Mail and Other Self-Administered Surveys in the 21st Century: The Beginning of a New Era*

by

Don A. Dillman

"If we could know where we are and whither we are tending we could better figure how to get there."

--Abraham Lincoln--

Early in the 20th century, probability sampling for surveys was a radical idea. The notion that about a thousand people selected from households throughout the United States could yield consistent and accurate/reliable estimates of characteristics of the entire population seemed to defy reason. Such sampling is now accepted as an essential cornerstone of the survey method, and the precision of resultant estimates is taken for granted.

Similarly, the idea that interviews on important survey topics could be conducted by telephone, rather than in person, was a radical idea in the 1960s. Telephone numbers had just become standardized with seven digits and area codes, and long distance calls still were conducted mostly through operators. Further, the idea that telephone would supplant personal interviews as the dominant method of surveying, conducted by interviewers who read answers from a computer screen and entered answers into software designed to tally their answers, was difficult to imagine. The few computers that did exist were kept in secured air-conditioned rooms, and only highly-trained operators could touch them; the personal computer had not been developed. Yet, as we

---

*A discussion paper written by Don A. Dillman, Professor of Sociology and Rural Sociology, and Deputy Director for Research and Development at the Social and Economic Sciences Research Center, Washington State University, Pullman, WA 99164-4014, 509/335-1511; e-mail: dillman@wsu.edu. Comments and suggestions from Al Paez and Dretha Phillips and many others on an earlier draft of this paper are acknowledged with thanks. This is a work in progress paper. Additional suggestions and comments will be appreciated.

near the end of the 20th century, computer-assisted interviewing by telephone has become the most accepted data collection method for national surveys.

In this paper I predict another radical shift. It is that self-administered surveys, which leave interviewers out of the data collection process entirely, will become the dominant method of surveying early in the 21st century. Just as the 1950s and 1960s saw the personal interview reach its zenith of importance, and the 1980s and 1990s have witnessed a similar emphasis on surveying by telephone, there is reason to expect that the self-administered survey will achieve prominence in the first two decades of the 21st century. A constellation of forces, which makes the change probable rather than only a distant possibility, will shift many surveys from those requiring an interviewer to methods that do not. These methods include regular mail, courier, fax, electronic mail, the Web, and touch-tone data entry; all of which provide new challenges to survey sponsors and methodologists. For this reason, I also identify research issues needing attention from the survey methodology community to facilitate conducting high quality self-administered surveys in the early 21st century.

Telephone Interviewing Is Being Challenged

There can be little doubt that the nation's dominant survey method in the late 1990s is telephone interviewing. It has been used increasingly as a replacement for in-person inter-viewing. In situations where the lack of telephones in households prevents its use as a stand-alone method, it is being integrated in mixed mode designs; an example is the Current Population Survey (used to estimate the nation's unemployment rate), in which most of the reinterviews are conducted by telephone. General public telephone surveys with households being selected by randomly generated telephone numbers have become synonymous with correct surveying. Most house-holds (95%) now have telephones, and random sampling of telephone numbers is routine.

Thus, it came as a shock in the 1996 general elections when the Reuter's telephone poll conducted by John Zogby came closer to predicting the outcome of the 1996 presidential election than any other national pollster. This surprise stemmed in large part from his methods:

He samples from national lists of telephone numbers and doesn't

> worry that he doesn't capture unlisted households in his sample. He interviews from 9A.M to 9 P.M. and weights down what he describes as the resulting mild over sampling of women and retirees. And, he adjusts his sample to match his estimation of party identification (34.5% Democrat; 34% Republican; the rest independent). Other pollsters recoil in horror at such practices (Morin, 1996).

Zogby predicted an 8-point victory for Clinton, the exact margin, while other major polls predicted differences ranging from 11 points (USA Today/CNN) to 18 points (CBS/New York Times). Closer examination of the data show that Zogby's raw data were further from the actual results (with undecideds excluded) and that his success was based almost entirely on his allocation model for undecided voters, including weighting (Moore, 1996). The fact that all major polls must be weighted for underrepresentation of certain segments of the likely voter population, and that these adjustments and decisions on the allocation of undecided voters moved some predictions further away from the election outcome, underscores the complexities now associated with telephone interviewing. Weighting for individual characteristics to compensate for underresponse by certain segments of the population (e.g., younger people) in telephone surveys has become a standard means of improving their accuracy.

