

## **PRESIDENTIAL ADDRESS**

### **NAVIGATING THE RAPIDS OF CHANGE: SOME OBSERVATIONS ON SURVEY METHODOLOGY IN THE EARLY TWENTY-FIRST CENTURY**

---

DON A. DILLMAN

These are challenging times for survey methodologists. The telephone, which once seemed the heir apparent to face-to-face interviews, is no longer an obvious choice for conducting many surveys. It is being adversely affected by the change in connectedness of telephone instruments from households to individuals and a cultural shift whereby people are able to control the telephone rather than vice versa, as was the case during its rapid rise to prominence as a mode of data collection in the late 1970s to early 1980s (Dillman 1999).

The internet has not served as an immediate source of salvation. It is limited by restrictions on access, an inability to develop sample frames, and response rates that are mostly lower than those achievable by telephone (Shonlau, Fricker, and Elliott 2002). Other data-collection methodologies, from Interactive Voice Response (IVR) to classroom and home administration on computers, and mixed-mode designs are being proposed as unique solutions for specific survey topics and populations. Random sampling from defined populations is being forgone by some, replaced by modeling of data collected from volunteer panels. The current situation, seen from a methodological perspective, has elements of being both the best and worst of times.

Less than a decade into the era of electronic surveying is not a time for either despair or jubilation, but I think it is an appropriate time to take a look at where survey methodology has been and where it might be going. Specifically, I will talk about changes in our uses of survey modes, new areas of needed research that result from those changes, and some implications for AAPOR. The context for these observations is set by perspectives gained over the course of my career with a variety of survey modes. These experiences

DON A. DILLMAN is the Thomas S. Foley Distinguished Professor of Public Policy in the Departments of Sociology and Rural Sociology and deputy director of the Social and Economic Sciences Research Center at Washington State University in Pullman, WA 99164-4014. This article is a revised version of the presidential address to the American Association for Public Opinion Research, May 18, 2002, St. Petersburg, FL.

Public Opinion Quarterly Volume 66:473-494 © 2002 by the American Association for Public Opinion Research  
All rights reserved. 0033-362X/2002/6603-0013\$10.00

underscore my belief that in the early twenty-first century we are experiencing an extraordinary, but exciting, rapids of change.

### Survey Mode Choices in Personal Perspective

In the spring of 1964, only weeks after becoming a graduate student in sociology, I was handed a cardboard box filled with maps, blank interview forms, and a set of show cards. At the motor pool I was assigned a green state car with "Iowa State University" painted in black letters on the driver's door. Following instructions I drove to Madison County, IA (of covered bridges fame), and began to locate the little black squares marked yellow on the map. My purpose was to interview the farmers, presumed to live in each of the black squares, and to learn how pesticides were being used on Iowa farms.

At the first black square, I carefully parked my car so that the driver's side door would be visible, just as the farmer, who introduced himself as George, came out to greet me. We exchanged pleasantries about the weather, an Iowa custom much older than surveying, and he invited me into his home even before I explained why I was there.

The interview was long, and the house completely quiet. I read 70 pages of carefully constructed questions before coming to the final page and what I had been trained to believe was a sensitive question. "Where do you store your pesticides?" I asked. For the first time in the 2-hour interview, George hesitated. And I waited, as trained to do, in expectant silence.

Finally, the silence was broken by a loud and confident voice from an adjacent room: "George, tell him the truth!" Until that moment I had no idea that George's spouse was in the house, let alone listening to the interview. I felt somewhat triumphant at learning that the pesticides were stored in the basement near the cream separator, but never once questioned whether the answers I got on the remaining 69 pages were accurate. Very long face-to-face interviews were the only legitimate means for collecting data in those days, and we just assumed our data were as good as we could get. No other means of data collection would have been considered for this, or any, of our many surveys.

Yet, only 5 years later I designed and implemented my first telephone survey, and in 1970 established a data-collection facility for conducting telephone surveys, the Washington State University Social Research Center Public Opinion Laboratory. My main memory of that experience is going to meetings to face the skeptical discussants who invariably seemed to begin with "perhaps Professor Dillman has not heard of the Literary Digest Survey debacle" and would continue with how doubtful they were that I could collect accurate information on the telephone. Salvation seemed to be delivered in Airlie House, VA, in 1974, at the first Health Survey Methods Conference,

when Seymour Sudman summarized an unscripted round-robin discussion by declaring that it was apparent that we could now conduct health surveys of 30 minutes or even longer over the telephone. In less than a decade the telephone had moved from being an unrecognized possibility to an accepted method of data collection.

However, these early telephone-survey experiences are somewhat less memorable to me than two more recent ones from my graduate survey-methods practicum that illustrate how quickly certain taken-for-granted aspects of telephone methodology could be misunderstood. In one lecture my description of RDD sampling methods was stopped midsentence by a student's raised hand. "What is RDD?" she asked.

"Oh, random digit dialing," I responded, somewhat apologetically, at my use of an undefined acronym.

"No, I mean what is it?"

After a brief pause to absorb the impact of a graduate student perhaps not having knowledge of the fundamental sampling method that allowed general public surveys by telephone to be done, and detecting a wonderfully teachable moment, I proceeded to give the long answer. It began with how telephone numbers once consisted of two-letter and five-number combinations, such as Blackwood 82248. I also explained how it was often necessary to make calls through operators, who in my community knew a lot more about you than your phone number. Gaining speed, I went on to explain with enthusiasm the development of standardized numbers and the significance of rotary dial telephones. Fortunately, the student interrupted a final time: "Oh, now I understand, I had a princess phone with push-buttons in my room at home. It was the rotary dialing thing I didn't understand!"

This was the same course in which another graduate student dutifully reported back on an examination that the major development in RDD sampling methods was contributed by a person with the last name Mitofsky-Wacksburg. As a result, I permanently changed my teaching notes to reflect the contributions of Warren Mitofsky (1970) and Joe Wacksburg (1978) to the development of a much-needed procedure for drawing national samples of telephone numbers.

