not included in the determination of metrics, but is imperative in ensuring standardization across data collection staff.

The staff whom are to use the system will also be trained on the different aspects of the system. This training should also vary based on their different roles, such as call center supervisor or forms processing group leader. Thus, a formal training plan will be developed to ensure proper use of the system. This training will also specifically describe some steps that must be taken when an issue is noticed, such as an interviewer having too high of a percentage of refusals. It is also important to note there that some staff may be reluctant to accept the new quality control procedures. This risk will be mitigated by ensuring that the staff understand the reasons for the new implementation. However, this is always a problem seen in the implementation of any new system or policy.

Issues

NASS is currently developing multiple new systems. Many of these systems are high in demand and are necessary in the development on the NOC. The quality control system is very important, but not often considered a requirement to begin operations. Thus, these specialized personnel that are required for the development of the system are utilized for these other "higher priority" systems.

As with most projects, the acquisition of resources is a significant impediment. Many projects require staff with very specialized functions and thus difficulty arises in acquiring these specific personnel. Creating this quality control system requires some specific functions that can only be performed by specialized personnel. These few staff are also high in demand. Thus, the resources needed for this project are difficult to obtain, slowing the development of the project. Unfortunately, some steps must be completed before subsequent steps can begin, i.e. some steps to complete the project lie on the critical path. For example, the tasks involving the database must be concluded so that the data may be available for the development of the system.

Data originating in different systems must be standardized before being put into the centralized database. Since many of NASS's systems were created by different groups (inside and outside of the agency) and at different times, the unique interviewer identifiers are not standardized. Thus, when the data is pulled into the centralized database, it must be ensured that the data is being merged properly through some sort of transformation or script to match the interviewers' data.

Not only is it often difficult to prepare the database for the storage of the data, the acquisition of the software to create the dashboard is a difficult process. The process is not as simple as choosing the software, purchasing it, and installing it; deployment of the software on a company- or agency-wide network is an immense task. Security and enterprise architecture personnel approval must be obtained in most cases, which may take a large amount of time. It is often best to utilize current services if possible. For example, Microsoft SharePoint is commonly used by many organizations and is a viable dashboard creation tool. Unfortunately, some organizations also have standardized procedures for the development of SharePoint sites, which may conflict with the desired functions of the dashboard. Thus, even if the software is installed and available, it may not be configured properly for this specific use.

Conclusion

A quality control system for data collection yields many benefits. Many of these benefits may be realized before the system is even in its final form. As the system is being developed, some issues in the data collection process may be noticed, providing an opportunity to smooth out some of these issues. Although the ultimate goal of the system is to ensure high quality data and proper control of the data collection process, other opportunities are present. For example, one of the first steps of process improvement is the collection of data about the process (paradata in the case of survey data collection). The quality control system collects this data, allowing this stage of the process to be complete, and therefore process improvement projects may be completed in a much more efficient manner. The system may also provide a means for data collection optimization across multiple sites, identifying areas where training may be needed, improving hiring procedures, survey development (seeing what works and what doesn't in a survey), and implementing responsive survey design.

Our path to creating a quality control system has been difficult, but progress has been made. Through diligence, we have been able to overcome many of the barriers we have encountered. There are many obstacles left to surmount, but since we have shown that it is possible to get past them, we are optimistic in completing the project. Obtaining the resources necessary to build the system is very difficult, but with patience, the system can be created.

So far in the development of the quality control system, we have identified metrics that will be tracked relating to both productivity and data quality, determined data storage requirements, developed monitoring methods involving an interviewer evaluation solution and data viewer, and investigated the use of dashboards. Future goals include creating the method of transferring the quality control data from the various systems into the centralized database, developing the dashboard for easy access to the necessary information, testing an early version of the system, and deploying the final product. Once these final tasks are accomplished, the comprehensive data collection quality control system will be available for production. Remember, however, that the process does not end here; once the system is in place, it should continue to grow into a more efficient and effective tool for quality control. Never let the system remain stagnant; it must grow to accommodate any changes in the overall data collection scheme.