For election surveys, telephone interviews are used by all major pollsters. Other methods are considered too slow or subject to fatal errors. It's against this background that Peggy Visser, Jon Krosnick, Jesse Marquette, and Michael Curtin (1996) analyzed results from three longtime election polls in Ohio. Beginning early in the 1980s, each had attempted to predict outcomes of the Ohio elections. One poll was a mail survey by the Columbus Dispatch newspaper. The other polls were telephone surveys by the universities of Cincinnati and Akron. The Columbus Dispatch surveyed only registered voters, sending them a facsimile of the state ballot, whereas the other polls were conducted by standard random digit dialing (RDD) telephone methods.

The result of this comparison came as a surprise. The mail Dispatch poll more precisely predicted the outcomes of the 32 races, exhibiting an average error of 1.6 percentage points compared with 5.4 percent for the University of Akron poll and 4.9 percent for the University of Cincinnati poll.

Individual election differences also were consistently in favor of the Columbus Dispatch poll. All but two of 32 predictions by the Dispatch were closer to election outcomes than those of the Akron poll, with one being a tie. Similarly, 14 of 18 predictions for the same elections were less accurate for the Cincinnati poll than for the Dispatch.

How could this happen? It is conventional wisdom that mail surveys should not be used for election forecasting. There are several reasons for expecting the Columbus Dispatch mail surveys to do worse than the telephone surveys. The mail polls had lower response rates (about 25 percent compared with about 60 percent) and opinions were elicited as much as ten days before the election. In addition, the mail returns were cut off at least three days before the election, so last-minute voter decisions could not be captured. Also, the mail polls did not ask questions in the manner commonly done for telephone. They listed only the office, candidate names and party, rather than wording the questions in what has become the conventional poll format, i.e., "If the election were held today, which of these candidates would you vote for . . . ."

These observations reflect to some extent the overwhelming acceptance of telephone surveys as the way by which election polls should be done. Thus, the attributes of the telephone method are the standard against which alternative methods are to be judged. When viewed from a formal survey error perspective, the more accurate results of the mail surveys seem less surprising. Four major sources of errors in surveys--sampling, coverage, measurement, and nonresponse--are deemed critical to achieving survey accuracy (Groves, 1989).

From a coverage standpoint, the mail questionnaires were sent only to registered voters; the telephone pollsters had to ask whether a person was a registered voter, a question for which socially desirable "yes" answers are commonplace (Visser et al., 1996). Thus, the telephone surveys entailed considerable risk of obtaining opinions of nonregistered voters. From a measurement standpoint the telephone surveys posed the voting preference questions in a hypothetical manner, i.e., "If the election were held today, which of the following candidates . . . ." Party identification and candidate order were not always done in the same way. In contrast, the mail poll question was a replica of the election day ballot. Therefore, if order of candidate listing and mention of party identification had an effect in the election, it could be expected to have the same effect in the mail survey.

From a sampling error standpoint, the mail surveys had larger completed samples (about 1,600 compared to 800), and therefore had less sampling error. It's only from a response rate (25 percent by mail, 60 percent by telephone) perspective that the mail survey came up short, which is only an indirect indicator of nonresponse error. Overall, the mail survey scores better on three of the four aspects of survey error, so perhaps it should not be too surprising that the results more precisely predict the outcome of the elections.

Neither of these examples, the success of the Columbus Dispatch mail poll or that of the effects of Zogby's unconventional adjustment methods, suggests the imminent demise of conventional telephone survey methods. However, they do serve as illustrations of an unease that I believe many survey methodologists are feeling with telephone interviewing methods in the late 1990s. They also serve as a reminder, if any is needed, that the 20th century has been a time of rapid change in survey methods and that more changes may lie ahead as we begin the 21st century. In the paragraphs that follow, I consider the nature of these changes and what they may portend for the future.

Why Survey Methodologies Have Changed During the 20th Century

Changes in the nature of survey methods and the rise of telephone interviewing in the last third of the 20th century can hardly be traced to a single innovation or other cause. Modifications of institutionalized methods, of which the conduct of surveys certainly is an example, occur slowly and usually result from the coalescence of factors in the survey environment that encourage the use of one method over another. These forces of change can be grouped conveniently under four categories: (1) changes in societal organization and culture; (2) available technology; (3) sources of cost and efficiency; and (4) a consideration of contributors to survey error. Each of these is discussed below with regard its their influence on surveys since the 1940s.