Even these reminders of the constancy of change on matters large and small, and of how even the descriptors of our techniques lose meaning, did little to prepare me for the technological revolution in survey methods that we are now experiencing. These classroom experiences occurred at approximately the same time that I was being asked to look at initial attempts to develop a way for survey respondents to answer questionnaires via the internet and that I began participating in discussions of whether we could really get people to punch numbers of answer choices into their telephone keypads. I began these discussions as a skeptic but have since changed views in ways that I did not anticipate.

## Factors Reshaping the Survey Research Environment

Data-collection procedures that seemed curiosities have achieved use and permanency in ways that I had not anticipated, creating a survey landscape that is quite different from the one I described to students as recently as the early 1990s. The five features described below encapsulate many of the factors that are reshaping the environment in which survey methodologists work. These features affect research priorities that need to be addressed. Underlying my description of these features is my belief that our survey methods are more a dependent variable of society than an independent variable. In other words, our methods inevitably reflect the culture around us, much more than they represent anyone's conception of what would be an ideal survey methodology. Fundamentally, we are dependent upon what society at a particular time in history allows and then encourages us to do.

## More Survey Modes from Which to Choose

When telephone-survey methods were being developed in the 1970s, many of us thought that they would replace face-to-face methods more or less completely. That has not turned out to be the case, especially for surveys that require the highest quality household-sampling frames. Now, I hear similar pronouncements about internet surveys replacing all other modes, a future that I do not believe is likely. My own view is that we are going to have not only many different methods in competition with one another but many sub-variations of modes as well.

There are at least five survey modes in extensive use today—face-to-face procedures, telephone interviews, mail surveys, internet surveys, and touch-tone entry (or IVR) surveys. However, these methods only touch the surface of the specific procedures in use. Also being used are self-administered paper surveys, audio- or videotaped questions answered by entering responses into a computer, and questionnaires embedded in e-mails rather than being posted on a Web page. Telephone and face-to-face interviews may be conducted using paper- or computer-based scripts, and paper questionnaires may be optically scanned and imaged or manually entered in computer software. The explosion of acronyms—PAPI, CATI, CAPI, CASAS, audi-CASI, and others—give evidence of the alternatives available for survey data collection and the search for succinct descriptors (Couper and Nicholls 1998). The future of surveying is far more likely to evolve toward the use of different survey modes for different studies than it is to be the disappearance of older modes in favor of only one or two new ones.

### Fitting the Mode to the Survey Situation

The effect of these possibilities for administering questionnaires is to provide an unprecedented opportunity for tailoring (or targeting) the data-collection mode to the population and the survey problem, instead of the reverse procedure for fitting the problem to the survey mode (Dillman 2000).<sup>1</sup> Whereas it may be deemed best to ask school students to answer questions on computers while hearing questions through headphones, for reasons of higher response rates and/or better data quality, other survey populations may be surveyed more effectively in their homes with the use of traditional show-cards. In the last few years, I have helped design employee and customer-satisfaction surveys for establishments for which I was reluctant to recommend anything except touch-tone data entry by telephone and for other situations (e.g., general public surveys such as the U.S. decennial census) in which face-to-face methods were essential for part of the data collection. In other instances (e.g., professional organization member surveys), the internet mode seemed far more desirable and was much less expensive.

We live at a time in which surveyors are more likely to be creating a survey methodology that works best for their population than simply adopting a standard methodology recommended in survey texts. The trajectory of survey methodology is toward diversity rather than unity.

### Greater Use of Mixed-Mode Designs

An important spin-off of the move toward matching modes to the survey situation is the mixing of survey modes for the collection of data from a single population. Several factors contribute to this situation. Lower response rates to surveys, and to telephone surveys in particular, are a driving factor. However, a more fundamental driver is the attempt to reduce costs, using modes for which the marginal costs per interview are less than would be the case for alternatives, for example, substituting the internet or IVR for telephone interviews. Some organizations have attempted to improve the cost-effectiveness of their survey operations by making telephone contacts through an interviewer. However, once a respondent agrees to be interviewed, he or she is routinely switched to an IVR system so that the interviewer can contact the next respondent while the previous interview is being completed.

A third factor that appears to be driving the trend toward mixed-mode surveys is the belief, not well researched, that offering individuals a choice of modes is likely to increase response rates. Earlier evidence (e.g., Dillman,

1. My use of tailoring is somewhat different from that used by Groves and Couper (1998) and Dijkstra and Smit (2002), where they describe using feedback in an interview to decide what the interviewer should say next in an effort to get individuals to agree to be interviewed. Choosing a particular survey mode based on evidence that it will be better received by the potential respondent and/or produce better quality data is also a form of tailoring.

Clark, and West 1994) suggests that offering an alternative mode will cause some people to switch to another mode but will not improve the overall response rate, whereas more recent evidence suggests sequencing mode requests will improve response (Dillman et al. 2001).

As more and more surveys mix modes, the measurement consequences may be substantial. Past research suggests that different survey modes sometimes produce different responses, particularly for opinion questions (de Leeuw 1991; Dillman et al. 1996).

A far more profound implication of using different modes to collect the same data is that researchers tend to ask questions in different ways that they deem best suited for their particular mode. Whereas face-to-face interviewers and mail surveys often used long, fully labeled scales, telephone interviews tend to use fewer scale categories, and many have found it convenient to label only the terminal categories (Dillman 2000, chap. 6). The emphasis on short scales with only the endpoints labeled has increased as designers of IVR surveys have tried to minimize script lengths beyond the brevity of typical telephone interviews. Users of internet surveys are exhibiting a marked preference for the use of check-all-that-apply questions because of the reliance on HTML programming for which radio buttons and check boxes are offered as fundamental response alternatives. But such a question format does not work on the telephone and is usually replaced by a yes/no format. Changes in survey modes used for longitudinal studies are often made without consideration being given to potential mode consequences that confound attempts to measure change from one year to the next.

The same information technology advancements that offer us alternative survey modes also make it possible to easily build a second or third mode of data collection. Question files can be moved easily from one mode format to another. The ease by which alternative-mode questionnaires can be constructed has helped move us from a time in which introduction of a second mode was a major undertaking to it being more of an editing activity.

