

# Why Innovation Is Difficult in Government Surveys<sup>1</sup>

*Don A. Dillman*

Needed innovation and change in government survey organizations are often difficult to accomplish, especially for dealing with measurement and nonresponse survey error. It is argued here that three important contributors to this difficulty are (1) the necessary co-existence in such organizations of research and operations cultures; (2) major differences in the dominant value systems of those cultures; and (3) the difficulty of resolving those differences in hierarchically-oriented organizations. Four steps towards a solution to these innovation barriers are recommended.

*Key words:* Innovation; work cultures; organizational hierarchy; survey error; government surveys.

## 1. Introduction

Innovation in large survey organizations is as essential as breathing—there is no alternative. Technologies continue to change, redefining what is possible and what is demanded by survey sponsors. In addition, the receptivity of survey respondents to being asked to participate in surveys, and towards particular survey methods and technologies is also changing.

A result is to place demands for change upon large survey organizations, and the people who staff them, in directions that are uncharted and often difficult to accept. Indeed, the central tendency of most large organizations is towards preservation of the status quo rather than towards change. A long-time employee of a federal statistical agency summed up this tendency by advising an advocate of a proposal for change that would affect several divisions of the organization: “You should expect that your proposal will produce at least five reasons why it won’t work, five more

<sup>1</sup> Revised version of the invited banquet presentation to U.S. Bureau of the Census Annual Research Conference, Arlington, VA, March 21, 1994. Don A. Dillman is Professor of Sociology and Rural Sociology and Deputy Director of the Social and Economic Sciences Research Center, Washington State University, Pullman, Washington 99164-4014, U.S.A. From 1991-1995, he served under an Intergovernmental Personnel Act Appointment as Senior Survey Methodologist, Office of the Director, U.S. Bureau of the Census. The assistance of many people from several government agencies, including the National Agricultural Statistics Service, the National Science Foundation, the General Accounting Office, the Bureau of Labor Statistics, the National Center for Education Statistics, the Office of Management and Budget, and the Bureau of the Census in helping the author to understand how government surveys get done is gratefully acknowledged. I also appreciate the encouragement of Robert Tortora, Robert Groves, Harry Scarr, and Maxine Anderson-Brown to contemplate, systematically, the issues addressed in this article. Nevertheless, the opinions expressed here are those of the author and should not be attributed to any of these individuals or agencies or Washington State University.

reasons why even if it were to work it shouldn't be done, and at least twenty-five memos explaining each and everyone of those reasons in detail!"

Two considerations offer perspective on why innovation may be difficult to achieve in survey organizations. The first is the fact that fundamental change has come to survey methodology with great rapidity over the last two decades. Only twenty years ago mail surveys were defined as something to avoid by all reputable survey texts and telephone survey methods were not considered acceptable for any important surveys. Computer Assisted Telephone Interviewing, or CATI, was in its infancy and the personal computers essential for computer assisted personal interviewing did not exist. Measurement error was something people occasionally talked about, but the extent of its detrimental effects on survey results was not well understood. And, although mixed-mode surveys were occasionally done, little thought was given to the possibility that people provided answers to interviewers that differed from those they would have provided to self-administered surveys.

The second consideration is that an extensive research literature on the diffusion of innovations has shown that the acceptance of innovations almost always occurs far more slowly than necessary (Rogers 1981). Although the speed by which the acceptance of new methods occurs depends upon a variety of factors, ranging from personal characteristics to the attributes of the innovation, organizational characteristics are also an inhibitor of rapid adoption. Usually, it is organizational characteristics that inhibit organizational acceptance of innovations more than it is personal or individual characteristics. This fact is integral to the arguments laid out here.

In this article I argue that three interconnected features of large government survey organizations make it especially difficult for such organizations to foster innovation and change. They are (1) the co-existence of two quite different work cultures (research and operations); (2) major differences in the dominant value systems of those cultures; and (3) the difficulty of resolving those differences in hierarchically-oriented organizations. Four steps towards a solution to these challenges are recommended for improving the innovative climate in federal statistical agencies.

