Journal of Official Statistics, Vol. 12, No. 2, 1996, pp. 191–197

Rejoinder

Don A. Dillman

Reading these excellent, but quite varied, commentaries produced two immediate reactions. My first thought was one of being very impressed with the detailed and thoughtful analyses. Most of the commentators have provided important leadership to government statistical agencies and have much experience with factors they believe to have influenced survey innovation. Their comments are rich with insights into how statistical agencies from Australia to The Netherlands, and New Zealand to Canada deal effectively with the need to innovate.

My second reaction to the nearly eighty pages of comments was to immediately reread my original article to answer the question of "Did I really say (or do) that"? Had I really painted a "...rather gloomy picture of government survey statisticians, their organizations, abilities and attitudes" (Colledge) or said that "vertical organizations are bad" (Plewes). I was also surprised by Fellegi's conclusion "[Dillman] attributes an entirely negative influence to management..."

These comments, and some others, are at variance with the conclusions I believe I reached in the article. However, the strong statements of some commentators suggest that my analysis may have touched upon important but controversial issues, thus encouraging the articulation of well-formed opinions on concerns important to the future of statistical agencies. I hope that the result of this exchange will be to open a much needed discussion about the difficulties of achieving innovation in government statistical organizations. And, I welcome the opportunity to extend that discussion somewhat further in this response.

Some commentators provide opposing perspectives on the same issue. For example, Plewes argues " ... that my case for the need for change is weakened as much of that change is already taking place" and uses the redesign of the U.S. Current Population Survey (CPS) as his supporting example of "an amazingly innovative result." Fienberg and Tanur also offer the CPS as a "shining example of the innovation possible in government statistical agencies."

Bryant, on the other hand notes her amazement upon coming to the Census Bureau as director in 1990 to discover revisions to the CPS questionnaire being made "*for the first time in a generation*" (emphasis hers), even though, everyone involved with CPS had recognized for years that the wordings of some questions were completely out-of-date, resulting in significant measurement error. The time and effort required for revision of the CPS was, in fact, one of several events that encouraged me to contemplate why innovation in government surveys typically takes so long to accomplish, and to write this article.

Bryant also notes "Any survey organization other than a government one would

have revised the (CPS) questionnaire a number of times in the generation time period." Hoffman, on the other hand, argues that innovation is no more difficult in government survey organizations than in nongovernment ones. Fienberg and Tanur appear to agree with Hoffman by citing Fienberg's experiences as a senior academic administrator in a large university where he found researchers reluctant to embrace change.

Disagreement among the commentators is not surprising. They write from varying perspectives and collectively have, I estimate, at least 400 years of professional experience, much of which has been in high level administrative positions.

Several of the authors note correctly that my listing of barriers to innovation is incomplete, as are my proposed solutions. Bryant adds four additional barriers – lack of cross-fertilization with outside thinking, the tendency to avoid small scale experiments in favor of large scale tests which tend to inhibit needed experimentation, long budget cycles, and caution over changing measurement procedures, regardless of their validity. Tortora adds the lack of an agency policy on how change should occur. Kincannon points to the imbalance between the penalty for failure and the reward for success, noting that the failure of a change is likely to attract an inordinate amount of negative attention.

On the solution side, Tortora, Clark, and Schaible argue for rotation and training of personnel and the bringing in of new leadership, ideas with which I strongly agree. Trewin goes further, offering an insightful analysis of organizational culture that deals with leadership, communications, understanding one's market and the need for competition. Fellegi discusses the importance of interdisciplinary teams for fostering innovation, at least in the case of Statistics Canada. Bailar, Bethlehem, and Binder discuss the importance of high level discussion and support of change, and in particular statisticians who "champion" measurement and nonresponse issues.

On the whole I have little disagreement with the addition of any of these factors influencing innovation. Had I been writing a longer article I would have undoubtedly discussed some, though not all, of them. In the two years that have elapsed since the article was first drafted, I have come to appreciate in particular the importance of high level sponsorship. In this regard a retired statistical agency administrator pointed out to me that when an administrator gains a reputation for continually asking about an issue, such as measurement error, that tends to generate a process by which lower level administrators seek information and hold discussions they would otherwise simply ignore. However, such support is not sufficient.