### **More Surveyors and Survey Organizations**

When face-to-face methods were our dominant survey mode, there were relatively few survey organizations in the United States. Few firms or agencies could muster the necessary resources for developing complex multistage sample designs that made such interviewing efficient.

A major consequence of the development of viable telephone and mail-back methodologies for more survey situations was to expand the number of organizations that could conduct legitimate surveys. While I was director of a university survey organization in the western United States, our first national probability survey by telephone was conducted, and I recall our feeling that the monopoly of a few large survey organizations had been broken. Such

accomplishments are now routine for smaller, regionally based survey organizations. Similar feelings of accomplishment are now being experienced by organizations, and even individuals, who now conduct national and international surveys of groups by electronic mail and the Web.

It is also obvious that new groups doing internet surveys are ones that have not previously conducted surveys by other modes. At my university, many units now conduct their own surveys, whereas a few years ago my research center had a near monopoly on survey design and implementation. Some of these units no longer feel the need for services from a full-service survey organization. I think these groups and their survey activities are here to stay.

A negative impact of this proliferation of surveyors is for many of the new players not to understand, or appreciate, the need for controlling survey errors through sampling or coverage considerations, for example, respondent access codes. Examples abound of new surveyors equating responses from undefined volunteer samples with results that can be generalized beyond the specific respondents.

### **A Substantial Increase in the Number and Importance of Surveys**

Early in my career, surveyors had a ready response for people who felt they were oversurveyed. We explained to skeptics how low the probabilities were for any individual to be asked to respond to a survey of any kind. No longer do I even intimate such a response, as I would have done had George, my first interviewee, been a reluctant respondent. People are being asked to complete many, many surveys, sometimes on a daily basis. We are called on the telephone, sent surveys by mail, asked to fill them out in restaurants and hotels, and exposed to questionnaires by e-mail. That which was once a rare experience has become an almost daily one for some people. Our survey world, and that of our respondents, has changed.

I have listened to colleagues both applaud and deplore the changes outlined above. Some are pleased with the entry of new organizations into the field of surveying, while others worry that the current survey mania tends to cheapen and threaten the entire enterprise of surveying. In my view, the heavy use of surveys that now exists is much more a celebration of what we do than it is a situation to deplore. That is because sample surveys provide an essential function for people, organizations, and society not provided by other means, and we are closer to meeting that need than at any time in the past.

When I came to Washington State University in 1969, with a freshly minted Ph.D., one of my first research activities required getting a state car in order to drive to Seattle. The person behind the desk explained that would not be possible because I had not yet passed the state-required defensive driving course. Thus, on a sunny Saturday in September I found myself sitting through

a day-long course, taught mostly from a manual by a state highway patrol officer. The officer, who correctly sensed by late afternoon that most of his audience was no longer paying attention, asked if we would like to know how to drive down a highway at 110 miles per hour. Recharged by our obvious interest, he began by telling us that we should never try to drive that fast, but anyone who did should ignore the rules of posture taught by most driving instructors. Instead of sitting upright with hands at 10:00 and 2:00, we should crouch in the direction of the door, place our hands at 2:00 and 8:00, and press the sensitive area between our left kneecap and leg bone, hard against the door of the car. He then waited in dramatic silence until one of the audience asked, "Why?"

"Because if the car starts to go out of control you will feel it first through your knee instead of the steering wheel, and that will give you another tenth of a second or so to get the car back under control. You need that small margin of time. When one goes down a street at 20 or 30 miles per hour you can weave back and forth with little concern, but at high speeds things happen too quickly. In order to have a chance of maintaining control you need the earliest possible feedback."

The highway patrol officer's story illustrates the point that when systems are run very hard or fast, it is important to have constant feedback on their performance in order to maintain stability of the larger system. This point has been made more elegantly and in different ways by systems theorists of the same period, such as James Miller (1978) and V. L. Parsegian (1972).

Sample surveys have the remarkable feature of being able to survey only a few hundred or thousand people in order to make statistically reliable statements about tens of thousands or millions of units in the population, provided steps are taken to avoid or take into account survey errors from sampling, coverage, measurement, and nonresponse.

It should not be surprising to us that in our information age in which economic and technological systems are run incredibly fast and hard, relying to a great extent on just-in-time manufacturing and other immediate response mechanisms, that sample surveys have become both prevalent and essential. From the Federal Reserve Board (Woodward 2000) to local physician offices, and from automobile manufacturers to fast food restaurants, data from repeated surveys are used to make essential operating decisions. The use of sample surveys for societal operations feedback to individuals and organizations exceeded greatly their importance when the founders of our organization came together in Central City, CO, in 1946 (Sheatsley and Mitofsky 1992), and this is likely to remain so.

Society and organizations have other rapid feedback mechanisms as well. We experience these daily at grocery stores that record our complete purchase patterns and with other tracking systems. And these systems clearly provide a substitute for some previously done surveys (Baker 1998). However, much of the data needed for running an economy, an organization, or a retail outlet

are not data that run through a recorded transaction. Sample surveys are likely to remain a critical component of feedback systems for monitoring, controlling, and adjusting the systems that make the highly organized life of the twenty-first century possible.

### Some Research Consequences of the New Survey Landscape

What does it mean to have more people and organizations doing surveys, using more different modes of surveying, sometimes separately and sometimes together, at a time when more surveys are being done on more topics, with more consequences, than at any time in our history?