The arguments presented here draw heavily upon my experiences at the U.S. Bureau of the Census from 1991–1995, where I served as Senior Survey Methodologist in the Office of the Director. While there I participated in a research program, the aim of which was to define and test possibilities for fundamental change in procedures used to conduct the U.S. Decennial Census. In addition I was able to work with survey methodologists in other areas of the Census Bureau and several other government survey agencies on a variety of projects that suggested similar difficulties for achieving effective innovation.

From these observations, an irony became evident. Without exception these organizations had many well-trained professionals who saw a need to facilitate innovation. Yet, the adoption of ideas, particularly for reducing measurement and nonresponse error, seemed difficult to accomplish. It was also apparent that final survey decisions were sometimes an amalgamation of incompatible ideas, some of which had undesirable consequences for survey error, and, ultimately, seemed beyond the ability of any single individual to control. This article attempts to provide an understanding of some of the factors contributing to this result.

## 2. Operations vs. Research Cultures

Government survey agencies face tasks unlike those usually faced in universities or private sector work. Some government surveys are of great scale and complexity, so that not only do they present huge operational problems, but much of the specific knowledge for designing and implementing them must come from research which only the host agency can design and implement.

A circumstance is thereby created that might be likened to, for example, an aircraft manufacturer attempting to operate an airline while continuing to design aircraft. Putting pilots and flight attendants into the same room with aeronautical and thermal systems engineers – each representing multi-million dollar enterprises with equal investment in the outcome of their common research project – could produce some unpredictable as well as strange outcomes.

One of my early experiences at a federal statistical agency was to chair a working group tasked with developing a priority listing of research undertakings. There was one research endeavor about which I felt very strongly, so I did an elaborate job of justifying why an experiment needed to be conducted – it had to do with development of respondent-friendly design for questionnaires. It seemed to be well-received by the group. Perhaps you can imagine my surprise, when at the time the rankings were to be done, a person whom I thought I had convinced of its importance was ready to rank it dead-last. When I recovered enough to ask why, the answer was simple. “Oh, I agree it’s important, but you’ve convinced me it’ll work. Therefore, we don’t need to test it, let’s just do it!”

A large government statistical agency will have staff who are part of an *operations* culture – there is a job to be done, and research should not waste efforts on things that we are confident will work. Impressionistic evaluations of an idea are often considered adequate, and control groups, although desirable, are often viewed as expendable.

Other employees are part of a *research* culture. They would declare just as emphatically it is the nature of research to test what we believe will work to learn exactly *why* and *how it works*.

Given this environment, it should not be surprising that proposed tests of innovative ideas sometimes become a tug-of-war. To some, a good research project is “practice” in order to form impressions of whether something works. Control groups are not really necessary, and the fewer treatment groups the better. From this perspective the real value of research is as the rehearsal is to a stage performance, where one dares not fail. At the same time, those from the research culture have in mind carefully designed treatment factors and a full factorial design. Preordained rules of assignment to treatment and control groups as well as rules for interpretation of evidence must be scrupulously followed.

Both “research” and “practice” are essential to the success of government survey organizations, and the involvement of both is crucial for achieving the adoption of new technologies and procedures for large government surveys. However, the perceived needs of one often interfere significantly with the needs of the other. When “test” dollars are scarce, a frequent result is to attempt to get individual projects to serve both as practice and research vehicles. The operations culture often wants to

test a microcosm of the survey organization with the development of team skills being part of the test, while those from the research culture often want to limit the objectives of a project so results can be interpreted unequivocally. The unfortunate and frequent result of these competing concerns is that neither objective is accomplished as well as one would like.

Unfortunately, the major barrier to needed innovations that stems from co-existence of the research and operations cultures in government survey agencies is far more complex than differences in philosophy of what innovative ideas are important to test and how they should be tested. This barrier stems from differences in the core value systems of each culture and a division of responsibility that results in the over-emphasis of some issues at the expense of others.

### **3. Core Value Systems of the Research and Operations Cultures**

There are four commonly recognized sources of data collection error (Groves 1989). When surveys are designed, attempts should be made to limit error from each of them.

