Using Real-Time Process Measures to Improve Data Collection

David A. Marker
Westat
DavidMarker@Westat.com

Abstract:

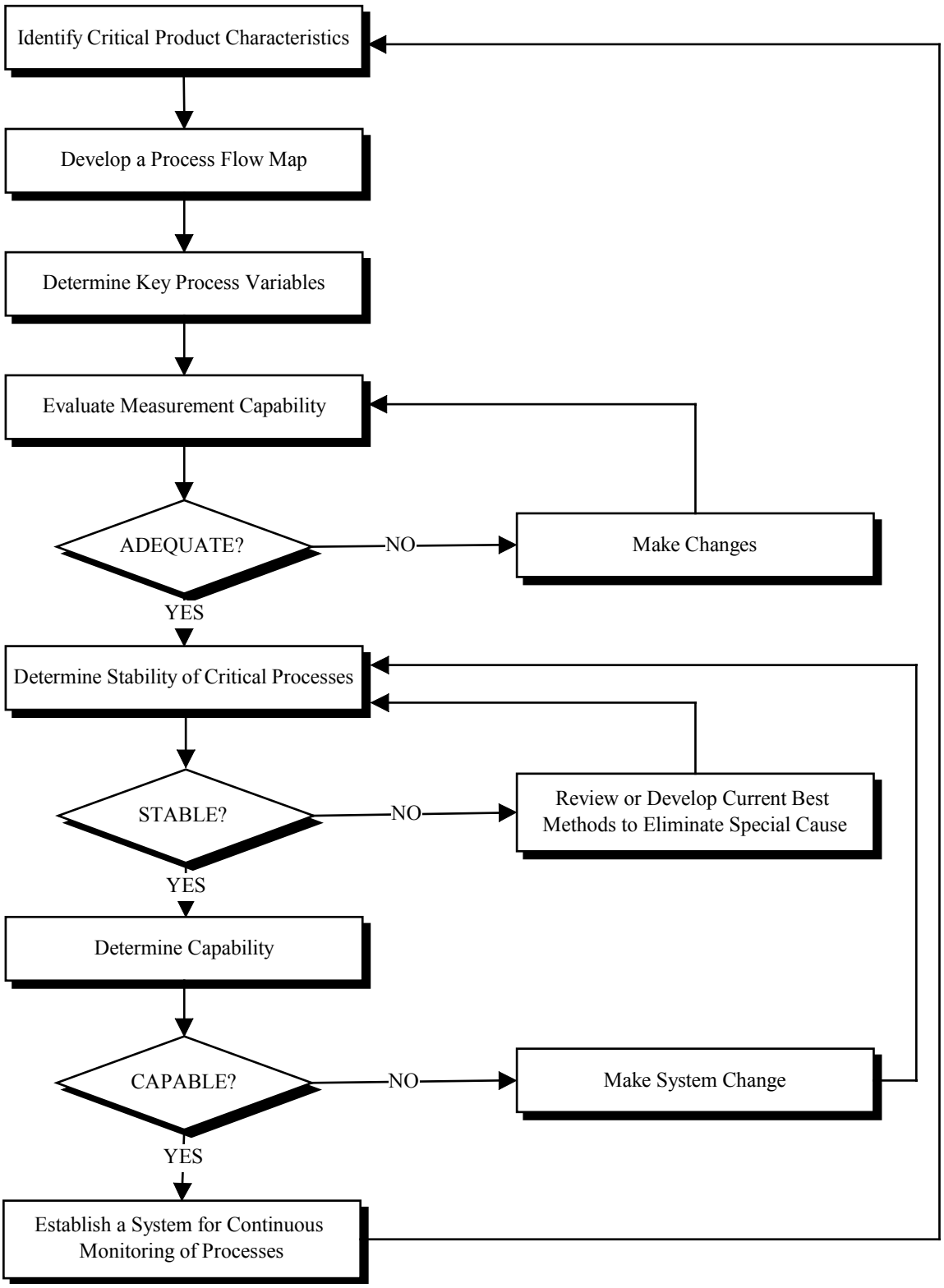
Many statistical organizations have come to realize that measuring process variables is key to improving quality. This involves identifying key process variables, measuring them, and acting on them in a timely fashion. The goal is not to analyse them after a survey is completed, but to get feedback on these variables when decisions could still be made that would improve the quality of the current round of data collection. The importance of process measures is seen by the current effort of the European Statistical System to develop a "Handbook on Improving Quality by Analysis of Process Variables" as part of the LEG Recommendation #3.