Without the existence of people positioned at high levels in statistical agencies who are trained in the social sciences and can therefore articulate the scientific basis for strategies aimed at reducing measurement and nonresponse errors, progress in dealing with many issues outlined in my article will be at best uneven. This concern brings me to the one disappointment I felt upon reading this provocative and informative set of commentaries. I did not get the sense that more than one or two of the authors recognized a need to bring into government statistical agencies social scientists trained in the concepts and theoretical orientations relevant to reduction of measurement and nonresponse errors so that the science associated with that work could advance, as has been the case for basic statistical research.

Bethlehem is the most negative, and candid, on this issue: "I do not believe in cognitive psychology, or any other science, providing solutions for these problems [errors other than sampling]. Useful models for psychological or social phenomena with sufficient explanatory power simply do not exist." This reaction is similar to feedback I received immediately after the article was written, usually stated cautiously as a question: "We know what the science is that makes sampling work, but is there really something scientific that can be done on the measurement and nonresponse issues you are so concerned about?" My response in these conversations was always the same, "Yes, but the concepts, propositions and theories are not from statistics; they come from other disciplines such as psychology and sociology".

Usually I went on to point out multiple references that developed some theoretical aspect of such research and noted that theoretical development and testing on these types of errors, though plentiful, was far from finished[1]. The vast majority of this work has or is being done by scientists whose predominant training is in a discipline other than statistics or mathematics.

I do not quarrel with Bailar's contention that "... at the bureau it was statisticians who brought in other types of research.... From the earliest days at the Census Bureau, statisticians have encouraged, funded, and provided a framework for the study of measurement and nonresponse errors.... Thus I find Dillman's argument that statisticians are primarily interested in sampling error not to have been the case... "

And, I agree heartily with Bailar's conclusion that for innovation to be accepted quickly, strong champions are needed. My concern is that there is a huge difference between being "interested" in an area of inquiry or seeing the need for it on the one hand, and having adequate training in the theories and concepts that will allow scientific work in that area to advance and accumulate. This is not a devaluation of what statisticians can do; it is simply a recognition that for an area of inquiry to advance scientifically with solid theory guiding the development of experiments and other research, a critical mass of scientists with appropriate training is needed. Depending upon outsiders to do such work, as suggested by the Joint Statistical Agreements or the Fellowship programs which Bailar mentions, almost guarantees that such work will be subordinated to what is defined as mainstream statistical work. In government survey organizations, "consultants" are not usually present when major decisions get made. And, they lack the continued presence needed to follow through on issues that others are attempting to implement without a thorough understanding of the concepts underlying them.

Others seem to echo Bailar's defense of statistics as the discipline that really counts! Binder notes agreement in principle that measurement and nonresponse error issues should be studied, but "... we must ensure that their advice is relevant... " and suggests that the status quo be maintained, i.e., modest expenditures, but "greater

[1] A few of the hundreds of useful references on the understanding of respondent behavior that include diverse theoretical arguments include: Schuman and Presser (1981), Tourangeau (1984), Groves, Cialdini, and Couper (1992), Goyder (1982), Dillman (1978), Schwarz and Sudman (1988), Jenkins and Dillman (In Press), Sudman, Bradburn, and Schwarz (1992), Schaeffer (1995), Groves and Couper (1995), and Schwarz, Strack, Hippler, and Bishop (1991).

than zero." So long as measurement and nonresponse issues are relegated to something for which the "champions" of these concerns, in which category I would place many of the commentators, are the principle decision-makers I suspect that allocations in most statistical agencies will be "greater than zero" but small, and the advancements also meager.