References

Bates, N., Dalhammer, J., Phipps, P., Safir, A., and Tan, L. (2010). "Assessing contact history paradata quality across several Federal surveys". *Proceedings of the American Statistical Association 2010 Joint Statistical Meeting*. American Statistical Association.
https://www.amstat.org/sections/srms/Proceedings/y2010/Files/306005_55654.pdf.

De Santis, A., and Carbini, R. (2010). "Quality of quality: an analysis of quality indicators and their quality". *Proceedings of the European Conference on Quality in Official Statistics*, http://q2010.stat.fi/papersbig/.

George, M. (2003). *Lean Six Sigma for Service: How to Use Lean Speed and Six Sigma Quality to Improve Services and Transactions*. New York: McGraw-Hill.

International Organization for Standardization (2005). "ISO 9000". *Quality Management Systems-Fundamentals and Vocabulary*. Geneva: ISO Press.

LaFlamme, F., Mayden, M., and Miller, A. (2008). "Using paradata to actively manage data collection process". *Proceedings of the American Statistical Association 2008 Joint Statistical Meeting*. American Statistical Association. http://www.amstat.org/sections/srms/proceedings/y2008/Files/300608.pdf.

Laflamme, F. (2009). "Experiences in assessing, monitoring and controlling survey productivity and costs at Statistics Canada". *Proceedings of the 57th Session of the International Statistical Institute*. Retrieved from http://www.statssa.gov.za/isi2009/ScientificProgramme/IPMS/0049.pdf.

LaFlamme, F and St-Jean, H. (2011). "Proposed indicators to assess interviewer performance in CATI surveys." To be presented at the Joint Statistical Meetings 2011, Miami, FL.

Lepkowski, J. M., Axinn, W., Kirgis, N., Kruger, S. N., Mosher, W., and Groves, R. M.. (2010). "Use of paradata in a responsive design framework to manage a field data collection." NSFG Paper No. 10-012. http://www.psc.isr.umich.edu/pubs/pdf/ng10-012.pdf.

Lyberg, L. (2009). "The paradata concept in survey research". Presentation. Presented at NCRM Paradata Network, London, UK, August 24, 2009. Retrieved from http://www.natcen.ac.uk/ncrm-paradata-network/docs/Lyberg-paradata-concept.ppt.

Mockovak, W. and Powers, R. (2008). "The use of paradata for evaluating interviewer training and performance". Washington, DC: U.S. Bureau of Labor Statistics. Retrieved from http://www.bls.gov/osmr/abstract/st/st080180.htm.

"Office of Management and Budget Circular A-11". Retrieved on 2/11/2011 from http://www.whitehouse.gov/omb/circulars_a11_current_year_a11_toc/.

"OMB's Standards And Guidelines For Statistical Surveys". Retrieved on 2/11/2011 from http://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf.

O'Reilly, J. (2010). "Paradata and Blaise: a review of recent applications and research". *Proceedings of the 12th International Blaise Users Conference*. Retrieved from http://ibuc2009.blaiseusers.org/papers/7d.pdf.

Pipino, L. L., Lee, Y. W., and Wang, R. Y. (2002). "Data quality assessment". *Communication of the ACM, 45*(4), Retrieved from http://web.mit.edu/tdqm/www/tdqmpub/PipinoLeeWangCACMApr02.pdf.

Project Management Institute. 2008. *A Guide to the Project Management Body of Knowledge*. 4th ed. Newton Square, PA: Project Management Institute, Inc.

Stieger, S. and Reips, U. (2010) “What are participants doing while filling in an online questionnaire: A paradata collection tool and an empirical study”. Computers in Human Behavior, Vol. 26 (6). Amsterdam: Elsevier Science Publishers.

MicroStrategy. “U.S. Postal Service: MicroStrategy Customer Success Story”. Retrieved on 2/11/2011 from http://www.microstrategy.com/Customers/Successes/detail.asp?ID=259.