This presentation will provide examples from surveys that collected real-time process measures and used them to improve quality of the existing data collection. Examples will come from a variety of surveys and organizations, including a new longitudinal survey of businesses in the United States

1. Introduction

Many national statistical offices (NSOs) have been working to systematically improve the quality of their operations. A weakness in many of these efforts has been the lack of identifying and measuring process variables. Until this new emphasis on quality, no one focused on process variables, just the final product. We asked when the data register was ready, not how we updated births and deaths. We asked what the response rate was, not how interviewers shared knowledge to improve rates. The difference is demonstrated by the first three boxes in the Plan for Continuous Improvement that Westat has introduced at a number of NSOs (see Figure 1). The Office of National Statistics' (ONS) Statistical Value Chain also emphasizes processes, although the change in emphasis is not quite as explicit as it is in Figure 1.

A Plan for Continuous Quality Improvement



Without process measurement it is difficult to know how to improve. If you change the process, what will happen to the product? You don't know. NSOs are accustomed to measuring products. Did they meet specifications? Then they started measuring a numeric score on the product; how close did it come to specifications? To improve, they realize they need to measure processes.

We like to use the example of developing a photograph. You can take a traditional photograph on film and develop it. You look at the picture a week later and realize that everyone in the photograph had their eyes closed. An improvement is to use a digital camera, so as soon as you take the picture, you realize that their eyes were closed. You can always go back and take another picture. The best would be to use a crystal ball, so that before you took the picture you would know the problem and make the appropriate adjustments. That is what real-time measurement of process variables is all about.

The NSO equivalent of the photograph can be the publication of a report or press release. Initially we track if it was issued on time. Then we get better by tracking how many days it took to release, differentiating between 1 day late and 4 days late. We still don't know what needs to be changed to be able to accurately predict the time it will take to have the report ready to release. Only by measuring the key processes involved in preparing the report will we know what to do. How long does data collection take? Data cleaning? Report writing? Only by measuring these variables real-time will we be able to use that information to adjust our processes before the report is late, or if necessary to notify users in advance of potential problems.

The need for measuring process variables has become evident to many NSOs that have gone beyond simple platitudes and actually tried to systematically improve. This is evident in Eurostat's Leadership Expert Group Recommendation #3, approved at the International Conference on Quality in Official Statistics in Stockholm in May 2001.

> Process measurements are vital for all improvement work. A handbook on the identification of key process variables, their measurement, and measurement analysis should be developed.

Implementation of this recommendation is the focus of Eurostat's Handbook on Improving Quality by Analysis of Process Variables, described by Nia Jones (2003). This handbook provides many wonderful examples of identification of process variables. One of the underlying points is how infrequently we are prepared to measure the key variables, even if they

have been identified. Our data systems, including our information technology systems, were not designed for this task.

2. Real Time Measurement

The key is not just to identify key process variables, not just to measure them, but also to measure them in real time. Can we measure key processes while the process is ongoing? Such real-time measures allow us to make adjustments to improve that product we are working on; not wait until the next time we try and produce a similar product. Even if we aren't sure how to modify our actions to make the improvement, real-time measurement will allow us to communicate the anticipated difficulties to customers in advance, when they may be able to adjust their expectations.

Let me focus on three European examples where they have successfully measured processes, but have not yet made the step to monitor and react real time (Jones et al., 2003). I will then talk about some examples from Westat.

The 2001 ONS Purchases Inquiry was extended to include the construction industry. The survey team noticed that respondents were having difficulty answering some of the questions. Item nonresponse rates were analysed and used as evidence to drop the question from analyses. Any question with low response rates over a two-year period is dropped from future waves of data collection. ONS is making good use of their process measurements to improve future rounds of data collection. I assume they are pre-testing question wording, but that these difficulties are not being detected at that time. I would urge them to go further, and monitor these rates during data collection and try to communicate with respondents when difficulties are identified.

The Portuguese Statistical Office (INE) has been evaluating the quality of their face-to-face interviewing for a number of years. They have surveyed their Labour Force Survey interviewers and mapped their interviewing process. They identified a number of key process variables, including delays in completing assignments and consistency rates between interviews and re-interviews. INE used a Pareto chart to separate out the main sources of shortfalls in number of completed interviews. Figure 2 clearly demonstrates that over half of the reduction is due to "lost units," what we refer to as ineligible units, such as vacant units or non-primary residences. Further investigation found that while refusal rates have been dropping somewhat

across waves of data collection, lost units are becoming more frequent. These are excellent examples of measuring process variables. What is missing (at least from the write up) was real-time feedback so that processes could be adjusted. When it was discovered that the list of addresses included many lost units, were the number of listings given to interviewers adjusted to enable them to reach their anticipated level of completes? What actions might have been implemented to clean out these lost units before they are sent to the field?