An example of viewing measurement and nonresponse research in an inappropriately, even nonscientific way is the tendency I have observed among some excellent statisticians who are responsible for research on measurement and nonresponse issues to ask "Can't you do a few cognitive interviews or focus groups and tell us how people will respond so that we don't have to run a field test?" It is true that a dozen or so such interviews can identify certain measurement problems with specific questionnaires. However, they cannot estimate from those interviews the incidence of a problem in a study population, leaving the researcher to guess whether something is likely to be a big or a small problem, and providing little information about how correcting the problem for some will result in problems for others. Focus groups often create a reality of their own as people take one another into account in the registering of opinions. Such groups have sometimes produced advice that is dramatically different, even the reverse, of people's actual behavior[2]. Cognitive interviews and focus groups can help develop good hypotheses, but are no substitute for field experiments. I have observed the tendency to want to rely on a few cognitive think-aloud interviews or focus groups and thereby avoid field tests in statistical agencies outside as well as inside the United States. If a scientific body of knowledge is to develop that will allow the correct predictions of the occurrence of measurement and nonresponse errors and solutions, then experiments, field tests, and basic theoretical work must all be supported.

It is important for all of us to recognize our intellectual heritage and in this case the ground-breaking contributions of statisticians, as noted, for example, by Fienberg and Tanur, Fellegi, and Bailar to opening-up and championing research on measurement and nonresponse error. It is also important to recognize how science usually advances. People within a discipline realize that their own theories and concepts cannot account for a seemingly important phenomenon or data discrepancy. As good scientists they attempt to find people who can help them understand it, which is exactly what the pioneers listed by Bailar did. The problem comes when beliefs are formed of the nature: "The theoretical basis underlying measurement error and response error is ... more intuitively accessible to nontechnicians than is that of sampling error" (Kincannon) with the result that it is assumed that nontechnicians and statisticians have the training to provide leadership for whatever research needs to be done. This tendency leads to the related one described in my article of resorting to personal experience as an equivalent substitute for scientific research. Needed innovation in resolving measurement and nonresponse errors in surveys is unlikely to be accomplished until we recognize that people cannot achieve the necessary theoretical skills by work experience while ignoring the conceptual and theoretical underpinnings, as is so often done now.

[2] For example, focus group tests of possible motivational messages for use in census questionnaire tests produced a nearly unanimous recommendation that a "benefits" message should be used rather than one that emphasized that response was mandatory. A subsequent field test revealed a dramatic 10% gain for the mandatory message and no effect from the benefits message. (Dillman, Singer, Treat, and Clark, In Press.)

Another important theme addressed in several of the commentaries is the influence of the research vs. operations culture. No one seems to dispute the existence of the divergent operations and research cultures, which was a goal of my article to describe, or the importance of these concepts. However, Binder correctly points out that I have ignored other distinctive subcultures, notably subject matter and information technology specialists who bring still other perspectives to bear on innovation.

Bethlehem describes the basis of the resistance of production departments as being "... trained to keep existing systems running ... not ... to think about redesigning their processes ... innovation may lead to different work, or no work at all." Miskura argues that the concerns of both the research (on measurement and nonresponse considerations) and operations cultures are legitimate and that stating hypotheses, interpreting results and suggesting solutions should not be the sole domain of the researcher. I agree! But, to do so calls for researchers to gain a greater understanding of and appreciation for operational issues, rather than orienting their work just to their peers. Similarly, operational employees must understand and seek to apply scientific knowledge of how and why nonresponse and measurement errors occur; the time has passed when an operational employee (whether questionnaire designer or implementation planner) could simply ignore scientific underpinnings or consider such knowledge as someone else's responsibility. Just as the information era demands that production-line workers in factories be computer and mathematically literate, production workers in statistical agencies must understand the conceptual underpinnings of their work. This need calls for a dramatic change in hiring and training practices.

I do not agree with Colledge when he says "Pure research belongs primarily in universities ... and it is out of place within a survey organization." Nor, as he suggests, do I identify myself principally as a pure researcher. In survey research I have not found such a distinction useful. To do so invites university researchers to ignore the reality of empirically testing their ideas in actual surveys, and invites agencies to ignore theoretical guidance while continuing to base survey design decisions on personal experience and opinion, e.g., "How I would respond if this cover letter was sent to me with a questionnaire." That, in my view, brings us back to one of my compelling reasons for writing the article in the first place: The difficulty of achieving innovation in an environment that has been slow to accept the best that social science has to offer.

