

PRESIDENTIAL ADDRESS

TOWARD A NEW RESEARCH RELEVANCY: WINNING OUR SEAT AT THE TABLE

MARK A. SCHULMAN

At a professional meeting that I attended late last year, a participant asked me, “What to you is the single best way to judge the quality of a survey?” I was, frankly, dumbstruck about how to answer such an overarching question. He opined, “To me, the single best measure of a survey’s quality is the response rate. Don’t you agree?”

When I disagreed passionately with my colleague’s assertion, I felt compelled to start a dialogue today about what is a “job well done.” I chose this topic because I fear that we risk marginalizing our profession when we become too narrowly focused on reaching certain survey research benchmarks, such as response rate calculations and standard errors, and not focused enough on providing the most useful strategic guidance to policy makers and other clients. I fear that too often we have become technicians who toil in the back rooms and have relinquished our seat at the decision makers’ table. Our society is much the poorer for it.

One key reason for our profession’s misplaced emphasis, I believe, is that we suffer from what Robert M. Groves (1989, pp. 32–33) and others have labeled “the tyranny of the measurable,” or, to rephrase slightly, “the tyranny of the ‘easily’ measurable.” I want to argue today for a rebalancing of our profession so that we recognize the limitations of “response rates” and other simple measures of survey precision and quality, while, more importantly, placing more weight on our profession’s insights and contributions to social and public policy decision making.

We certainly do need to be sure that we cover the technical bases of our surveys and base our conclusions on solid data. I want to emphasize the tremendous contributions that our survey methodologists have made in advancing our survey science. But we need to go a step further: Do we understand fully the dimensions of the problems we are seeking to inform? Do we un-

MARK A. SCHULMAN is president of Schulman, Ronca, & Bucuvalas, Inc. The author gratefully acknowledges thoughtful comments on earlier drafts of this speech by Don Dillman, Peter Miller, Stanley Presser, and Thomas Guterbock.

derstand the everyday life contexts of the people whose responses we are gathering? Do we frame our studies and our analyses in such a way that they have relevance for decision makers?

In short, I feel that we have reached a crossroads. We need to ask ourselves whether we should be technicians or problem solvers. For example, do we as researchers have the experience and vision to understand:

- How people's underlying functional needs might be better served?
- What constitutes policy and service satisfaction?
- What unique challenges face various segments of our population now and in the future—segments that include the elderly, the disabled, the poor, children, families, single-parent households, and so on?
- How we might better anticipate future needs using our research methods?

Let me emphasize that my perspective is that of an applied researcher. My main criteria in evaluating a study are these. First, did the study inform a decision-making process and provide insights into the issues? And, second, were the results actionable?

Sometimes I feel that we at AAPOR become so focused on methods that we neglect the reasons we conduct surveys: to support decision making. Hence, I am seeking to present my personal perspective, while respecting the skills and accomplishments of the methodologists who have advanced our profession. To start my dialogue today about “what is a quality study,” I want to do three things: tease out both the origins of this “tyranny of the easily measurable” and its consequences on how we conduct our research; attempt to shift our focus toward relevance and problem solving, in addition to technical and methodological issues; and suggest a shift in how we practice our profession and how we train and socialize the next generation of survey researchers—not just by thinking outside the box but by physically getting outside the box and away from our desks, observing and interacting with real-world people.

The Evolution of the “Tyranny”

Let me trace the evolution of this “tyranny” of response rates and sampling error by noting its simplest manifestation: the ubiquitous reporting of “sampling error” both by the media polls and by researchers generally. Even the esteemed television news anchor Dan Rather will add “. . . and the sampling error is plus or minus 3 percentage points” with a reassuring nod after reporting the poll numbers. The reporting of just the standard errors is still considered a triumph in educating the public and the media about the nature of survey estimates.

Our profession still bears the indelible mark of the more simplistic,

non-mean square error approach. Many, particularly in the media, still report and focus on “sample size” or sampling error as the barometer of survey quality. Many still believe, mistakenly, that increasing sample size is the easiest way to reduce overall survey error. Worse, many Internet surveys of the general public mistakenly rely on reporting so-called sample size, even if their opt-in respondents are from no known population universe and not really “sampled,” which raises coverage error issues as well. We well know that sampling error may be, in many cases, the least important mean square error component. And, importantly, AAPOR’s emphasis on standardizing response rate reporting, while critically important, has had another and perhaps unintended consequence for many researchers: elevating the “response rate” to seemingly *the* most critical factor in judging the quality of their research.