These sources of error include:

coverage error: the result of all units of a population (e.g., people, households, or organizations) not having a known, nonzero probability of inclusion in the sample drawn to represent that population;

sampling error: the result of surveying a sample of the population rather than the entire population;

measurement error: the result of inaccurate responses stemming from poor question wording, poor interviewing, survey mode effects, or the respondent's behavior; and,

nonresponse error: the result of nonresponse from people who, had they been surveyed, would have provided different answers than those who did respond to the survey.

Each source of error has a scientific foundation which explains its occurrence. Leaving aside coverage error for the present, the three remaining sources of survey error are described and attempts made to assess their occurrence.

Sampling error is generally considered the province of statistics and subject to principles derived from probability and other mathematical and statistical theories about which much knowledge has accumulated throughout this century. In contrast, our knowledge about why survey questions are not answered accurately, or as intended by the author of those questions, draws from an entirely different body of knowledge. Much of this work is guided by theories of human cognition developed within the discipline of cognitive psychology (e.g., Eysenck 1987; Jabine et al. 1984). Statistical theories offer relatively little help in understanding why certain wordings of questions and orderings of response categories influence answers, just as theories of how humans process information in responding to questionnaire items offer no insight into, for example, the relationship between sample size and sampling error.

Concepts associated with understanding and explaining nonresponse, and ultimately nonresponse error, can draw somewhat from cognitive psychology, but do so mostly from yet another body of knowledge. Not as well developed or integrated

as statistical theories appear to be, or even as cognitive psychology strives to be, this area of theory is a combination of influence. Theories of social exchange, cognitive dissonance, foot-in-the-door and reciprocal obligation have variously been used to organize and articulate procedures for improving response to surveys (e.g., Dillman 1978; Goyder 1987; Groves et al. 1992). These theories seek to combine concepts of how people are individually motivated (psychology) with the influence of group affiliations and shared values (sociology) in guiding efforts to reduce nonresponse.

Individuals who design surveys are typically trained in statistical science but not behavioural science. I do not expect this situation to change significantly – statistical science is very different from behavioral science and the specialization of knowledge is becoming greater. It is here that the differences between the research and operations cultures become magnified in a way that has particularly unfortunate consequences for survey innovation.

For example, the last two years have provided an opportunity to observe numerous working groups, some of them of substantial size, each charged with designing a survey. Typically, some individuals who align mostly with the research culture and others who align with the operations culture are included in these groups. When a report is made on sampling design, questions are seldom raised, and when they are, formulas are discussed and the participation is limited to those with strong statistical backgrounds, especially people from the research culture.

However, when discussion shifts to the way questions should be written, the form of the questionnaire, or the procedures that should be used to improve response, virtually everyone enters the discussion. For example, in a discussion of designing a prenotice mailing, a suggestion like this might be made, “When I open my mail at home I read some things and ignore others – I’d read a postcard, but wouldn’t take time to open a letter; consequently, I think we should use a postcard.” Reasoning based on personal experience has been substituted for reasoning derived from the available body of relevant scientific knowledge.

A response to such a recommendation, appropriately based on relevant theory, and in this case drawing from several theoretical orientations, might begin with: “The purpose of a pre-notice is to make people aware that a questionnaire is coming and encourage them to recognize it, open it and begin to fill it out. It takes a little longer to open and read a letter, and that improves the chances that the information will be stored in a person’s long-term memory and, therefore, recalled when the envelope with the questionnaire is received.”

When an argument is couched in personal experience or common sense terms, other arguments that have little to do with reducing nonresponse error but everything to do with how government questionnaires ultimately get printed, packaged, sent, and processed are given equal or greater weight in design decisions:

A larger space for the mailing label will allow placing more numerical information on it to facilitate mailing and processing.

More prominent source codes and key punch instructions will allow data entry to be done more efficiently.

Sending the questionnaires by bulk-rate postage will save money.

It is more efficient not to put mailing dates on questionnaire cover letters and not to address them individually to people in the sample.

Squeezing questions in on one page is okay even if the respondent is required to follow a rather serpentine path in order to read instructions and answer questions. It is desirable to have the same construction for all mail questionnaires regardless of content so that the same envelopes can be used.

