WHAT IS SURVEY QUALITY: BACK TO THE FUTURE

Ron Fecso
National Agricultural Statistics Service, USDA

"Eternal vigilance is the price of freedom from serious statistical blunders."
Wallis and Roberts

In recent years survey research presentations at the ASA Annual Meetings, Census Bureau Research Conference and Washington Statistical Society seminars have often filled the room. This is very exciting compared to my memories of attendance ten years ago. I sometimes heard it said that there were no more interesting problems in survey research. Now, several areas of survey research are considered "interesting" by many. Cognitive aspects of surveys and computer assisted telephone or personal interviewing are examples. More specialized topics such as missing data procedures, raking and analysis of data from complex surveys also draw crowds. USDA Graduate School seminars on advanced topics in survey methods or sampling theory fill the largest classrooms. Journals devote special issues to survey topics. Why has interest picked up? Can the catalyst be survey quality or as Morris Hanson implied (Census Bureau Annual Research Conference V, ARCV), do we have a problem distinguishing fad's from progress?

Let's start with the question "How happy are we with survey quality?" When I asked that question at this year's Quality Assurance in Government Symposium, the overwhelming majority of hands were raised for the response "survey quality is a big problem," versus choices of "not a big problem" and "unsure."(Fecso,1989) We need to find out why we're unhappy with survey quality. Although it would take more than a book to fully examine our discomfort with survey quality, I hope these condensed ideas will help you address your questions more easily.

This paper will examine (1) how total quality management philosophies, which are becoming increasingly popular in industry, can be used in a survey organization, (2) the relationship between survey research methods and quality and/or process control tools, and (3) some comments on implementing total quality management in a survey organization.

SURVEY QUALITY

How big a problem is survey quality? To answer the question we need to measure survey quality. To measure survey quality, we need a definition. We have our first hurdle - you will not easily find a definition of survey quality in the statistical literature. It's not in the encyclopedia of statistics! Index citations usually contain the term "quality" as part of quality of data, data processing, sampling frames, etc. The concept of survey quality is at first vague, as it is when thinking about the quality of any complex product. For example, a quality car brings to mind ideas such as dependability, fit, finish, speed, cost, recalls, market share, etc. with each of us as either consumer or producer weighing the importance of each concept differently. Don't we worry about sampling errors, nonresponse, respondent burden, coverage error, wording problems, or other errors which are part of a long list of negative sounding survey research topics? No wonder we're uneasy about survey quality. How could we be happy with something so far from perfect? To begin with, we must distinguish quality from luxury. There are lemon luxury cars and well made economy cars. Similarly, we'll need to learn how a quality survey need not be a perfect survey. Let's now define survey quality. Simply add the word "survey" to the definition of quality found in the Glossary and Tables of Statistical Quality Control.(American Society for Quality Control[ASQC])

"[SURVEY]QUALITY -- The totality of features and characteristics of a [survey] that bear on its ability to satisfy given needs."

Further, let's also modify the definition of quality control:

"[SURVEY]QUALITY CONTROL -- The operational techniques and the activities which sustain a quality of [survey] that will satisfy given needs; also the use of such techniques and activities."

The ASQC definition continues to mention the "aim" and "integration" of quality control. The aim of quality control emphasizes the role of the organization and management's responsibility to ensure that the survey is:

SATISFACTORY,
ADEQUATE,
DEPENDABLE, and
ECONOMIC.

The integration of quality control emphasizes each person's role in the process and how they can assure that the process meets the organizational aim. Integration requires attention to:

SPECIFICATIONS,
DESIGN,
PRODUCTION,
INSPECTION, and
REVIEW OF USAGE.

Except for adding the word survey, the concepts of quality do not differ for service, industrial or survey processes. Raj points this out nicely. "A large scale survey is an exercise in statistical engineering . Each step in the production line is a potential source of error. The sample unit may not be identified correctly, enumerators may make errors in the field, there may be errors of coding or punching the cards in the office, and so on. Thus it becomes important to ensure that the production process is under control [Integration], and that the outgoing quality is acceptable [Aim]." Raj finishes the paragraph with a sentence which ties survey research to quality control. "Sampling methods can play an important part in achieving this." Although our survey process may itself be a sample, we can use subsampling, experimental design and charting techniques in ways similar to industry.

To understand the use of control methods, recognize that the basis of control is the same for all processes: **"a differentiation of the causes of variation... When only random variations are present in a process, the process is considered to be ' in statistical control'."(Oakland, p. 52)** Yet, there is a critical difference in how quality control methodology, as found in industrial QC, is applied to survey processes.

To understand why we cannot easily use the traditional quality methodology books (most of which address industrial applications) to solve survey quality problems, we need to compare the processes and the basis for statistical measurement. Table 1 illustrates the differences between the basic tasks in industrial and survey processes and the philosophy of statistical measurements within the processes. We find a case of opposites. Industry is trying to repeatedly make the same item of known measurements, while we do a survey because there is the need to know an unknown value, using a process that is at times repetitive yet not one of mass production. As a result of the differing processes, the assumptions for measurement of the process also contrast. In industry, measurement error is assumed to be present, but experimental error dominates, making tests for minor changes in the process efficient. However, in a survey, sampling error often dominates the error of an experiment, for example, when testing for wording or sequence differences. This makes the power of tests too low to be useful or the experiments too costly for current budgets.