Let me say again, I am not for a moment suggesting that response rates are unimportant to survey quality. I do not want to mobilize the AAPOR response-rate police to hunt me down. I laud the efforts of the AAPOR Standard Definitions Committee, whose work has finally standardized the reporting of response rates after what had been years of chaos and noncomparability in reporting. We now have a “common currency” for comparison among sample surveys. Also, please be assured that I, myself, am extremely demanding on issues of response rates and field procedures.

However, my concern is we have become so focused now on response rates because of our standard measurement that we need to be reminded that response rate reporting is

- blind to actual nonresponse error and bias,
- blind to coverage errors,
- blind to poor question wording, and
- blind to sometimes poor study design that fails to understand the issues under study.

We know that response rates are only an ambiguous and sometimes troubling proxy for nonresponse error and bias, particularly how nonrespondents differ from respondents. Some research suggests that generally used response-rate boosting measures may, in fact, increase overall bias. Groves, Singer, and Corning (2000) found that each response-rate boosting measure might differently affect the mix of persons in our samples, thereby potentially increasing overall error. Teitler, Reichman, and Sprachman (2003, p. 136) make a related point, that the hardest-to-complete cases in their study after extensive and very expensive response-rate-boosting measures were still “poor proxies for nonrespondents.”

Just as challenging to our sense that “higher is better” are the findings of Keeter et al. (2000) in “Consequences of Reducing Nonreponse in a National Telephone Survey.” This study tested the claim that methodological shortcuts taken to collect timely public opinion data bias the results. They used identical questionnaires but fielded the questionnaire by telephone in two different ways:

a “standard” survey conducted over a 5-day period using an “at home” sample, and a “rigorous” field design. Even though the standard 5-day design achieved a 36 percent response rate and the rigorous design achieved a 60 percent response rate, the two surveys produced largely similar results on all items except some demographics, a then heretical finding. In fact, the less rigorous approach produced an unweighted sample a bit closer to census estimates. We have heard some further, tantalizing findings at this conference, particularly by Paul Biemer, which provide some further support that higher response rates may even increase bias.

Finally, I have several longer-term and practical concerns about our push for higher response rates. Again, I feel conflicted in raising these issues because, personally, I use response rates as a key yardstick of survey quality. However, a number of potential problems concern me. First, through the flagrant incentives that some firms, many Internet based, offer people to participate in surveys, are we training the next generation of respondents to expect to be paid to participate? How many of you routinely receive the following types of e-mail? “\$20 Cash to the First 20 to take our Survey Today!”

Fortunately, our pre-Internet research finds little long-term effect from token payments, that is, a few dollars, on whether respondents expect payment for survey participation (e.g., see Singer, Van Hoewyk, and Maher 2000). Clearly, these small incentives do boost response rates measurably with little long-term damage.

However, these studies were done prior to the rise of volunteer, nonrandom Internet surveys, which often pay incentives or treat survey participation as an extension of the Powerball Lottery. Will we find a respondent training effect from the proliferation of these kinds of “fee for participation” Internet surveys?

Oddly enough, we are not even sure of the impact of any of these “after the fact” incentives on response rates. In fact, some researchers argue that these postincentives have little effect on participation but may result in nonsensical answers from people responding to get the money. In any event, we do need to determine whether these types of incentives have a long-term impact on respondents’ willingness to participate in our surveys.

Another practical concern that I have is whether respondents may feel harassed by our intensive refusal-conversion methods, that is, recontacting initial refusals again and again to gain their cooperation. Sometimes we can convert as many as 30 percent or more of the initial refusers.

At one level, this is a very effective strategy to reduce nonresponse. However, there may be a darker side of this strategy: I fear that we may burn bridges to respondents who are doubly annoyed at us and feel harassed as a result of not just one initial contact and one refusal-conversion call but, in some cases, two or more conversion attempts. I can report from personal experience the sometimes angry calls we get from respondents after multiple

refusal-conversion attempts. I believe that we need to tease out the extent to which we may be alienating people in the longer term.

Also, in an era when millions of households are seeking refuge from telemarketers through “do not call” lists, we may risk turning respondents’ ire from telemarketers to survey researchers. Survey research is now exempt from do not call lists. Let us hope that pressure does not build on politicians and regulators to require us to abide by these lists.

In summary, I fear that our sometimes obsession with response rates as the major measure of survey precision may be distracting us by establishing a measurable, but limited, standard. Moreover, I am conflicted about whether we may be doing long-term damage to the goodwill of our respondents, who are already resisting our efforts more each year.