It is okay to double-bank the responses (place answer choices in multiple columns) in order to fit a question on a single page.

In mixed-mode surveys, simply giving the mail questionnaire to interviewers to read is okay.

One of the major reasons that government surveys are often disorganized in appearance, questions are not clearly communicated, and mailout procedures sometimes less effective than they could be is that the potential scientific base for such decisions is mostly ignored or "squeezed" out of consideration in the design process. A major reason for this state of affairs is described below.

The core value system of the research culture in many major government survey organizations is statistics. Professionals trained in sciences whose concepts define measurement and nonresponse error issues are few in number and generally lack influence in the design process. The give and take of working groups result in statistical decisions becoming the province of the research culture, while matters related to measurement and nonresponse often get turned over to individuals who represent the operations culture.

Becoming successful in an operations culture requires well-developed skills that have almost nothing to do with reducing measurement and nonresponse error. Organizing large numbers of people to get tasks done accurately, on time and at a low unit cost are deemed the more relevant skills. Success in the operations culture does not require familiarity with research specializing in nonresponse and measurement issues. Therefore, it should not be surprising that such critical error issues can be ignored in the design process.

The result is that while measurement and nonresponse issues have emerged as increasingly important sources of data collection error, there has not been a corresponding emergence of significant numbers of professionals to design theoretically-based projects needed to assure the development and implementation of appropriate innovations for resolving those concerns.

The lack of concern about how questionnaires affect respondents is especially evident for self-administered questionnaires, many of which seem devoid of respondent-friendly qualities. That characteristic is no longer surprising to me. Much of the historical tradition of most survey agencies has been to do personal interview surveys, and thus, when some survey forms have been switched to self-administered, the tradition of placing questions on the left-hand side of the page and answers on the right has been kept. Similarly, the tradition of using the difficult format of requiring respondents to match headings of rows and columns, a skill which interviewer training provides to interviewers, persists (Jenkins and Dillman. In Press).

#### 4. Division of Labor and Hierarchy

Attempts to resolve the competing pulls of operations and research are frequently made through the application of procedures that can be described as hierarchical decision-making.

Previous study of bureaucratic theories of organization and teaching graduate courses on principles of social organization did not prepare me for understanding the hierarchical complexity of government. That initiation started by talking with a person who I thought had authored a past memo about results of research projects, only to find out that it had really been authored by someone else – and this individual simply *issued* it. Only by learning to decipher the distribution lists for memos could one find the name of the original author. Government bureaus consist of a large number of divisions, branches and sections; the organization is highly complex, with many different tasks. One soon learns that section heads report to branch chiefs, who report to assistant division chiefs, who report to division chiefs, etc. Communication flows more easily vertically than it does across divisions, where individuals of similar rank seem most able to communicate.

More importantly, though, hierarchy extends beyond the agencies as well. I came to the realization that any experiment designed in the Census Bureau's 2000 Research and Development Program might be commented on by a minimum of eight entities outside the Census Bureau, each of whose comments should be carefully heard. Organizational hierarchies, no matter whether inside or outside government, have ways of resolving differences of opinion, whether from operation or research cultures or other opposing viewpoints. A matter may be decided based upon who is most powerful, or who won last time, or what the consequences will be for other activities of each division. Compromises are normal, with dynamics added by occasional consideration of what a friend described as the choice of, "Which hill are you willing to die on?"

Common sense decisions that emanate from organizational hierarchies are often predicated on very different considerations than the dilemma which produced the question in the first place. In particular, it is here that measurement and nonresponse issues are decided, losing out to operations issues on the one hand and statistical issues on the other.

In the end, I have no quarrel with the existence of a certain amount of hierarchy. It helps enforce accountability, and I find it hard to imagine how change might be achieved without some degree of hierarchical order. In government, with changing political leaders and changing political climates, hierarchy may have importance that is less understood and appreciated by those who work in universities or smaller organizations.