Because the survey process is so different from the industrial process, use of industrial oriented textbooks for training purposes in a survey environment is often a mixed blessing. On the positive side, knowledge of the techniques provides a larger bag of tools to stimulate ideas. The drawback occurs when detailed methodological training is widespread too early in an organization's pursuit of survey quality methods. Training is successful when the material is specifically developed for application to a known task, and practice examples incorporate ideas related to the work. When there are no concrete examples for the student to apply to their process, the old cliche, "the tools are a hammer and every problem must look like a nail," occurs. After learning the traditional material in a QC course, one is easily drawn to create "control charts" for many indirect measures of survey performance (measures of the potential for bias). Response rate and data entry error rate are two "causes of variation" which are usually the first to be charted. I'm not implying that these charts cannot provide useful information, my worry is that indirect measures become **the** quality control process. This concern is based on having heard questions such as: How do you set upper and lower "acceptance" regions? Should there be a lower bound? What do we do when we go out of tolerance? Can we really change response rates? If we do, can it hurt survey quality more than help it? What about specification error, response error, coverage error? How do these measures inform us about survey quality?

TABLE 1 -- THE DIFFERENCE BETWEEN INDUSTRIAL AND SURVEY PROCESSES

	PROCESS	
	INDUSTRIAL	**SURVEY**
Basic Task	Try to repeatedly make the same things, each with the same measurements.	Try to measure many things, each with a different measurement.
Measurement Assumptions	Measurement error is always present but experimental error dominates	measurement error (sampling or response error) usually dominates

Problems in QC training also have occurred in service industries, where service specific textbooks have only recently been introduced. Survey quality, as does a service process, requires different measures and management approaches than those found in traditional QC textbooks. There is no time here for a detailed description of the content of a survey training program, but a few survey related references come to mind for those who need to try to develop this training. The books by Wright, Zarkovich and Rossi, Wright and Anderson are recommended. There are materials specific to subject matter surveys (health, business establishments, and education), which I hope to provide more detail about in a paper at the Winter Meetings.

Before we can progress toward quality or training, we need to return to our question, what is survey quality? When looking for ideas on survey quality in the survey research literature, one quickly realizes why most people have the perception that survey quality is a big problem. There is an overwhelming amount of survey research literature, both current and vintage. A new convert to quality thinking in a survey organization might be faced with a dilemma--reading the literature would preclude having time to accomplish any work. A quick review of some important survey research history illustrates the amount of survey error knowledge which has been available (See Table 2).

Obviously, much more has been done than listed here. The point is that many of our "quality" problems are not new nor are many of their solutions. If you were to glance through the 28 pages of articles listed in Dalenius' "Bibliography of Nonsampling Errors in Surveys," you'd find everything from broad error profiles to very specific treatments of a specific error in a specialty survey. Morgenstern's "On the Accuracy of Economic Observations" first published in 1950 is another good place to obtain a sense of the breath of the survey quality problem. If such long and old lists of problems, many of which remain in the "product," existed in an industry, we might expect that the "product" will soon be stamped "made in Japan." Fortunately for our jobs, it's difficult to export a survey. Yet, the importance of surveys in this "information age" demands that we not allow lack of competition to fool us into believing that we don't have to continually work on survey quality.

TABLE 2--SOME HIGHLIGHTS OF SURVEY ERROR HISTORY

Many of the references in this listing come from Forsman, (ARC III) and Converse.

<u>18th AND 19th Centuries</u>
Measurement errors in censuses recognized

<u>1920's</u>
Problems with purposive selection attract attention (Gini and Galvani)
Interviewer effects on response error shown (Rice,1929)

<u>1930's</u>
Straw poll problems become famous, the Literary Digest Poll
Sample Surveys (Kiaer, Neyman)
Area Sampling (Jessen)

<u>1940's</u>
Interpenetrating subsamples to study nonsampling errors (Mahalanobis, 1946)
Total Error Model development begins (Deming and Geoffrey, 1942; Palmer, 1943; Hansen and Hurwitz, 1946; to name a few)
Raking (Deming and Stephan, 1940)
Debate on whether or not to publish response rates.