It was surprising to me that the concerns I raised about hierarchical management evoked relatively few comments. Miskura noted that managers tend to get what they reward, and Bethlehem suggested that hierarchical organizations have an advantage when management feels a change is important. Tortora described another aspect of the hierarchy problem which I did not develop noting the frustration felt by high level administrators when they made decisions but were unable to make them stick. That too is a symptom of the dysfunctions of hierarchical management. The insulation of many layers of management each with carefully prescribed supervisory duties makes innovation harder to achieve for both high and low level employees; the system itself works against innovation.

Kincannon notes that he is less concerned about the hierarchy problems than I, and

perceptively offers as a reason the difficulty in tackling them directly. He also offers some optimism that changes in technology of communication and theories of management will, "...continue to whittle away at the information rigidities and even the decision functions of hierarchy." In addition he predicts that, "...the great, crashing budget debate in the United States will flatten hierarchical structures..." I sense that his prediction may be much closer to reality than most thought possible only a year ago, and that what once seemed impossible to attack directly may suddenly be subject to direct action. In this context it has probably never been more important that agency administrators, at least in the U.S., understand the concepts of matrix management and other modern management concepts mentioned by both Kincannon and Tortora.

Overall, as I reflect on the circumstances which prompted me to write the original article, informal feedback on the article (which has been substantial), the insightful commentaries published here, and my government agency experiences since then, my main conclusion is what some may see as dreadfully academic. I wished I had written a longer article! Had I done that I would have incorporated some of the commentators' ideas, and explained my position on the original ideas somewhat more fully. The article's main ideas about the challenge of managing competing organizational cultures, value system differences, and the difficulties that will persist for hierarchically-oriented organizations in the information age strike me as being as relevant now as when the article was written.

Finally, three of the commentators reminded me of the advice given to me by the publisher of my first book, "Once you think you have written a particularly fine phrase, immediately go back and strike it out"! My references to the role of engineers, pilots and flight attendants in designing aircraft undoubtedly fall into this category as appropriately commented on by Tortora, Kincannon and Fienberg and Tanur. I concur that surely there is a better way to characterize the difficulties of innovation in government survey organizations, the importance of and urgency for which have never been greater.

References

Dillman, D.A. (1978). Mail and Telephone Surveys: The Total Design Method. New York: Wiley-Interscience.

Dillman, D.A., Singer, E., Clark, J., and Treat, J. (1996). Effects of Benefits Appeals, Mandatory Appeals and Variations in Statements of Confidentiality on Completion Rates for Census Questionnaires. Public Opinion Quarterly (In Press).

Goyder, J. (1988). The Silent Minority. United Kingdom: Oxford University Press.

Groves, R.M. and Couper, M.P. (1995). Theoretical Motivation for Post-Survey Nonresponse Adjustment in Household Surveys. Journal of Official Statistics, 11, 93–106.

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.

Jenkins, C.R. and Dillman, D.A. (In Press). Towards a Theory of Self-Administered Questionnaire Design. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo,

N. Schwarz, and D. Trewin (eds.), Survey Measurement and Process Quality. New York: Wiley-Interscience.

Schaeffer, N.C. (1995). A Decade of Questions. Journal of Official Statistics, 11, 79–92.

Schuman, H. and Presser, S. (1981). Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context. New York: Academic Press.

Schwarz, N., Strack, F., Hippler, H.-J. and Bishop, G. (1991). The Impact of Administration Mode on Response Effects in Survey Measurement. Applied Cognitive Psychology, 5, 193–212.

Schwarz, N. and Sudman, S. (1992). Context Effects in Social and Psychological Research. New York: Springer-Verlag.

Sudman, S., Bradburn, N., and Schwarz, N. (1995). Thinking About Answers: The Application of Cognitive Process to Survey Methodology. San Francisco, CA: Jossey-Bass.

Tourangeau, R. (1984). Cognitive Sciences and Survey Methods. In T. Jabine, E. Loftus, M. Straf, J. Tanur, and R. Tourangeau (eds.), Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines. Washington DC: National Academy of Science.

Received January 1996