Societal organization and culture

In the post-War era of the 1940s, probability sampling and the survey method were "discovered" by U.S. society as having enormous power to provide badly needed knowledge. Sample surveying embodied the idea that one could estimate the characteristics of households,

individuals, or other large units of society within a few percentage points by surveying only a thousand or two of those units. The development of multi-stage sampling methods with stratification and clustering, in conjunction with principles for in-person interviewing, occurred at the same time that a more centralized mass society was developing, as described by sociologists such as Daniel Bell (1965).

As national corporations replaced local companies and more local, as well as state, governmental functions were assumed by the federal government, an appropriate emphasis developed in understanding the characteristics of national populations. For members of the general public, households represented the only available sampling unit for locating and selecting random samples. There was no methodological alternative to in-person interviewing. This method was consonant with U.S. culture inasmuch as virtually all transactions of importance in society took place via in-person contacts.

Dramatic increases in levels of living gradually made the telephone a standard fixture in most American homes, with the coverage rate reaching 90 percent as early as 1970, an increase of more than 100 percent from the 1930s. This coverage rate, though not complete, became acceptable for many surveys. The prospect of telephones being available in nearly every home occurred just as it was becoming increasingly evident that national surveys were not a financially viable alternative for surveys that could produce valid small geographic area estimates. The fact that local-area surveys needed sample sizes nearly as large as national surveys meant that state and even city surveys were extremely costly. The need for an alternative to the face-to-face interviews increased as the limitations of using national estimates for local discussions in a heterogenous nation became increasingly evident.

When the telephone commenced being used for interviewing in the 1970s, this voice method of communi-cation was viewed very differently than it is today. In the 1960s and before, it was an instrument for brief, short-distance conversations. Long-distance calls were expensive and required going through successive operators. People often had to shout to be heard. For a young person growing up in the 1940s and 1950s, the simple words of a parent, “it’s long distance,” also meant “Please be quiet!” The cultural norms associated with telephone calls meant that any telephone call was considered important, and all callers were to be treated politely. Behavior on

the telephone was simply an extension of in-person interaction norms and mores that were small-town oriented. Also, the high cost of long distance calls meant that such calls were short and to the point. Typically, people could not ignore a ringing telephone; the call was probably important.

Yet, the telephone was not considered a legitimate means for serious data collection. When rigorous attempts to develop telephone interviewing methods began in the early 1970s, it came as something of a surprise that surveys could be conducted by telephone. I recall participating in the 1972 "Health Survey Methods Conference" at Airlie House in Virginia, one of the first times telephone interviewing was discussed at a major national conference. The significant conclusion offered against considerable skepticism was that it was possible to get people on the telephone to complete meaningful interviews for perhaps as much as an hour. Serious efforts followed to develop precise telephone methodologies with three independently-conceived books being published in rapid succession (Blankenship, 1977; Dillman, 1978; Groves and Kahn, 1979). Also during the 1970s, research showed that mail questionnaires could obtain data of reasonable quality (Dillman, 1978; Heberlein and Baumgartner, 1978). However, the mail method continued to face formidable coverage problems; there were no acceptable national address lists, a problem that still persists. By the mid-1970s, methodologists talked frequently about the telephone being the heir-apparent to the face-to-face interview.

An important contributor to the rise of telephone interviewing and replacement of the in-person interview was the speed of project completion. The advent of the information age, in which organizations quickly obtained information to make equally quick decisions, also helped propel the telephone interview into its current dominant position. As early as 1974, during the Watergate hearings, prominent national overnight surveys were being done to obtain people's reactions. Now, it is simply routine in election campaigns to conduct nightly surveys to track a candidate's fortunes and make decisions on campaign strategies. As society has shifted towards tracking fluctuations as well as enduring concerns, the telephone showed it could perform very well.

The rapid development of telephone interviewing, in conjunction with technological developments discussed below, contributed to the decentralization of telephone interviewing. For the first time, a small organization located anywhere in the United States could do national as well as local

surveys, and do them economically (Dillman, 1978). By the early 1990s, the telephone effectively had replaced in-person interviews for virtually all important interview surveys, or was being used in tandem with in-person interviews to reduce the costs of geographical dispersion. It had become the nation's most important survey method.