<u>1950's</u>
Census Bureau Survey Model (Hansen, et.al., 1951)
Linear Survey Models (Stock and Hochstim, 1951)

<u>1960's</u>
Index of inconsistency to measure response variance contribution to Total Error (Hansen,1964)
Randomized response model (Warner, 1965)

<u>1970's</u>
Total Survey Design (Dalenius, 1974)
Bibliography of Nonsampling Errors in Surveys (Dalenius, 1977)

Groves stated that studying survey quality brings out two reactions:

1. Why do we know so little about survey quality?
2. Why don't we use all that we know?

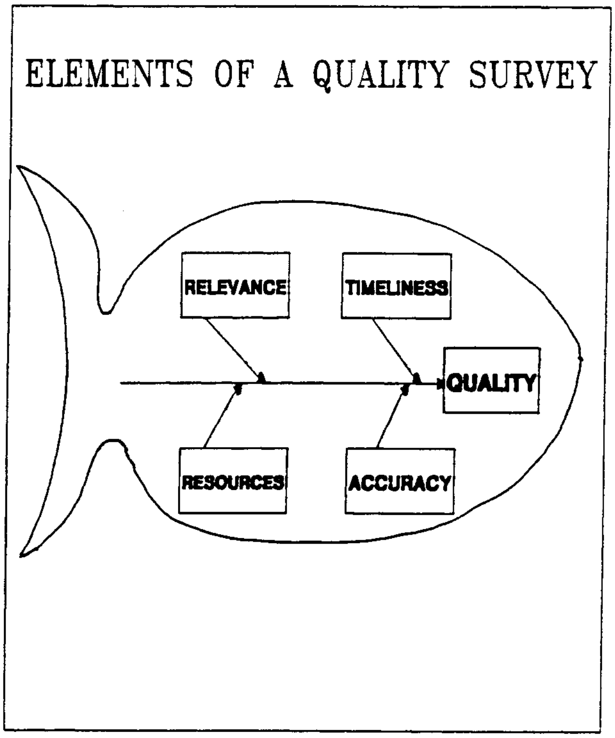
With all the survey research that has been done, why must we feel uneasy when we hear the phrase "survey quality?" Do we feel uneasy because we are unsure about the relative importance of the many outputs of our surveys? Which output should be more accurate? Are we overwhelmed by ideas and at a loss over which we can afford to do or have the time to do? These are the questions which characterize a "state of confusion" for which the quality philosophy textbooks indicate a pressing need for a formal total quality management approach.

An excellent start towards a formal basis for comprehensive quality control was written by Dalenius (1983). The first of two components in Dalenius' formal basis, measuring the usefulness of a survey considers traditional QC concerns such as accuracy, cost, and relevance, as well as topics such as privacy protection and wealth of detail. To this list I add timeliness, even though it may be considered a part of relevance, since it is often a management priority and it directly affects both accuracy and cost. Relevance, accuracy, resources (cost) and timeliness are the concerns or AIMS of upper level survey managers (Figure 1), knowing that the product is satisfactory, adequate, dependable and economic. Management's task is to keep the concerns in balance with respect to the needs of the users and the capabilities of the organization. Dalenius' second component, a comprehensive QC program, has two subprograms: accounting for the control of the survey design and accounting for the control of the survey operations. These are the INTEGRATING items which are the task of everyone. They include knowing the specifications, proper design, care in production, inspection and review.

GETTING STARTED

How does a survey organization begin the quest for survey quality? What should management do? What measures should be done now? It's tempting to begin a mass quality training program. What tempts us is that we can continue to get our own work done while "delegating" the uncertainty of what to do to the trainers, consultants, or quality section. Avoid the trap! Although training is a necessity, mass training of the "workers" succeeds when the work involves a few simple repetitive tasks which have well specified engineering tolerances, that is in cases where the simple tools of charting can be understood and easily applied. A mass training approach would likely suffer in a survey organization where little of the standard training material would immediately apply. It's best that we listen to Deming and the other quality philosophy gurus and have the top level of management get involved to provide the road map to quality. The "top down" approach

Figure 1--A "Fishbone" chart of Survey Quality



is important because we haven't answered the question, "Is there a quality problem?" To answer the question, management needs to know the causes of variation in the survey process and determine whether all but the random sources of error are removed. Management can do this by being the link between cost and quality; deciding what price we pay for quality. There are hundreds (at least) of survey research tools which can be used to measure or improve accuracy, resources, relevance, and timeliness, far too many tools for an organization to implement all of them. Management therefore must decide which survey research tools deserve the resources. Management also decides when new areas of research are ready to go operational or show too little promise to consider funding. These decisions are made, like most decisions, with uncertainty. However, as managers and statisticians we should be among the best prepared to be comfortable with decision making. The tools of our profession are made to help us live with uncertainty, but we must measure the uncertainty to be sure we can live with it. Only with measures of our processes can management decisions lead the way to quality.

A key to statistical control of a complex process is the development of a management system which uses statistically based decisions to attain the specified goals. This is the surest way to organizational quality of any process within the organization. Deming expresses this as management's obligation to provide a road map to improvement.