The Cognitive Revolution

One important and constructive step forward and beyond sample size and response rates has been our growing focus on cognitive psychology. These efforts date back at least to the Cognitive Aspects of Survey Methodology (CASM) efforts in the late 1970s and early 1980s. This was also the subject of an AAPOR cosponsored conference last year in Charleston, SC, which many of you attended. Two of our council members, Roger Tourangeau and Jon Krosnick, have been leaders in this area.

The underlying premise of the cognitive revolution is that “reporting errors in surveys arise from problems in the underlying cognitive processes through which respondents generate their answers to survey questions” (Tourangeau 2003, p. 5). In other words, our question wording can be a major source of measurement error. The cognitive psychologist asks not what the question was but, rather, how did the respondent interpret the question and formulate the response. Bradburn (2002) notes that concepts from cognitive psychology, such as primacy and recency, assimilation and contrast, and schemata, are becoming familiar to survey researchers and are used in their work.

With the growing use of cognitive testing, we have made some badly needed progress in broadening our definition of survey precision. This type of testing gets us closer to my goal of better understanding our respondents, their needs, and their aspirations.

However, too often we are on autopilot in how we develop our studies and our questionnaires, particularly when we are under tight deadlines. We review the literature, cull out, and electronically cut and paste what appear to be the relevant questionnaire items from past studies. We all know this drill, right? (By the way, in the old days, we really did cut with scissors and paste with tape!) “Yeah, here are 32 questions about leadership qualities that some grey-haired eminence asked 10 years ago . . . just slap them into the questionnaire . . . that was easy! . . . Instant questionnaire!” Of course those of us who

have been around a few years may remember that the questions did not really work 10 years ago and are even less comprehensible now.

I discovered the worth of cognitive testing early on. I was working on a substance use assessment in certain catchment areas with a group of experienced researchers who had assured the sponsors that we would use a well-known and normed question series to compare our data to the national norms. Because our target population was mainly low-income, low-education people, we did some cognitive testing of these standard questionnaire items. I can only describe our reaction as “startled” as we observed our test respondents stumble through the question series, obviously challenged in comprehending these items: “Well, I do have heart disease, . . . but I don’t think that it’s coronary heart disease.”

The growth of cognitive testing has at least shaken many of us from this lethargy: the sense that if it was good enough for Professor X 10 years ago, it is good enough for me. We are becoming more critical consumers of questionnaire items, thanks to what some have called the “cognitive revolution.”

In short, we need to avoid becoming prisoners of desk research, and not just the cut-and-paste kind. Cognitive testing has been a step forward. However, even asking “better questions” still does not always produce the end result of understanding the underlying issues, as opposed to respondents’ comprehension of our questions.

Training and Orientation

Now, to be a bit more provocative, I have a nagging sense that part of the reason we tend to undervalue research impact over technical and methodological issues is rooted in our traditional training as social scientists. I am not suggesting here that I can definitively describe all of our training programs, which are very diverse in their content. However, let me contrast the approach traditionally used to train social scientists to that commonly used to train market researchers.

When we train social scientists, we often do not emphasize problem solving through research but, rather, the quest for “scientific” theory development and for replication. Our mode of training and our textbooks sometimes position the practice of social research in Parsonian terms. Talcott Parsons, in his widely read book *The Social System* (1951), posits that our research should be guided by the desire to create an almost mechanistic “general theory of action,” similar to the field theory in physics. Earlier editions of the widely used Babbie text, *The Practice of Social Research* (6th ed., 1991, p. 40), for example, closely followed Parsons’s approach in describing the underlying approach of his text as being “grounded in the position that human social behavior can be subjected to scientific study as legitimately as can the physicist’s atoms.” I am pleased

to note that these references have been removed from the latest edition of the Babbie text (9th ed., 2000).

For many social researchers, social research, to be scientific, must be grounded first in theory, then operationalization, and then, finally, observation. That was our training theology for many years. Observation takes a back seat to theory. Again, in older editions of Babbie's text (6th ed., 1991, p. 47), he writes, "The final step in the traditional model of science involves actual observation, looking at the world and making measurements of what we have seen." There is little mention here of what I consider to be the critical first step, observation and problem identification. These earlier text editions devoted little space to so-called field research. Again, I am pleased to report that the latest edition of the Babbie text (2000) acknowledges more fully the importance of observational and qualitative methods in research design.