From the standpoint of innovation in a rapidly changing technological environment, though, hierarchical processes make such cultural and value system concerns more difficult to resolve. The down-side of hierarchy for innovation is that it forces large amounts of critical information upwards through a series of smaller and smaller funnels. This process is slow, but the information that eventually gets through represents only a part of the original message. In addition, the information that is

finally communicated may bear very little resemblance to the original message that was pushed upwards through the multiple administrative levels.

I am reminded of the childhood game in which the first person in line whispers a message to the second person, who whispers it to a third person, and so forth. When the last person in line announces the message for all to hear, it often bears little resemblance to the original message. Hierarchical organizations try to keep messages confidential, in effect whispering them upwards and, as they are passed from one level to the next, distortion is inevitable. When the message is about innovative ideas not well understood by those at higher levels, the problem of distortion is particularly great. One newcomer to a government organization was led to ask, innocently but perceptively, "Why is it that most of the important decisions usually get made at least two levels above those who possess the relevant scientific knowledge?"

A related problem is that whenever a significant conflict develops between organizational divisions, a tendency exists in strongly hierarchical organizations for communication to be forced vertically and for meaningful horizontal communication (across divisions) to be discouraged. The manifestation of this phenomenon is illustrated by what happens as people learn that a conflict exists and a memorandum is being prepared. Lower level workers write statements to higher level people who then write to the next level, etc. I once observed work on a project come to a complete standstill for three weeks while a memo was being formulated to object to a decision of a working group. When the memo finally came with attachments, the package was more than 50 pages long. In the information age in which we live there *has* to be a more efficient way!

Hierarchy, as it is expressed in some government survey organizations, also has a direct deleterious effect on the early development and dissemination of innovative ideas for discussion. Because memoranda are used mostly to issue a final opinion or decision under an administrator's signature, the use of memos to articulate an original "what if . . ." idea in writing is discouraged. Further, because draft memoranda follow a vertical path to sign-off, even if a "what if . . ." idea were to be expressed in writing, it too would be inclined to follow such a slow vertical path. As a result, the horizontal flow across divisions of innovative ideas and, therefore, the promotion of active discussion by all relevant parties at a very early stage is discouraged by the hierarchical orientation of employees. When newly developed ideas for solving problems do flow horizontally, it is usually by word-of-mouth, where they inadvertently become distorted and misunderstood in the same way that unwritten ideas get distorted when communicated verbally up and down the normal decision-making hierarchy.

In this way the potential innovative power of the word processor and electronic mail as generators and disseminators of innovative ideas or discussion thus remains unrealized.

In a stable, nonchanging environment, hierarchy has distinct advantages. Years of experience accumulate understandings of how each facet of an organization functions and can often identify indicators of problems well before they are visible at sub-unit levels. In a dynamic environment where new technologies are forcing change to occur in pursuit of both efficiency and effectiveness, and the best understanding of those technologies often rests with younger, lower ranking individuals, the effects of



hierarchical processes is to slow down and sometimes thwart altogether needed innovation and change.

## **5. Summary of the Three-Part Problem**

To summarize, the conduct of government surveys is often thwarted by three interconnected problems. First, two cultural orientations of professionals in large statistical agencies exist side by side, both of which are essential to conducting surveys important to the nation – they are the research and operations cultures. The research culture by its nature is the one most frequently looked to as the source for defining and experimentally testing potential survey innovations. The operations culture is, by its nature, the one most frequently looked to for assuring that successful large-scale surveys are accomplished on time and within budget. People who subscribe to each of these cultures often bring very different goals to the work groups charged with testing innovative ideas, which I have described as practicing and rehearsing on the one hand vs. formal experimentation on the other.

By tradition and by numbers, the research culture is dominated by the discipline of statistics, whose theoretical expertise rests primarily in the area of sampling error and general statistical design. Measurement error and nonresponse error – the theoretical bases of which reside in cognitive psychology, influence psychology, and sociology – generally are not well represented in either the research or operations cultures, especially at higher administrative levels. Although these skills are clearly represented in the research cultures by staff in some agencies, these staffs are typically far too small to muster the needed influence within the research culture for dealing adequately with nonresponse and measurement error issues. Professionals with these skills and orientation are almost entirely lacking in the operational divisions.