MANAGING IN QUALITY

Upper level management must be sold on quality as the route to organizational excellence. Of course they want quality, yet many do not easily buy into the philosophy because the packaging of the Deming and other gurus' messages may be confrontational with top management (Joiner). The management "lecture" may be too blunt and offensive, even (or perhaps especially) where needed. A rash of "don'ts" appears. Don't disregard workers knowledge. Don't appear uncommitted. Don't give vague instructions. What managers would admit that this was their style? After all, knowledge, training and standing as a survey researcher are often the criteria for entering management. (Not unlike engineering in industry.) The message is simple and can be stated softly:

- Encourage discussion
- Act for the long run
- Improve through education and recruitment
- Show commitment by using measurements to drive decision making

Who can deny that these are good management practices? Who will admit they don't practice them? Increasing information demands, tightening/uncertain budgets and competing for the supply of quantitatively oriented students pose serious problems for managing for the long run. These issues certainly deserve more attention, although they are not the focus of this paper. The focus will be actions which we can begin immediately.

Management can begin their leading role in survey quality by assessing the state of quality in the survey organization. A good tool comes from the criteria for the Japan QC prize. The ten issues involved can easily be worded for a survey organization (Fecso, 1986) and summarized under three management functions.

1. What measures does top management use to quantify the effectiveness of the surveys and organization as a whole?

2. How are these measures used in a top-down system of quality management? Specifically:
   - What are the survey standards?
   - How are data gathered and turned into information? (Charts, etc.)
   - How is QC organized and promoted by management?
   - What statistical and QC methods training is available?

3. How does management use the data to plan for the future?
   - How are problems identified?
   - How are new initiatives or research ideas prioritized and allocated resources?
   - What data sources are used as a basis for further planning?
   - How are long- and short-range plans kept consistent?

Our overall uncertainty about survey quality makes me believe that too little use is made of measures and planning strategies.

DEVELOPING MANAGEMENT'S QUALITY CONTROL TOOLS

The role of management in a survey organization is varied and complex. I recommend that managers read Kenneth Prewitt's article, "Management of Survey Organizations" (Rossi, Wright and Anderson) for insights on survey organizational goals (prestige, survival and growth) and operational issues. Thus, the role of management in survey quality control is to determine when the system (a survey or the organization as a whole) is outside the standards it sets for organizational goals and operations. Management must have tools which identify the source of deviation from the standard and decide what resources of the organization should be used to control or remove the source. The problem is one of resource allocation.

For now, let's look at accuracy, one of the four components of survey quality. Few managers would question the need to measure the extent of nonresponse bias, response deviation and bias, and coverage error, to name a few. Yet, how many surveys measure all the "perceived problems?" For some idea's, I suggest reading the report "Quality in Establishment Surveys" (Office of Management and Budget). Over a quarter of the surveys examined did not compute an estimate of variance. Measures of specification error (such as comparisons with independent estimates, pretests, cognitive studies or records check studies) had only modest use. Specific measures of coverage error such as birth rate, duplication rate, and out of scope rate were computed on half the surveys at best. Interviewer or response variance studies were done on only about 10 percent of the surveys. Measurement of nonresponse and processing errors were similarly lacking.

These error sources, which we know are important in determining how we feel about survey quality are not being universally addressed. It's no wonder we "perceive" that survey quality is a big problem. So many techniques are available and so few are used. Yet, the true issue is how much of a problem are these error sources? It's time to begin to answer this question and to understand the role of timeliness, resources and relevance. The lead belongs to management, since we don't have the resources to implement all the known survey research tools, nor are they all appropriate in a particular survey.

Management can lead the way out of the cloud of uncertainty about survey quality. Careful selection of your approaches to specific problems is required. As is true in industry, don't try to solve all the organizations problems at once. It's important that your initial efforts be successful in order to convince the doubtful to "buy in" on quality. You'll need time to learn and feel comfortable with the methodology, so start with one survey. Prioritize the importance of your surveys and choose one of the most important. Sure, they're all important, yet some have higher visibility or a stronger perception of having a quality problem. Next, develop a small number of the most important standards or goals you have for that survey for each quality concern: resources, accuracy, timeliness, and relevance. Feel free to adjust these quality concerns if necessary. After all, you know the "needs" of your organization better than me. Some ideas to consider, which should be widely applicable (but certainly not all inclusive) follow:

Ideas for Survey Quality Standards

Quality Component	Standard or Goal
Timeliness	Major due dates met. Publication on time.
Resources	Survey within budget. More resources allocated to measure quality. Staffing adequately used.
Relevance	User meetings conducted with concerns and priorities for resource allocation identified.
Accuracy	Total error models developed. Major "perceived" errors measured.

Avoid too long a list. Use nominal group techniques or other "brainstorming" methods to reach a top management consensus on the important standards or goals. You'll need to make time to develop your measures for items on the list. The point is to convey your desires to the next level of management in a measurable way. The next level of management would then develop more micro level standards for their function in the survey process.

The standards and goals chosen will also depend upon the survey methods and purpose. Bailar presented a nice cross tabulation of survey types (single time, repeated with or without overlap, longitudinal with complete or partial overlap) by the kind of estimates that can be produced (one point in time, durations, relationships, net or gross change, trends, cumulated data). Of the 40 cells in the table, 31 were indicated as survey-type by kind-of-estimate-desired possibilities (e.g. single time survey to estimate for one point in time, repeated partial overlap survey for trends). The survey process for each combination differs and thus each requires development of a specific quality management strategy. Without development and measurement of some basic management standards and goals, how can the "workers" in the organization be anything but confused about the quality of their work?