I have contemplated this question a lot during the last few years as I have responded to edgy, if not angry, novice surveyors who found it difficult to accept the idea that the opportunities given to people to respond to questionnaires needed to have been controlled by access codes. I have also thought about this question as people showed me panel data that fluctuated between surveys and were unaccepting of the idea that switching from telephone to the Web might change their trend lines on survey satisfaction questions (Dillman et al. 2001). Time has also been spent trying to convince people that just because they managed to get several thousand responses to a survey the results cannot be applied beyond the specific people who responded. In addition, internet surveys have been shared with me that were designed as if a new piece of software were being created, with the assumption that respondents would be willing to spend the time needed for mastering the software prior to filling out the survey. In e-mails, people have asked what the acceptable response rate is for Web surveys, sometimes with an accompanying argument that the standard should be lower than for other surveys because it is harder to get people to respond to them. I also have memories of visits to some large survey organizations where one unit was creating IVR surveys, another unit creating Web surveys, another unit conducting the mail surveys, and yet another the telephone surveys, and where each seemed to be creating and refining their own methodology. And, the more they refined their individual procedures, the more divergent their question formats became.

On the one hand, these kinds of questions and situations can be rather discouraging: it would appear that survey methodologists have not been very effective in conveying to others why and how surveys work, and convincing people as to best practices. On the other hand, I find reason for celebration when I am asked these questions. There has not been a period in my lifetime when as large a portion of the organizations and people who run them needed to understand and use surveys effectively.

One consequence of the five trends outlined above is the need to pursue somewhat different research questions than those that dominated research agendas during the last few decades. One such issue derives from the greater

importance that self-administration is assuming as a means of responding to questionnaires. Its investigation requires multidisciplinary knowledge that goes beyond the training most of us have received on our way to becoming survey methodologists.

The dominant view of many survey methodologists in the last century was that answering questionnaires was much too complex for respondents to do on their own; interviewers were essential. This view fostered an appropriate focus on improving interviewer models for collecting data, rather than research on how to eliminate the interviewer for most data collection.

This model was mostly adhered to even as the greater use of information technologies fostered a dramatic shift in society toward disintermediation in the provision of all manner of goods and services. People were required to do for themselves activities that were assumed to need assistance. Examples of disintermediation include learning to pump one's own fuel in gasoline stations, to deposit and withdraw money through ATMs rather than tellers, to renew library books and to start and stop newspaper delivery through use of the keypads on touch-tone phones, and for professionals, to do much of their own secretarial work. If other intermediaries are seeing less use, why should we expect the survey interviewer to be an exception?

The decision for members of our profession to start using a particular survey method is not mostly a matter of choice. We adopted the telephone method in the 1970s not because it was inherently better than face-to-face interviews, but because it was cheaper. Self-administered modes such as the internet and IVR are being adopted not because they are inherently better for obtaining answers but rather because the cost factors are advantageous. The computer processing advantages that benefited telephone interviewing in the 1980s and 1990s through the elimination of keypunching are now accruing to these newer methods, as well as to postal surveys, through imaging and processing changes (Dillman 2000, chap. 12). The fact that self-administration has shifted from easy to ignore to being necessary to understand is defining a new research agenda for survey methodology, three examples of which follow.

### **Example 1: Visual Design and Layout**

It is essential that the concepts that underlie the processing, comprehension, and use of visual information be understood. I believe that our approach to understanding potential sources of mode differences has so far developed only part of the equation. Our current understanding of established mode effects (e.g., social desirability, acquiescence, addition and subtraction effects, and category order effects) has relied mostly on analysis of the control features associated with the presence or absence of interviewers, namely, factoring presumed interviewer expectations into people's answers and whether the

interviewer or respondent is in control of delivering the response stimulus (Dillman 2000, chap. 6).

A recent national comparison of the telephone, mail, Web, and IVR modes for collecting long-distance satisfaction information, using multiple 5-point scales with labeled endpoints, revealed that respondents to the telephone and IVR, namely, the aural data-collection modes, were significantly more likely to use the labeled endpoints on the satisfied end of the scale. For example, the percent of “completely satisfied” respondents to telephone and IVR were 39 percent each versus 21 percent for mail and 26 percent by the internet (Dillman et al. 2001). These differences could not be accounted for by non-response error or by traditional models of primacy versus recency effects. It is likely that the visual presentation of response categories gave equal visual prominence to all categories, whereas the aural presentation gave greater prominence to the labeled end categories.

One of the main methodological benefits of the increased use of internet surveys is that more survey methodologists are being required to understand and use concepts that underlie visual processing of information, for example, figure-ground composition; preattentive versus attentive processing of information; the effects of relative size, color, brightness, and location of information on questionnaire pages; reading behavior; and related concepts (Jenkins and Dillman 1997). We have also come to understand that self-administered questionnaires are composed of information being presented in four different languages: the verbal (words), the numerical (numbers), the symbolic (e.g., arrows), and graphical modes of communication (Redline et al. 2002). When these languages are developed in compatible ways they provide stronger guidance on how visual information is processed, comprehended, and used. The appropriate use of these languages to guide respondents through a predefined navigational path offers the potential of bringing the questionnaire stimulus delivered to respondents from self-administered questionnaires more in line with that which is delivered by well-trained interviewers and of reducing mode effects.

Empirical evidence now exists that people’s questionnaire-answering behavior can be influenced by differential applications of these concepts and therefore can be used to reduce respondent errors when answering questionnaires. Specifically, two studies have shown that the failure to follow branching instructions correctly can be reduced significantly through the combined manipulation of graphical (brightness, location, and contrast), symbolic (arrow), and verbal languages used in questionnaire design (Redline and Dillman 2002; Redline et al. 2002). To date, these studies only scratch the surface of providing knowledge that will fulfill the promise of reducing nonresponse error through questionnaire design as well as giving us a better understanding of the causes of measurement errors in surveys.

### Example 2: Understanding the Effects of Culture as Well as Cognition

The 1980s and 1990s are likely to be remembered as the decades when cognitive psychology was brought back into survey methodology (e.g., Krosnick and Alwin 1987; Schuman and Presser 1981; Schwarz and Sudman 1992; Tourangeau, Rips, and Rasinski 2000). These contributions have been tremendously important to gaining an understanding of respondent behavior, and I expect future contributions from cognitive psychology to be of even greater importance. However, the combined emphasis on interviewing as the dominant model for surveying and cognitive science as the equally dominant perspective for understanding what is communicated to and from respondents has resulted in another set of influences (that of culture) receiving considerably less emphasis. By cultural influence I mean shared meanings and understandings of people that are group-based and have motivational importance for human behavior.