The detrimental and unfortunate result is that in the inevitable give-and-take that goes into designs and tests, matters of measurement and nonresponse are relegated to a considerable degree to being matters of “procedure” more appropriate to the domain of those who subscribe to the operations culture. There, often stripped of their theoretical underpinnings and no longer seen as matters of critical survey error, they become trade-offs against matters considered of equal or greater importance, e.g., postal procedures, per-unit mailing costs, and so forth, by those who subscribe to the operations culture.

The hierarchical nature of government acts as an additional enforcer of this process in multiple ways, including personnel and printing policies that are largely outside the control of any individual agency. It also encourages issues to move slowly up and down levels of an organization, where proposals for change run a gauntlet biased towards the status quo. Other things being equal, individuals who were not in the room when the nature of a possible innovation was discussed seem likely to find current procedures and policies more attractive than alternatives, which themselves require changes in other parts of an operating organization.

## **6. Towards a Solution**

A first step towards a meaningful solution to these interconnected issues that thwart

needed innovation is to bring into government agencies significant numbers of professionals with training in the theories relevant to defining, identifying and resolving measurement and nonresponse error issues. For research to positively affect government survey practice, much of that research must be done on government surveys. At present, the professionals are not there in sufficient numbers to make a big difference. Furthermore, the organizational structure tends to leave little doubt as to the priority of errors. When I came to Washington, D.C., I was reminded that the descriptions of error sources tended to divide the world into sampling and nonsampling error. Whenever something is categorized as “non,” that usually implies not well defined or understood.

We have reached the time in survey history when both measurement and non-response error need direct and focused attention by more than a few individuals. Much of the progress that needs to be made in, for example, designing mixed-mode surveys, utilizing telephone surveys for a public increasingly inclined not to respond willingly by that means, and bringing new survey technologies that may lack inherent respondent-friendliness (e.g., responding to recorded messages by punching answers into a touchtone phone) on-line depends on it.

A second step is to build the capability for understanding and working to overcome measurement and nonresponse error into the *operations*, as well as the research, culture. This is not to suggest that efforts to deal with measurement and nonresponse belong mostly in the operations divisions. However, decisions that are made at operational levels are the ones that most directly affect measurement and nonresponse. Examples include: question wording and order, the way a form or questionnaire is constructed on a page, the information included as part of an address, the class of postage used, whether letters include dates, the contents of those letters, whether letterhead stationery is used, and the kind of mail processing equipment purchased. Unless an understanding of measurement and nonresponse issues permeates the operations culture, I suspect these aspects of error will continue to be traded off to per-unit mailing cost and other considerations that completely ignore considerations of survey error. The avenues for accomplishing this include prioritizing in-service training, hiring people with previous training in these areas, and building research capability and interest into the operational divisions of the organization.

The second step is especially important with regard to those parts of national survey organizations charged with testing and acquiring new computer and other information technologies. The criteria applied to deciding within the operations culture what advancements in such technologies are important tend *not* to be concerned with the reduction of survey error. Rather, they are likely to be influenced most by per-unit interviewing or printing costs and other efficiency considerations that are separate from measurement and nonresponse error considerations.

The acquisition of new computer technologies is one of the main driving forces behind innovation in government survey organizations. Therefore, it is especially important that professionals with expertise in reducing measurement and non-response error be located in these units so that new technologies are likely to help mitigate rather than confound efforts to reduce measurement and nonresponse error. Just as survey innovation aimed at reducing all aspects of survey error is too critical to

rest solely on the individual judgments of statisticians or cognitive psychologists, or sociologists, the adoption of new information technologies is too critical to have such decisions rest solely with technologists.

The third step is to deal effectively with organizational structure, or the problem of hierarchy as I have referred to it here. It is useful to remember that the reasons for building organizational structure, originally, represented an effort to develop an effective communication device that would facilitate accomplishing a task which requires the efforts of a large number of people. Communication needs to be two-way, and it needs to be reasonably fast. At one time, hierarchical command structures were the best available means for accomplishing that in large organizations.