In the remainder of this paper I will concentrate on quality management of repeated surveys. Repeated surveys are likely to have a higher management priority than single time surveys in most organizations, some techniques are more readily implemented, and many ideas will generalize to other types of surveys. The remainder of the paper is an example of getting started on one of the survey quality concerns.

AN EXAMPLE - ACCURACY IN A REPEATED SURVEY

Assessing quality with respect to accuracy cannot be done without first relating the accuracy measures to the relevance component of survey quality. Spencer (1985) presents an interesting analysis which relates probability of use, precision and Bayes regret. In this analysis, the more likely the use of the data the more important data quality becomes. Also, an interesting note was that the **perception** of quality directly influences the likeliness of use. His 1982 paper presents some motivation for data improvement and ideas on cost-benefit analysis.

An immediate problem when evaluating surveys is that most surveys result in a large number of estimates. Since the sources of error for each estimate will vary, you may need to again limit the scope of your effort. Prioritize the estimates made for the survey and pick the several most important variables. Many of the quality issues (coverage, nonresponse, etc.) which affect the most important variables will also affect those not chosen, thus your efforts will indirectly affect all the survey items. I suggest having this limited focus for several reasons:

1. Can you do a quality job "mass producing" quality before you're sure what it is and are comfortable doing it? Most survey organizations are heavily burdened to keep up with the workload; a "do-able" effort carried to completion

will do more good than trying to attack every error and having no time to really see any results.

2. Staff who have the training to determine accuracy problems and to develop accuracy measures may be limited; focus scarce resources on the most important issue.

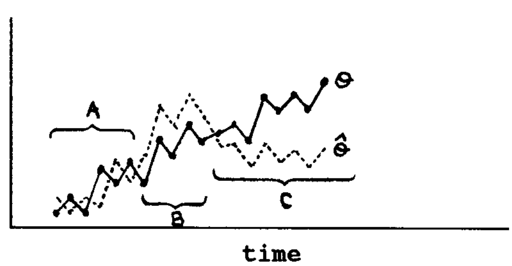
3. Once comfortable with measuring the main items of interest, resources can be allocated to develop generalized variance models and possibly to generalize nonsampling error models (Tortora) to provide cost effective measures of the lesser variables.

Next, the most important uses of the data must be identified. For this example, I'll assume that time series of unbiased estimates and the ability to identify differences of a specified amount between survey periods are identified as priority uses of the data. Obviously, other uses exist such as the relationship between data items or between time and the data items (moving averages, cycles, etc.).

Consider first the need for a time series of estimates. Figure 2 presents the impact of nonsampling error on the survey series (for now assume no sampling error). Obviously, the survey series would be misleading to the user if there were any unknown patterned biases. Thus, management should feel uneasy if there is no reliable measure of the total bias. The management action needed should be obvious here, to develop these bias measures. This will require personnel with the proper training and experience and the time and resources to develop the measures. Are we lacking the human or financial resources or the time to do this?

Next, assume there are no nonsampling errors and consider how we'll address the "ability to identify differences of a specified amount." The key here is to identify the "specified amount." Let's call the difference of interest "d." The first simple check is to examine the time series itself. Figure 3 shows some potential problems. In effect, if the time series does not behave like the area marked "D," then there is an accuracy or relevance problem with the survey. The problem could be operational

Figure -2- Some difficulties with biases



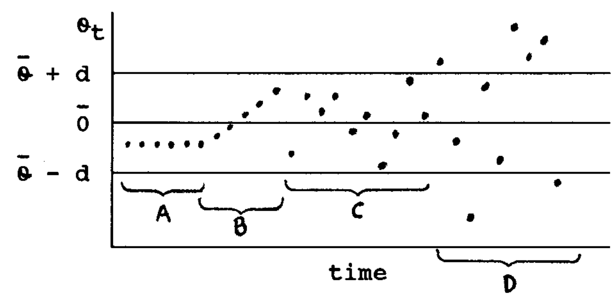
A. Bias negatively correlated with change
B. Constant bias
C. Cumulative biases

(systematic biases) or a lack of need (the parameter does not deviate from the long term mean by more than d).

Based on the idea that an important part of relevance is the ability to recognize change, Figure 4 addresses the impact of the precision component of accuracy on relevance. The difference between successive estimates is plotted as a run chart with the 95% confidence (your choice of size) interval indicated. Run A shows a survey which does not appear to differ by d, making the need to do the survey or the size of d questionable. Run B indicates a survey with too little power to detect a difference from zero. Why use resources on such a survey? If we need the estimate, we should reallocate resources to reduce the variance to a relevant level. Run C can detect non zero differences, meeting the objective of the survey. The smallness of the confidence interval and d may be thought of as reflecting the luxury of the survey. Inability to detect the chosen d is a standards failure, that is, a quality problem in need of action.