When interviewing was our dominant survey model, questionnaire construction tended to focus on how question stimuli were cognitively processed and interpreted. It was left to the interviewer to find ways for motivating people to answer the survey questions. When the shift is made to self-administered questionnaires, then the question of how to get respondents to provide any answer, the role of question wording, and the trappings of how questions are delivered pose new issues that need to be understood.

The infamous butterfly ballot (fig. 1) episode provides a striking example of how cognitive and cultural perspectives worked simultaneously to produce outcomes unintended by many voters. There is compelling statistical evidence that people who intended to vote for Gore for president made errors that disqualified their ballot or produced unintentional votes for Patrick Buchanan (Sekhon et al. 2001). One estimate placed the losses to Gore from unintentional votes for Buchanan at 2,800, from incomplete punches at 2,300, and from double punches at 11,000. I don't believe this outcome can be explained solely by either cognitive or cultural considerations; it requires both. On the cultural side, people who are about to answer questionnaires do not expect the answer choices to be assigned to two different pages, nor do they expect to use the same column for answers with names on one page being located on the right and on the facing page on the left. In some ways that is like driving onto an interstate highway and finding that your side of the highway has two-way traffic. In addition, people have learned that the normal place to begin processing information on written pages is the upper-left portion of the left page, where Bush was listed in the first row and Gore in the next one below. It is also reasonable to expect under this arrangement that the second answer space in the column would be for Gore, rather than for a candidate (Buchanan) whose name was listed on the second page.

On the cognitive side, the task orientation that respondents are likely to

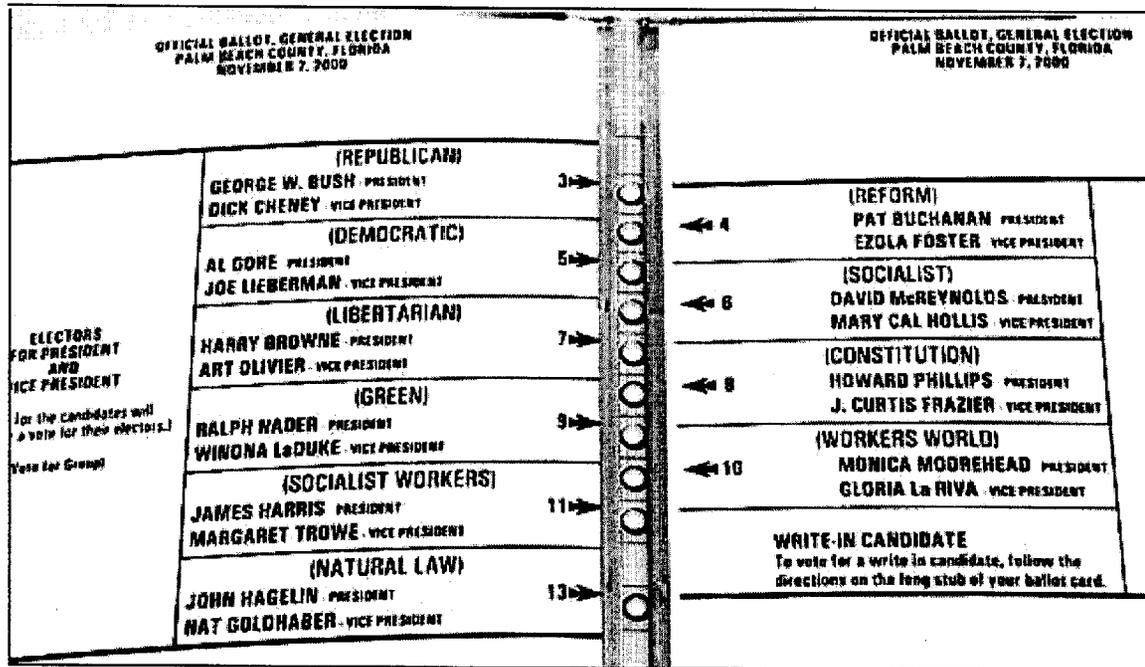
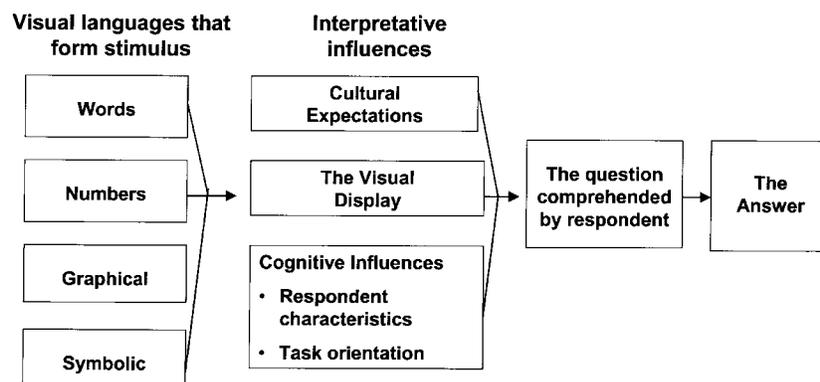


Figure 1. The 2000 Palm Beach County, FL, butterfly ballot shown from approximate voter perspective



**Figure 2.** A proposed model of how language, cognition, and culture influence responses to self-administered questionnaires.

have taken to this ballot is to look for the name of the candidate for whom they have already decided to vote, and there is no reason to expect them to read all of the answer choices prior to marking their preference. The dividing line between the Bush and Gore choice connects the Gore choice precisely to the second circle in the answer column, thus becoming a visual guide that connects widely separated pieces of visual information.

The large number of double punches may also be explained by joint effects of cultural and cognitive considerations. An arrow that points to the third box in the column that carries a cultural meaning of “go here” is located below the Bush/Gore separation line next to the third box in the column. It seems most likely to become visible after one lifts their fingers grasping the stick pin from the first punch. It is also a normal questionnaire-answering behavior, learned over time, and therefore culturally based, to attempt to correct a bad mark once one realizes it has been made. In addition, instructions on the first page, which may or may not have been read, inform respondents to punch “the number of the candidate and/or question, not to exceed the number of votes allowed for each office.” This information may have led some voters to punch a second hole, thinking it was for the vice-presidential candidate in the same box.