I know that I am treading on some sacred ground here. I am deliberately being provocative and painting things a bit "black and white" for the sake of argumentation, so forgive me on this. However, let us examine the very different and "inductive" orientation of market research, which uses the same survey methods. A widely used text in market research (Churchill 1987) suggests that problem identification is the most important step in the research process. The market-research text analogizes the role of the market researcher to that of a physician: "A patient who is not feeling well describes his symptoms to the doctor. The patient may or may not be able to diagnose his own case. . . . The creative challenge for the doctor is in the diagnosis. . . . The doctor's responsibility is to cure the patient; the marketing researcher's responsibility is to solve the client's problem" (Churchill 1987, p. 30). Hence, the market researcher applies his methodological tools to solving the problem. And the process is largely inductive in nature, seeking to observe and define problems at stage 1.

The training of the market researcher often involves the case study method, where they apply methodological tools to formulate issues and resolve problems. Students use their analytic tools to evaluate situations strategically and develop and test strategies to address problems and issues. This case study method provides guidelines to aid the student in analyzing a case situation by casting himself or herself in the role of protagonist, developing criteria for alternative decisions, and generalizing to other situations. The *Harvard Business Review* has a regular feature describing case studies. The case study method challenges students by bringing them as close as possible to the business situations of the real world. Do we in the social sciences have a similar pedagogical tool that emphasizes case study? More often than not, in my observation, the answer is "No."

Market research also has something of a yardstick, albeit a somewhat fuzzy one, to measure a study's contribution to the decision-making process: return on investment (ROI) in research. This is the ratio of net profit to investment. In other words, if we invest \$100,000 dollars in research, can we get a \$25

million return based on the study's conclusions and recommendations? While I am not sure that we can always reduce things to dollars and cents in social science research, I do wish that we had a similar metric that would quantify the contribution that we make.

Yes, theory has a role to play in helping us develop our hypotheses and study designs. And I have already warned you that I have painted things very black and white for contrast. In truth, the directional arrows go both ways, from theory to observation and from observation to theory. However, in my applied world, methodological tools become the servants, not the masters. By contrast, in training social researchers, I sense that too often our desire to truly be "scientists" and reduce behavior to the study of atoms has led us to an overemphasis on the deductive, that is, starting with theory, even if the theory does not exactly fit the problem we are addressing.

Curiously, as I wrote this section of my address, I had a nagging suspicion that, rather than setting out a new path for our profession, I was, in fact, advocating that we return to our roots. I had the joy of reacquainting myself with the book that was probably most influential in whetting my appetite for a career in survey research. The late Herbert Hyman's *Survey Design and Analysis*, first published in 1955 and long out of print, entices students with a series of case studies so that they understand that "there are the skills needed on the technical or methodological side, but consider the *substantive* problems for which surveys are undertaken" (p. 32). I hope that we can come full circle back to Hyman's research vision. Of course, if you open the Hyman book, you will also find a refresher course on the use of Hollerith cards and card-sorting machines.

Next Steps

I am calling on us today to rebalance our work and our training programs to emphasize not just the important technical and methodological issues but to focus on problem identification and making our research more relevant to our clients, whether in public policy or in the commercial realm. I want to infuse our students with the sense that, to answer society's issues, we need to move beyond our desks and our libraries and make contact with the real world.

One way that we can accomplish the task is by breaking down the wall between qualitative and quantitative so that we move closer to our respondents' world. I often find qualitative research indispensable not just in providing a better understanding generally of the issues and the study subjects but also in generating theory, hypothesis development, and, postsurvey, illuminating further our survey findings. As Yogi Berra said, "You can observe a lot by watching."

Too often I see researchers pigeonholed as "qualitative researchers" whose work always carries the footnote that it is not projectible to the larger pop-

ulation. From my perspective, qualitative research should be viewed as a critical first stage in this process and not just as a nongeneralizable alternative to survey research. Of course, too often I see qualitative researchers expressing disdain for quantitative research as well, claiming that our survey numbers cannot really undercover their hidden truths. Yet the two types of research approaches should work hand in glove.

The advancement of cognitive testing treats this more qualitative process as mainly a means to improve question wording and respondent comprehension. What I am suggesting goes far beyond improving question wording: gaining a better understanding of the issues and problems that we are confronting in our studies.

Here are some examples, the first using qualitative research prior to questionnaire development. Recently, we were asked to evaluate a model to deliver social services to a specific target group. Before doing the usual client feedback type of survey, I first asked to speak to both the clients and the caseworkers. I discovered from the caseworkers that the model being evaluated was not even in place yet. I found during a roundtable with the clients that they, not unexpectedly, were completely unaware of the changes in the delivery model. In fact, the clients expressed some major reservations about that model that, heretofore, had not been recognized. I am not sure that we would have captured these critical issues had we just launched our survey.