We've found that changes in the confidence interval usually indicate a new, undetected source of nonsampling error, and a time series of sampling errors shows this readily. I anticipate that time series ideas (moving averages, trends, etc.) will provide methods to further explore these charts.

Figure -3- Characteristics evident in the time series



A. Runs which should not be expected
B. Trends which should not be expected
C. Series not different from the mean by d.

Figure -4- Precision, A part of relevance

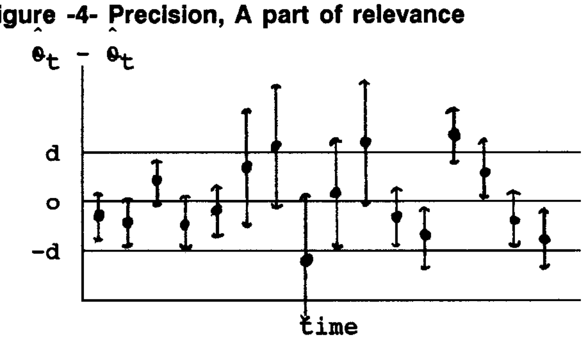


Figure 5 represents an example where breaking the "rules of graphics" helps surveys. We've learned that graphs should be kept uncluttered and simple to be effective for most users. In this example, the "clutter" helps us identify nonsampling errors which have eluded an extensive edit system. The plot shows estimated standard errors of a total by stratum and for the combined strata by year. The thick band of lines at the bottom of the chart indicates the usual variability in the estimates of the standard error. Steep spikes rising from the "control swarm" are evident. Investigation of the data in these strata found one or two observations which eluded the edit system and caused each spike. A steep decline of the estimated standard error of the total is evident at the right of the graph. Further investigation found that a new summary system had failed to square the stratum weights in the estimator. A chart like this is a useful acceptance measure of the "final" data run since it is quite sensitive to individual observations which would cause notable biases.

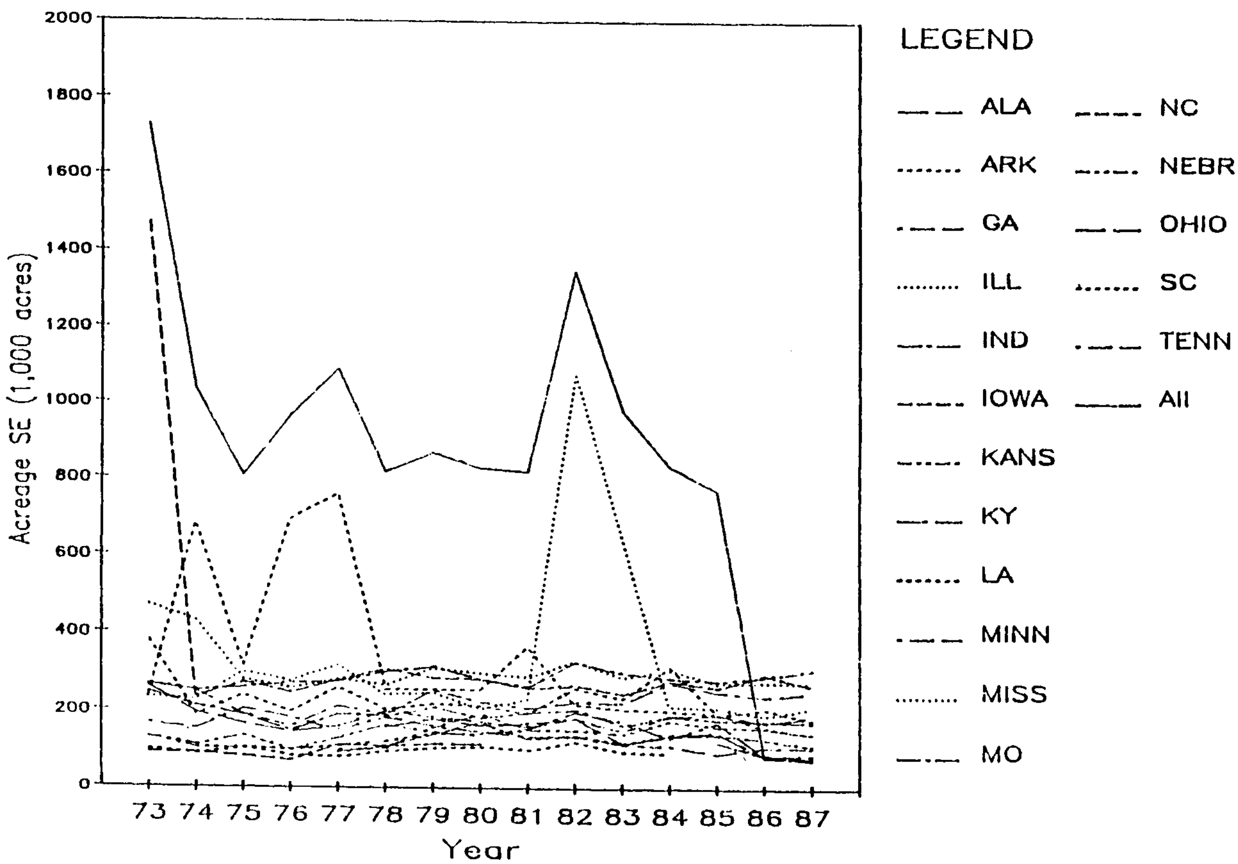
There are more examples of graphics like these which we've explored at NASS. There also are some interesting examples in a paper by Art Silverberg (in Wright). These graphics are simple and even obvious to many. I would fully agree with anyone who said these points should be made through total survey error models. There is no question this can be done, and that error models have been recommended since the early 50's. The problem is that few surveys have total error models. I'd like to propose that all surveys have total error models developed. But there are no comprehensive models which we all can use and general models tend to provide limited information. Since survey designs vary, application specific models will require management to find funds and qualified researchers to do this "engineering" work. We need to take the time to motivate the need for this management decision. Until then, the major benefits of creating simple graphics to explain more mathematical concepts is that operational personnel can grasp the ideas, create the measures and use the measures with little formal mathematical training.