Although we cannot know with certainty precisely what happened in the voting booths of Palm Beach County on November 8, 2000, I believe there is a compelling case that both cultural expectations and cognitive processing of visual information contributed to thousands of people not carrying out their voting intentions. The combined forces of visual composition, cognitive, and cultural influences are shown in figure 2. Such a model can also be used to

provide understanding of how people complete other self-administered questionnaires.

### **Example 3: Has the Implied Respondent Contract Changed?**

When I went to Madison County to interview George and other farmers in 1964, there was a lot of culturally based communication and influence at work, beginning with the parking of the car with the Iowa State University emblem, and the legitimacy it implied, in a location where the respondent could see it. And, it even extended to a cooperative spouse helping to assure I got accurate answers on where George stored his pesticides. Well-trained interviewers have continued to provide strong guidance for getting accurate answers, relying on societal norms for interpersonal communication. An implied reciprocity that guides conversations also guides interviews (Schwarz 1996). It is this mutual influence dynamic that I have attempted to capture with a social exchange approach to designing mail surveys (Dillman 1978).

I wonder if this dynamic, which I believe helped honest researchers get honest answers, is changing to one of providing truthful responses when it fits our purpose. Stated another way, has the implied social contract between researcher and respondents undergone a fundamental change that is having a negative impact on data quality? I don't believe there is a single cause for this possible change but rather a set of emergent features of the surveying environment that needs investigation. These features include requests to participate in many more surveys than in the past in many spheres of life. On Web surveys it involves the increasing use of volunteers who sign up for panels who are promised entry into drawings for chances at cash prizes, and even automobiles. Not many are as blunt as the e-mail I received recently: "Your opinions have value. Are you getting paid for them? Start now! Take surveys at home and get paid \$15–\$125/hr. for your opinions. There are over 1,500 surveys everyday for you to participate in. GUARANTEED!"

I also wonder if some of our emerging survey-design practices may be contributing to less commitment for providing honest answers. These include Web-construction practices that require answers to each question before the next page can be displayed, thus, slowing down the process of reaching the "enter the drawing" page. Another concern is excessively long questionnaires that have resulted from the realization by some designers that the marginal data-collection costs for asking additional questions are minimal.

Based upon observations made during cognitive interviews, I also wonder if a new style of question reading and comprehension is emerging for Web surveys, with questions less likely than those on mail surveys to be read in their entirety, resulting in truncated comprehension. Repeatedly, I have observed cognitive interview respondents read only part of each question before going to response categories. The back button is used only if they cannot

figure out the meaning of the next screen. Their behavior suggests a trial-and-error answer process, consistent with how they might figure out the way in which various software applications work. The sense of not wanting to mark a wrong answer because of the difficulty of the correction process that I have often observed in interviews of paper-questionnaire respondents seems not to be part of the response process. Occasionally, I wonder if there is a game mentality guiding responses to Web surveys for figuring out how to answer quickly and maximize chances for winning prizes. Does this produce as a side effect an increase in the proportion of meaningless or inaccurate responses?

We are too early in the innovation process to have good answers to these questions, but not too early to put their possible consequences onto our research agenda. Is the sense of obligation for providing honest answers declining, while responding to surveys is becoming more of a game than a case of serious reflection and response? Certainly this is not an issue limited only to Web surveys. Punching buttons to answer IVR surveys and the high volume of surveys on customers may contribute to respondents contemplating first what message they want to send to the survey sponsor, and only second whether they want to give one that is an honest reflection of their views. Web sites that provide “hotlinks” for voting for a hotel chain for the Freddie awards and others that ask site visitors to show their support for a candidate by responding to a candidate-preference survey provide a different motivational context for responding to surveys than prevailed in the past. Another example is the car salesman who explains the importance of a new car buyer giving all 5’s when the inevitable customer-satisfaction survey arrives.

### **AAPOR as an Important Meeting Place**

How well and how quickly researchers respond to new issues such as those I have outlined above depends in part upon AAPOR. When people at my university ask why I come to AAPOR, my brief answer is that is where the knowledge is of how to do surveys well. Whether intentionally or unintentionally, AAPOR has become the annual meeting place for people concerned with survey design. I know of no other organization whose members do as much scientific work on data-collection procedures, measurement, and non-response. Our conference, our journal, and AAPORNET bring together influential interchanges about most of the elements that produce quality in surveys. At one time AAPOR was viewed by some of my friends as a nice secondary organization to belong to. That is a view I do not hold. AAPOR is the best place to come if one wishes to be on the cutting edge of survey methodology and to learn how to do surveys well. Three features of our organization are especially important for maintaining AAPOR’s quality of being the most important meeting place for survey methodologists. Each needs

continued attention if the organization is to maintain its current importance for advancing research on survey methodology.

### **Maintaining the Balance between Commercial and Noncommercial**

One of AAPOR's most important features is our explicit recognition of commercial and noncommercial sectors, with the latter being further subdivided into academic, government, and nonprofit interests. We have a governance structure that assures that neither dominates the other, through requiring that certain officers alternate between sectors from one year to the next. This balance needs to be carefully protected.

Once I was invited to give a talk in small, very rural Washington community. A state senator from the area took me aside to offer some advice. He said that he would give me a joke that I could use to gain audience acceptance, but after that I was on my own for communicating to this small group of business people. I should tell them about the three biggest lies, he said. The first lie, which he explained with the help of outstretched hands, was to say I caught a fish that long. The second lie I should mention was that the check was in the mail. However, the third, and most important lie, was to say, "I am from the government and I'm here to solve your problem." Naturally, I proceeded to tell this to a government friend who reminded me of the fourth biggest lie, "I am from the University and I understand your problem." Yet another friend suggested a fifth lie: "I am from the private sector and although I'm not necessarily here to solve your problem, that will change if the contract you are thinking of is large enough."