A second case illustrates the value of qualitative research in the postsurvey phase. Recently, working with a group of specialists on the sociology and psychology of religions, we conducted an online bulletin board focus group, free from geographic limitations, to probe more deeply the attitudes of college students whom the researchers have been following through longitudinal panel studies since they were in their early teens. While the results of the panel survey were highly productive, the recent online bulletin board focus group, which gave the students time to consider their responses, provided some very powerful insights into the impact of their college years and subsequent experiences on their religious identity and how they feel about their religious views. Note that this was done postsurvey, not prior. Again, this qualitative, depth research enabled the researchers further to clarify and understand their survey findings, particularly on gaps between the subjects' religious attitudes, on the one hand, and their actual religious behaviors, on the other.

I dare say that if I asked for a show of hands from those of you who routinely build qualitative phases into your research designs, you will find the same level of enthusiasm. These qualitative tools include

- roundtables with respondents, stakeholders, and clients;
- in-depth semistructured interviews;
- focus groups;
- observational studies; and
- Internet-based qualitative groups.

Ethnography, used by anthropologists, is becoming increasingly important in market research. Research teams, typically including researchers and a camera operator to document the research, accompany consumers as they engage in a wide variety of activities, including grocery shopping, eating, using household items, and so on. They combine qualitative and quantitative measurement, though not usually surveys, in their observations. The increasing use of traditional focus groups reflects this desire as well—to stare the consumer in the eye and observe.

In summary, neither qualitative nor survey research has a monopoly on uncovering the hidden truths. They often have a multiplier effect when they are used together. I get very frustrated when either type of researcher claims that they have a monopoly on truth.

In Conclusion

Much of our research agenda and much of our program content at AAPOR is focused on such issues as nonresponse; sample coverage; questionnaire design issues, such as order effects; and cognitive testing. Some researchers have become almost fixated on response rates, which are readily measurable. Clearly, we need to understand the methodological underpinnings of our work. However, we should not worship simplistic benchmarks at the expense of research relevance.

Let us make sure that we devote the same high level of effort to serving decision making in our complex society that we devote to methodological issues. Let us not become a profession that focuses mainly on technical survey issues, but not on the people and issues that we are studying. Importantly, let us also train our students to go beyond tests of statistical significance and focus more on tests of real-world significance and impact.

I propose further that we establish a yearly AAPOR “Made a Difference Award” to recognize studies that have made a notable contribution to public policy in the past 5 years or so. I think that this will send a message about the need for studies to be more than the sum of the tabulated data. Let us make sure that we understand how our research can better support our clients. We want a seat at the table and a voice when decisions are made.

That seminal jazz musician Charlie Parker said, “Music is your own experience, your thoughts, your wisdom. If you don’t live it, it won’t come out of your horn.” Let us open our windows and doors, move beyond our desks and standard question batteries that we cut and paste. Let us develop a clearer understanding of our respondents, their wants and needs, and their aspirations, so that relevant, actionable research will “come out of our horns.”

References

Babbie, Earl R. 1991. *The Practice of Social Research*. 6th ed. Belmont, CA: Wadsworth.

- . 2000. *The Practice of Social Research*. 9th ed. Belmont, CA: Wadsworth/Thomson Learning.
- Bradburn, Norman. 2002. "Address to the 2002 QDET Conference." Paper presented at the International Conference on Questionnaire Development, Evaluation, and Testing Methods, Charleston, SC, November 14–17.
- Churchill, Gilbert A. 1987. *Marketing Research: Methodological Foundations*. 4th ed. Hinsdale, IL: CBS College.
- Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: Wiley.
- Groves, Robert M., Eleanor Singer, and Amy Corning. 2000. "Leverage-Saliency Theory of Survey Participation." *Public Opinion Quarterly* 64:299–308.
- Hyman, Herbert. 1955. *Survey Design and Analysis*. New York: Free Press.
- Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert M. Groves, and Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly* 64:125–48.
- Parsons, Talcott. 1951. *The Social System*. Glencoe, IL: Free Press.
- Singer, Eleanor, John Van Hoewyk, and Nancy Maher. 2000. "Experiments with Incentives in Telephone Surveys." *Public Opinion Quarterly* 64:171–88.
- Teitler, Julien O., Nancy F. Reichman, and Susan Sprachman. 2003. "Costs and Benefits of Improving Response Rates." *Public Opinion Quarterly* 67:126–38.
- Tourangeau, Roger. 2003. "Cognitive Aspects of Survey Measurement and Mismeasurement." *International Journal of Public Opinion Research* 15:3–7.