[insert Figure 2]

Similar analyses show that measures of consistency across time in re-interviews have been increasing as well. INE credits this to training sessions and communication between supervisors and interviewers explaining the meaning of the different possible answers to a question. On a quarterly survey like the Labour Force Survey, such improvements can be made between cycles. Unfortunately, other surveys have longer cycles and analyzing re-interviews after the data collection period are not very helpful. As an extreme example, in the United States 2000 Census, re-interviews were conducted on most of the over 400,000 interviewers. Unfortunately, the re-interviews were generally conducted as data collection in one location was concluding. Generally by the time any analyses of the results were conducted, the offending interviewer had moved on to another location and continued to follow their old patterns. Lessons learned from these re-interviews will not be implemented until the next census, 10 years later. Re-interviews were a quality control check to see how well interviewing had gone, not a quality assurance activity to *improve* data collection real time (Morganstein, Marker, and Levine 2003).

Statistics Sweden has been working "to identify and measure key process variables that are related to nonresponse errors, measurement errors, and productivity in interviewing activities." They begin by identifying *all* factors that might influence these outcomes, then try to define possible *key* process variables for each. Using cause-and-effect diagrams they found it relatively easy to identify factors, even if identification of measurable key process variables was difficult. It was still illuminating that this two-step process resulted in greater understanding of the factors that affect the outcomes, regardless of the success of the second step.

One of the measurable key process variables identified for both nonresponse errors and productivity was the interviewer burden, defined as the number of planned hours per survey per interviewer. This same variable was a key factor adversely affecting the response rates at Statistics Netherlands (CBS) when I visited there in the late 1990s. Interviewers typically worked on multiple surveys, but the sample was allocated to them independently from survey to survey.

As a result interviewers could be assigned impossible workloads. When this was combined with the minimal amount of supervision used by CBS, the result was "surprises" at the end of every data collection cycle, when interviewers turned in large numbers of cases that had never been worked. Monitoring these key variables real-time can identify potential problems and allow supervisors to shift responsibilities to avoid these problematic outcomes.

Fenwick and Tippen (2003, p 371) describe the quality management approach used for the UK Retail Price Index. Their documentation goes beyond just giving precise details of how tasks are to undertaken. It identifies "individual responsibilities, thereby increasing team understanding of roles and interrelationships between different team members." This allows staff to focus on key processes for which they are responsible.

As in many of the examples discussed above, there is an emphasis on post-cycle evaluation, rather than real-time measurement. ONS auditors re-check a random selection of items. The data collection contract allows for no more than 4 percent of the reported prices to be in error. Incentive payments are given if the rate is significantly below 4 percent. This methodology could be improved by requiring the contractors to re-check a sample of prices themselves and keep a control chart on this variable. ONS should clarify whether the standard of 4 percent is to reflect the average on the control chart or the upper control limit? Another way of saying that is do they want errors to typically be below 4 percent, or do they want this limit exceeded only in very rare situations? These require very different levels of accuracy, and have corresponding implications on the quality of the data.

Morganstein and Marker (1997) provided a number of examples of real-time process measurement. Telephone center phone calls to identify respondents and arrange in-person visits were monitored by primary sampling unit (PSU). A control chart was used to determine that procedures in urban PSUs needed to be modified while rural PSU procedures could remain unchanged. Edit failures were examined from initially processed questionnaires and a Pareto chart was used to identify a few questions that were frequently failing. Experts were called in and the edits modified, resulting in significantly reduced time required by these expensive experts.

At Westat we are currently conducting a survey evaluating the Terrorism Risk Insurance Act (TRIA) of 2001, passed by the U.S. Congress in response to the attacks on September 11, 2001. Respondents are given the option of filling out a mail questionnaire or providing the same information via the web. This is a very time consuming questionnaire. Approximately 30,000 businesses are asked to complete a 25-page questionnaire that typically

takes from 1 to 4 hours to complete. Three hundred and fifty insurance companies are given a 50-page questionnaire that takes about a month of calendar time (less actual work time) for many to fill out. While this is a voluntary survey, insurers have a strong interest in participating since TRIA provides re-insurance for them in case of a future international terrorist attack. The survey has much lower saliency for individual businesses, although for those in high-risk locations or activities the availability of such insurance also is very important.