How can managers decide when to allocate and the amount of resources to allocate to accuracy improvement when there is competition to use the resources to provide more timely or additional data? Pareto charts are useful in depicting the balance between accuracy and costs. Consider a simple model for our repeated survey example:

$$Y_t = T_t + e_t + B_t$$

Where Y_t = the survey result at time t,
T_t = the true value at time t,
e_t = sampling error at time t,
B_t = total survey bias at time t.

Figure 5--Estimated Standard Errors of the Total by Stratum and for combined Strata.



Looking at the example's most relevant statistics, the difference between periods, we find

$$Y_t - Y_{t-1} = (T_t - T_{t-1}) + (e_t - e_{t-1}) + (B_t - B_{t-1})$$

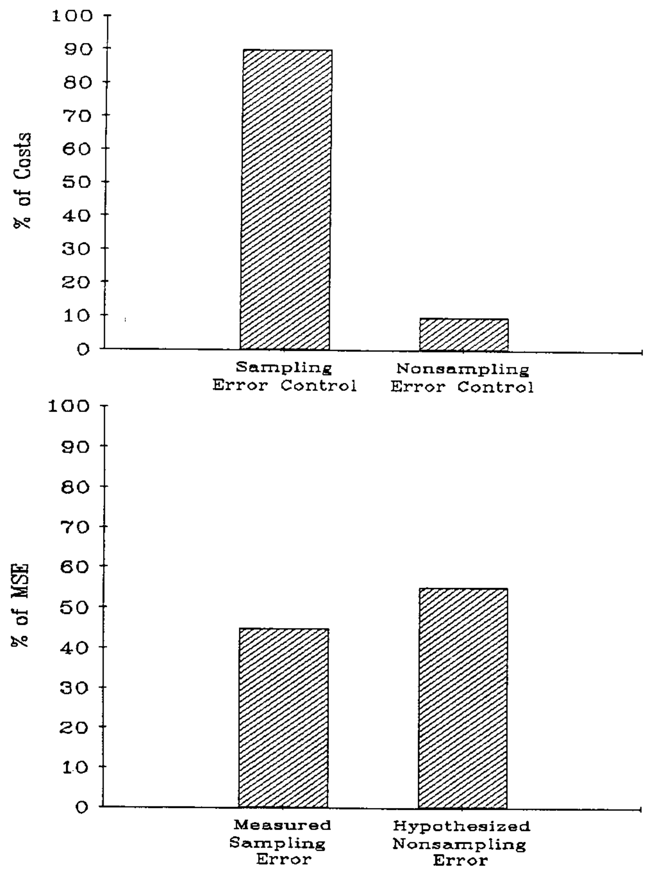
$$= D + E + B$$

Where D = the true difference,
E = the sampling error difference and
B = the bias difference due to all sources.

The usefulness and ultimate relevance of the survey value depends upon E and B. Management can compare the costs associated with these component using pareto charts as in Figure 6. We would like to see the shape of the histograms become reasonably alike. However, in the typical survey case, we hypothesize that nonsampling errors exceed sampling errors and yet most organizations spend little to measure and/or reduce their contribution to total error. Parato chart comparisons become more effective as you separate components of both sampling and nonsampling error. Similarly, comparisons of the importance of the variables collected in the survey with their cost and accuracy can be made to aid in resource distribution decisions.

Again, let me stress that I can't tell you that these are "the" tools to use. These ideas have worked well at NASS, especially in presentations to our Planning committee.

Figure 6
A Simplified Pareto Analysis
Of Costs and Error Sources



We've been able to quickly and convincingly convey the need to form quality teams and fund several nonsampling error studies. Hopefully, you'll find some ideas here and be motivated to develop others.

CONCLUSIONS

This paper certainly does not contain all the answers. My hope was to give ideas which encourage you to more formally approach quality issues and to pass on some "first things to read" which I found useful. If you are a top manager, the next step is yours, if not you must begin by convincing top management to buy-in. Top management must begin by providing the road map to quality.