However, modern day information systems make the communication function of bureaucratic organizational structures obsolete. There are faster, more efficient ways of gathering intelligence and distributing information that employees need. In fact, I sense that many-layer organizations tend to thwart effective communication more than they facilitate it. The central tendency of e-mail is to let everyone know everything. The central tendency of multi-layer hierarchies is to limit information to fewer people via a process that results in the original intent of a concern inevitably becoming distorted or subordinated to other concerns. Another result is for an intolerably long period of time to elapse between the identification of a problem and its resolution.

The issue of hierarchy in government survey organizations has two facets. One of them is the separation of research and operations cultures into different, often competing, divisions, so that disagreements must invariably get played out at higher levels. The second issue is how to link them together so that one does not entirely dominate the other. A government survey organization that allows either the research culture or operations culture to control the other will neither be innovative in an effective way nor will it conduct, in the long run, high quality surveys. The organizational structure needed is one that encourages each to influence the other and allows disagreements to be worked out quickly, at lower levels under an umbrella of shared purpose.

This concern is the reason that many organizations have developed matrix management. For example, the dollars to conduct a project are assigned to a project manager, who then recruits people from appropriate home divisions to staff it. Personnel get evaluated in their home division for professional competence and by the project manager for contribution to the success of the project. Therefore, dual accountability exists.

Organizational structure is a far more complicated issue than can be dealt with here, but in my view a core one that simply must get attention. We are trying to design 21st century surveys with organizational structures whose intellectual heritage stems from the 19th century. Giving serious and sustained consideration to concepts described in such books as *The Virtual Corporation* (Davidow and Malone 1992) would provide a critical beginning.

A fourth step towards resolving the concerns I have raised is to increase people's understanding of the multidimensional nature of survey error. For example, just as cognitive psychologists need to understand the fundamentals of sampling error,

statisticians need to understand the fundamentals of measurement error. Professionals who identify primarily with the operational culture need to understand and appreciate these fundamentals just as members of the research culture need to gain an understanding of the special operational difficulties associated with doing large-scale surveys. A noteworthy effort to provide cross-disciplinary training programs in the theoretical bases of multiple sources of survey error is the National Science Foundation supported University of Maryland/University of Michigan Joint Program in Survey Methodology.

Finally, I think it is important to place the ideas expressed here into the larger context that has influenced development of this presentation. The general nature of the issues and solutions that I have discussed here are *not* unique to government survey organizations. Universities, large corporations, and others all find themselves struggling with how to facilitate needed innovation, rather than unnecessarily thwarting it. Neither should the specific comments and examples be viewed as any sort of condemnation of the tremendous accomplishments of our nation's statistical system, which I greatly admire. Rather, they are intended as a discussion of the realities thrust upon us by the information age with its technological and social imperatives. The success of our government survey organizations in meeting the data needs of the 21st century depends on responding to these concerns.

## 7. References

- Davidow, W.H. and Malone, M.S. (1992). *The Virtual Corporation: Structuring and Revitalizing the Corporation for the 21st Century*. New York: Harper Business.
- Dillman, D.A. (1978). *Mail and Telephone Survey Methods*. New York: Wiley-Interscience.
- Eysenck, M.W. (1987). *A Handbook of Cognitive Psychology*. Hillsdale: Erlbaum.
- Goyder, J.C. (1987). *The Silent Minority: Non-Respondents in Sample Survey*. Cambridge, U.K.: Polity Press.
- Groves, R.M., Cialdini, R.B., and Couper, M.P. (1992). Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly*, 56, 475–495.
- Groves, R.M. (1989). *Survey Errors and Survey Costs*. New York: John Wiley.
- Jabine, T.B., Straf, M.L., Tanur, J.M., and Tourangeau, R. (eds.). (1984). *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*. Washington, DC: National Academy Press.
- Jenkins, C. and Dillman, D.A. (In Press). Towards a Theory of Self-Administered Questionnaire Design. Lyberg, L. et al. (eds.) *Measurement Error and Process Quality*. New York: John Wiley.
- Rogers, E.M. (1981). *Diffusion of Innovations* (3rd edition). New York: The Free Press.

Received November 1994