We are using real-time monitoring of a number of process variables to try and improve the survey. Let me give four examples of process variables we are monitoring. The progress in our telephone center in completing telephone calls prompting sampled cases to complete the survey; completes to date by mode of data collection; meeting internal schedules; and item response rates from returned surveys.

Telephone prompting calls identified many situations where the person who received our questionnaire was not the appropriate person to fill it out. It was very important to identify such situations and to re-mail the form if necessary. It was hard to predict how long this process would take, and thus how many interviewer hours to schedule. Monitoring this variable allowed us to modify work schedules and finally gain agreement from the survey sponsor to extend the field period.

We were not sure what response rate we would achieve, nor what proportion would choose to answer via the web rather than by mail. Again, monitoring the low response rates allowed us to keep the sponsor informed of our progress and led to the extension of the field period.

Before beginning many major activities the project team developed flow charts that identified who was responsible for key steps and expected dates for their completion (Figure 3). Each box in Figure 3 has the name of the responsible staff member in the top left and a completion date in the top right. Based on this flow chart the due date for the deliverable file was determined. This flow chart is periodically reviewed to determine if we are still on schedule, and to keep the sponsor informed of implications for the final deliverable. Note that these are not arbitrary deadlines set by outsiders, but rather dates jointly agreed upon by the sponsor and the project team.

[Insert Flow Chart]

Finally, once a reasonable number of forms had been received we examined their item response rates. A surprising pattern was observed for question B2, which had four parts. It read in part:

B2. Please indicate if the statements below describe the organization. Is the organization selected a …

	YES	NO
a. Subsidiary of another U.S. firm?	1	2
b. Subsidiary of a foreign-owned firm?	1	2
c. Headquarters of your organization in the U.S.	1	2
d. Something else? (Specify) _____	1	2

Apparently many respondents interpreted this instruction as asking them to circle the "1" for those that applied to their organization. In many cases they did not bother to circle the "2" for the other questions. We determined this problem by looking at the high rate of missing data and common occurrence of circling one of the "1" while leaving others blank.

Since we observed this during the data collection period, we were able to try and rectify this situation. We had the email address for approximately one-fourth of the sampled cases. We sent them an email encouraging cooperation and providing some clarifications. Part of the email read "When answering question B2, please answer 1 or 2 for EACH of the four subparts of this question." As coders reviewed each questionnaire that was returned, one of the items they were told to look for was this pattern in B2. If at least one of the responses had a circled "1", they were to circle "2" for any that were left blank. Table 1 shows the percent of respondents for which the coders had to clean up this problem. It is clear that after the email the coding effort was reduced. (The sample sizes are large enough that these differences are statistically significant.) Sending the email to 7,000 sampled businesses cost nothing (other than a few hours of staff time to put together the email list), but coding time was definitely saved.

Table 1. Percent of respondents who needed coders to clean up question B2.

	Percent of responses requiring coder cleaning	
	Before email (n=2,269)	After email (n=986)
B2a	10.4%	6.6%
B2b	11.7%	6.3%
B2c	2.9%	1.0%
B2d	34.7%	23.9%

3. Conclusion

Measuring process variables is key to improving quality. Measuring these variables real time allows organizations to make those improvements as soon as possible. Eurostat and other organizations realize the importance of making these changes, but the examples of successful real-time process measurement have been few. This paper provides some examples that can hopefully be used by NSOs around the world to help in their improvement efforts.

References:

Fenwick, D. and Tippen, G. (2003). Quality management using ISO 9000 for price indices in the UK. *Journal of Official Statistics*, 19:4, 365-382.

Jones, N. (2003). LEG Recommendation #3: Handbook on Improving Quality by Analysis of Process Variables.

Morganstein, D.R. and Marker, D.A. (1997). Continuous quality improvement in statistical agencies. In L. Lyberg et al. (Eds.), Survey Measurement and Process Quality (pp. 475-500). New York: John Wiley & Sons.

Morganstein, D.R.; Marker, D.A.; and Levine, D. (2003). Evaluation of the Census 2000 Quality Assurance Philosophy and Approach Used in the Address List Development and Enumeration Operations Final Report. Prepared by Westat for the U.S. Census Bureau.

Saebo, H.V.; Byfuglien, J.; and Johannessen, R. (2003). Quality issues at Statistics Norway. *Journal of Official Statistics*, 19:3, 287-303.