Suppose I were to ask the top manager in your organization to show me what is known about the quality of a particular survey. How would that person respond? Being presented with a long list of research papers or having the resident "expert" summoned to the office would be a typical response, but not one which indicates that quality management flows from the top down. The manager who had a working set of measures related to the organizations major concerns (e.g. performance, features, reliability, conformance, durability, serviceability, responsiveness, aesthetics, reputation, work environment, and employee characteristics) would have taken the first step toward clearly conveying their plan for improved quality and productivity. Thus, I suggest that initial efforts be directed at developing these top management measures. Possibly in the form of a "state of the survey" report.

Top management also needs to stimulate people to be involved and innovative. There are some simple guides which can help. Recognize contributions which show:

1. measured improvement, or
2. improved methods to measure improvement.

Finally, top management must reflect back on the long history of research which exists for most surveys. Why is there so much unused methodology? Have we done a poor job in selecting or conducting research? Or, are we failing to implement worthwhile ideas for the time honored, but quality and productivity reducing, excuses of lack of time, politically unfavorable, no money, next year, the client doesn't want it, etc. Let's look back at what we know about survey quality and organize the information so we can proceed with less uncertainty about the decisions we need to make and the status of survey quality. The concepts of total quality management can provide us with the framework for organizing our thoughts, but we must provide the thoughts and most importantly carry through the actions.

ACKNOWLEDGEMENTS

Thanks to Milt Goldsamt, Toni Tremblay, Read Johnston and Wendy Waters for editorial and compositional assistance.

REFERENCES

American Society for Quality Control (1983), Glossary and Tables for Statistical Quality Control: Milwaukee, WI.

Andersen, R., Kasper, J., Frankel, M.R., and Associates (1979), Total Survey Error, San Francisco: Jossey-Bass Publishers.

Andrews, Frank M. (1982), The Construct Validity and Error Components of Survey Measures: Estimates from a Structural Modeling Approach, Proceedings of the Fourth Conference on Health Survey Research Methods, National Center for Health Services Research.

ARC I-V (1985-1989), Proceedings of the Census Bureau Research Conferences, Commerce Dept., Washington, D.C.

Bailar, Barbara (1987), "Information Needs, Surveys, and Measurement Errors," Proceedings of the Census Bureau Research Conference.

Converse, Jean M.(1987), Survey Research in the United States: Roots and Emergence 1890-1960, Berkeley: University of California Press.

Dalenius, Tore (1977), "Bibliography of Nonsampling Errors in Surveys," Int. Stat. Rev., Vol.45, pp. 71-81 and 181-197.

Dalenius, Tore (1983), "Errors and Other Limitations of Surveys" in Statistical Methods and the Improvement of Data Quality, Tommy Wright ed., Orlando: Academic Press.

Deming, W.E. and F. Stephen (1940), "On a Least Squares Adjustment of a Sampled Frequency Table When the Marginal Totals are Known," Annals of Math. Stat., 11, 427-444.

Fecso, Ron (1986), "Sample Survey Quality: Issues and Examples from an Agricultural Survey," Proceedings of the Section on Survey Research Methods, American Statistical Association.

Fecso, Ron (1989 and 1988), "Survey Quality," Notes of the Quality Assurance in Government Symposium, Washington Statistical Society.

Groves, Robert M. (1987), "Research on Survey Data Quality", Public Opinion Quarterly, Vol. 51, 5156-5172.

Joiner, Brian L. (1985), "The Key Role of Statisticians in the Transformation of North American Industry", The American Statistician, Vol. 39, No. 3, pp. 224-7.

Morgenstern, Oskar (1963), On the Accuracy of Economic Observations, Princeton Univ. Press, Princeton.

Oakland, John S. (1986), Statistical Process Control, William Heinemann Ltd., London.

Public Health Service (1977, 1982), Health Survey Research Methods, DHEW Pub No. (PHS) 79-3207.

Raj, Des (1968), Sampling Theory, McGraw-Hill.

Rossi, Peter H., James D. Wright and Andy B. Anderson (1983), Handbook of Survey Research, Orlando, Academic Press.

Spencer, Bruce D. (1982), "Feasibility of Benefit-Cost Analysis of Data Programs," Evaluation Review, Vol. 6, No. 5, 649-672.

Spencer, Bruce D. (1985), "Optimal Data Quality," JASA, 80, 564-573.

Tortora, Robert D. (1987), "Quantifying Nonsampling Errors and Bias", Journal of Official Statistics, Vol.3, No. 4.

U.S. Office of Management and Budget (1988), "Quality in Establishment Surveys", Statistical Policy Working Paper 15, Statistical Policy Office, Washington, DC.

Wallis, W.A., and H.V. Roberts (1965), Statistics: A New Approach, New York: The Free Press.

Wright, Tommy, ed. (1983), Statistical Methods and the Improvement of Data Quality, Orlando: Academic Press.

Zarkovich, S.S. (1966), Quality of Statistical Data, Food and Agricultural Organization of the United Nations, Rome.