I have been fortunate in my career to spend a significant amount of time working with government and private organizations, as well as universities. Each has impressed me enormously with the quality of some of the surveys they do, and each does some work I find rather disappointing. I also believe that we innovate in different ways at different rates. Each of us has much to learn from the others.

If exposed only to people in our own employment sector, I believe each of us will respond less well to the need for research on the design of mixed-mode surveys and the specific issues I have outlined above. Government has a reputation for sometimes developing methods that, once in place, become enormously resistant to needed change (Bryant and Dunn 1995). Commercial firms have a reputation for developing their own private methodologies with algorithms they won't reveal, making it difficult for outsiders to understand whether they do or do not work. University professors also have a reputation of attempting to test difficult-to-evaluate theories in rigidly controlled ways for the sake of testing the theory, and without much regard to practical results. Left to our own devices, each of us is likely to propagate and promulgate

some mythologies about what makes surveys work in addition to good science. At this time of rapid change it is essential that we have a forum for learning from and influencing one another, and AAPOR provides such a forum.

### **Responding to the Challenges of Maintaining a Larger Tent**

In recent years our annual conference has expanded horizontally, that is, with more sessions in competition with one another—as many as eight sessions were scheduled simultaneously in this year's conference. This is a substantial departure from earlier times when conference attendees would attend the same sessions in the morning and discuss the common experiences at lunch. Although the emotional side of me is nostalgic for that sort of community experience, I am also aware that the complexity of modes and needs of their users requires learning experiences that will lack interest to other members. In addition, many of our members have little interest in survey methodology, but their areas of expertise are crucial to the mission of our organization.

AAPOR needs to attract the new surveyors of the world to our meetings. But I worry about whether AAPOR will welcome members whose interests are highly specialized and diverse, and whose concepts of surveying differ somewhat from those we have promulgated in the past. My preference is that we provide a large and accommodating tent that encourages professionals with diverse views on surveys and surveying to participate meaningfully as members. I believe that consciously encouraging professionals with diverse views on roles and methods of surveying will help keep surveyors, both large and small, from fragmenting into specialized camps that see survey quality in quite narrow ways.

Fragmentation on mode choices and their uneven application to a variety of survey situations provides one of the major threats to quality surveying in the United States. Some survey methodologists are interested in developing measurement scales that can be copyrighted for their company's use in measuring employee satisfaction. Others are mostly interested in modeling data from volunteer populations. Certain surveyors are mostly interested in diary methods, while others are concerned with panel techniques. The interests of still others start and stop with what can be accomplished using a particular mode, whether the internet, IVR or audio-CASI. The larger the different kinds of survey arenas in which we work in this era of tailored design, the more likely we will be able to learn from and influence one another.

### **Emphasizing Survey Standards through Education**

Traditionally, the strength of AAPOR as a meeting place was in the sessions and the discussion that ensued in our face-to-face meetings. Commensurate with the development of our information economy, I wonder if our greatest

potential strength might be in development of our potential as an educator for how to do surveys well. To some extent this is already under development. In recent years AAPOR has expanded its short courses for conference attendees, and this feature has attracted to membership individuals who need to learn about the broader uses of survey methodology. Yet, I believe we have only touched the surface of the topics that might be covered and what might be done for novice surveyors, regional chapters, and for nonmembers such as newspaper reporters covering elections. Such efforts might also be expanded to our Web pages. The organization that provides the best opportunity to learn about advances in survey methodology should logically be the one where people go to learn about the fundamentals of quality surveying. The entry of so many new people and organizations into the field of surveying and the need for basic education on how and why surveys get conducted motivate this concern.

I also wonder if in the long run emphasizing the development of educational offerings will be a more effective promoter of standards than our current procedures for enforcing standards. Throughout our history, AAPOR has maintained a heavy focus on survey standards and made it a practice to enforce those standards with members as well as nonmembers. It is a process that, in my opinion, does not work well. The length of time it takes to investigate cases and the amount of time it takes to follow our carefully stipulated procedures make it an unsatisfying process both for those who wish to enforce standards as well as those who are being investigated. This appears to be the case especially for the frequent allegations of push-polling that now occur during election campaigns. I am also struck by the enormous amount of time that standards cases take and how that time is competitive with other AAPOR Council activities.

With greater diversity in survey modes and the increased numbers of organizations conducting surveys, I expect that the enforcement of standards is going to become even more difficult to accomplish. It is time for AAPOR to take a fundamental look at the area of enforcing survey standards, to decide what's possible, and how it might best be accomplished in our more diverse survey environment.

Thus, a long-run education strategy for teaching all comers about survey methodology may serve AAPOR better than a strategy of identifying and enforcing specific violations of standards. It may also have the effect of making our organization a more hospitable environment for newcomers to the survey field and ultimately strengthen our capacity for asking the right research questions.

## Conclusion

It is hardly surprising that while conducting my first interviews with George and the other farmers of Madison County, IA, I had no idea of the nearly constant change that would affect the conducting of surveys across the next four decades or of my personal involvement in trying to understand those changes. Nor did I have an inkling of how one change would build upon another (e.g., the structure of long-distance calling, creation of desktop computers, touch-tone calling, laptop computers, fiber optic telecommunications, and the internet) and cumulate into new forms of surveying that I was unable to imagine. And I was least able to imagine how the demand for surveying and the institutional structures for meeting that demand would also change and continue to change throughout my career.

I cannot imagine a more exciting time to be a survey methodologist than the one now before us. The availability of many survey modes from which to choose, the tendency to tailor survey procedures to the situation, plus the far greater use of mixed-mode designs create both the need and the opportunity for exciting advancements in our craft. The entry of new professionals into the field of surveying and the great expansion in numbers of surveys also create enormous opportunities. And, as modes and structures for conducting surveys change, the research needs to change also.

I have described the need for new research perspectives and concepts that flow from the long-term trend toward greater use of self-administered survey procedures for collecting survey data. Visual design and layout, the influence of culture on response, and possible changes in respondent perceptions of the social contract for answering survey questions are only three of many issues that could have been discussed here. However, I believe that each reflects a significant paradigm shift that needs to be pursued.

Rapid change is never easy for professional organizations to handle. AAPOR faces the prospects of greater diversity in the survey backgrounds and interests of professionals who come to our meetings. We face the continuing challenge of maintaining our role as a meeting place for public opinion researchers and survey professionals from different sectors of society as the diversity both within and between sectors expands. These changes in turn raise important questions about education versus enforcement as means for improving survey quality.

Navigating these rapids of change will not be easy. The challenges to AAPOR and to the profession of survey methodology are enormous. However, I look forward with enthusiasm to the experience and believe it is a ride worth taking.

## References

- Baker, Reg. 1998. "The CASIC Future." In *Computer Assisted Survey Information Collection*, ed. M. P. Couper, R. P. Baker, J. Bethlehem, C. Z. F. Clark, J. Martin, W. L. Nicholls II, and J. M. O'Reilly, pp. 583–605. New York: Wiley.

- Bryant, Barbara Everitt, and William Dunn. 1995. *Moving Power and Money: The Politics of Census Taking*. Ithaca, NY: New Strategist Publications.
- Couper, Mick, and W. L. Nicholls II. 1998. "The History and Development of Computer-Assisted Survey Information Collection Methods." In *Computer Assisted Survey Information Collection*, ed. M. P. Couper, R. P. Baker, J. Bethlehem, C. Z. F. Clark, J. Martin, W. L. Nicholls II, and J. M. O'Reilly, pp. 1–21. New York: Wiley.
- De Leeuw, Edith. 1991. *Data Quality in Mail, Telephone and Face-to-Face Surveys*. Amsterdam: TT Publications.
- Dijkstra, Wil, and Johannes H. Smit. 2002. "Persuading Reluctant Recipients in Telephone Surveys." In *Survey Nonresponse*, ed. Robert M. Groves, Don A. Dillman, John L. Eltinge, John L. Little, and J. A. Roderick. New York: Wiley.
- Dillman, Don A. 1978. *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley.
- . 1999. "Mail and Other Self-Administered Surveys in the Twenty-First Century: The Beginning of a New Era." *Gallup Research Journal* 2(1):121–40.
- . 2000. *Mail and Internet Surveys: The Tailored Design Method*. 2d ed. New York: Wiley.
- Dillman, D. A., J. R. Clark, and K. K. West. 1994. "Influence of an Invitation to Answer by Telephone on Response to Census Questionnaires." *Public Opinion Quarterly* 58(4):557–68.
- Dillman, Don A., Glen Phelps, Robert Tortora, Karen Swift, Julie Kohrell, and Jodi Berck. 2001. "How Choice of Survey Mode Influences Answers to Customer Satisfaction Surveys." Paper presented to American Association for Public Opinion Research Annual Conference, Montreal, May 18.
- Dillman, Don A., Roberta L. Sangster, John Tarnai, and Todd Rockwood. 1996. "Understanding Differences in People's Answers to Telephone and Mail Surveys." Chap. 4 in *Current Issues in Survey Research, New Directions for Program Evaluation Series*, ed. Marc T. Braverman and Jana Kay Slater, pp. 45–62. San Francisco: Jossey-Bass.
- Groves, Robert M., and M. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: Wiley.
- Jenkins (Redline), Cleo, and Don A. Dillman. 1997. "Towards a Theory of Self-Administered Questionnaire Design." In *Survey Measurement and Process Quality*, ed. L. Lyberg, P. Biemer, M. Collins, L. Decker, E. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin. New York: Wiley-Interscience.
- Krosnick, J. A., and D. Alwin. 1987. "An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement." *Public Opinion Quarterly* 51:201–19.
- Miller, James Grier. 1978. *Living Systems*. New York: McGraw-Hill.
- Mitofsky, Warren. 1970. "Sampling of Telephone Households." Unpublished CBS memorandum, New York.
- Parsegian, V. L. 1972. *This Cybernetic World of Men, Machines and Earth Systems*. Garden City, NY: Doubleday.
- Redline, Cleo, and Don A. Dillman. 2002. "The Influence of Alternative Visual Designs on Respondents' Performance with Branching Instructions in Self-Administered Questionnaires." In *Survey Nonresponse*, ed. R. Groves, D. Dillman, J. Eltinge, and R. Little. New York: Wiley.
- Redline, Cleo, Don A. Dillman, Aref Dajani, and Mary Ann Scaggs. 2002. "The Effects of Altering the Design of Branching Instructions on Navigational Performance in Census 2000." Research Report Series, Survey Methodology #2002-02. U.S. Census Bureau, Statistical Research Division, Washington DC.
- Schuman, H., and S. Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments in Question Form, Wording, and Context*. New York: Academic Press.
- Schwarz, N. 1996. *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*. Hahwah, NJ: Erlbaum.
- Schwarz, N., and S. Sudman. 1992. *Context Effects in Social and Psychological Research*. New York: Springer.
- Sekhon, J., J. Wand, K. Shotts, W. Mebane, Jr., M. Herron, and H. Brady. 2001. "Voting Irregularities Caused by the Butterfly Ballot in Palm Beach County." Paper presented at Joint Statistical Meetings, Atlanta, August 6.
- Sheatsley, Paul B., and Warren J. Mitofsky. 1992. *A Meeting Place: The History of the American Association for Public Opinion Research*. Ann Arbor, MI: American Association for Public Opinion Research.

- Shonlau, Matthias, Ronald D. Fricker, Jr., and Marc N. Elliott. 2002. *Conducting Research Surveys via E-mail and the Web*. Santa Monica, CA: Rand.
- Wacksburg, J. 1978. "Sampling Methods for Random Digit Dialing." *Journal of the American Statistical Association* 73(361):40–46.
- Woodward, Bob. 2000. *Maestro: Greenspan's Fed and the American Boom*. New York: Simon & Schuster.
- Tourangeau, Roger, L. J. Rips, and K. Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press.