Technology

Until the 1960s, the United States had no telephone area codes and interviewing required assistance from long-distance operators. Telephone numbers often included letters, and so lacked numerical standardization. Clearly hearing the other person was difficult on long-distance calls. Changes in the nation's telephone system, which have occurred at an escalating rate, provided necessary technological improvements that allowed both sampling and interviewing by telephone.

Computer-assisted interviewing, whereby data were entered into a computer during the interview thereby eliminating keypunching, sped up the process of collecting and reporting data. However, when this process started, it was necessary to use mainframe computers with less memory than most of today's personal computers. Getting those computers to interpret a response and show the next question on the screen was a slow, even tortuous, process for both interviewers and respondents. The development of stand-alone micro-computers early in the 1980s gave powerful boon to telephone interviewing and its decentralization. The addition of automatic dialing, call scheduling, and data reporting software, in conjunction with continuing increases in the capabilities of software to customize interviews, contributed enormously to the ease of conducting telephone interviews.

Computer-assisted interviewing was extended to personal interviewing via laptop computers, allowing field interviewers to send results into their main office as soon as their work is finished. Mail surveying also has benefitted from technological changes, as desktop publishing has made it possible to print questionnaires, send correspondence, keep track of returns, and perform other activities that once required large amounts of hand labor and were subject to considerable error. Address files now can be created and used with great efficiency.

Cost and efficiency

Increased changes in telephone costs for in-person interviewing contributed to its decline. In the early 1970s, it came under pressure from changes in the traditional labor force (e.g., home-makers willing to work part-time), as well as higher interviewer and transportation costs and the need to make multiple callbacks to find people at home. Obtaining the final personal interviews needed to reach acceptable response rates became enormously expensive.

The ability of a telephone interviewer to sit in one location and make call after call, with geographical dispersion having virtually no effect on costs, contributed to telephone interviewing becoming far cheaper than the face-to-face alternative. The development of Computer-Assisted Telephone Interviewing (CATI) systems eliminated keypunching as a separate cost activity and reduced costs for handling and storing paper questionnaires. The development of automated dialing and other routines increased the number of calls interviewers could complete and added other efficiencies, making telephoning a more desirable alternative to in-person interviewing. The telephone method was the first of the survey methods to benefit from computerization.

Error structure considerations

It's tempting to think of improvements in survey error structure as the driving force behind changes in survey methods and procedures. As others have pointed out, survey decisions usually entail an accommodation between survey costs and errors (e.g., see Groves, 1989). In the shift from in-person interviewing to telephone interviewing, somewhat greater coverage error (nontelephone households) was accepted as a trade-off for lower costs and, as a result, promoted the ability to do surveys that otherwise would have been deemed not possible.

Although telephone interviewing began with skepticism that high quality data could be obtained, comparative research has shown that data collected by the two methods are reasonably comparable, and there have been no major changes in data quality as a result of the shift away from in-person interviewing (Groves and Kahn, 1979). From a sampling standpoint, it has become possible to eliminate some of the more complex sampling methods (e.g., multi-stage cluster samples) that were less cost efficient.

Forces of Change in the Late 1990s

Telephone interviewing is under considerable pressure as we approach the 21st century. Methods that seemed to work well in the 1980s are becoming less effective. Informal discussions with telephone survey organizations suggest that achievable response rates are decreasing, and more and more callbacks are required to obtain acceptable response rates. Perhaps most troublesome are the greater number of calls reported by one national agency to be as high as 30 percent that go unanswered by a real person after repeated attempts. It's been shown that while response rates for telephone interviewing methods have been decreasing in recent years, those for mail self-administered surveys have held steady (Hox and de Leeuw, 1994). In addition, the potential for other types of self-administered surveys is increasing. If the use of self-administered survey methods is to increase, it seems likely that the same four forces of change--societal organization and culture; technology; costs and efficiency; and error considerations--will again play important roles.

Societal organization and culture

We live in an organizational era of total quality management. Increasingly, organizations are attempting to build in feedback mechanisms for all of their units. Customer satisfaction surveys have become standard operating practice for many organizations. Although the benefits of sampling remain evident, providing reliable information on organizational components such as individual franchises in a restaurant chain requires overall sample sizes that are much larger than those necessary for overall feedback only to the parent organization. For the most part, the handling of enormous data sets and providing analyses by subunits is easy to accomplish. More and more the costs of data collection and initial processing become a controlling factor on what an organization can accomplish through its surveys.

The rapidly evolving information era has placed new demands on the population. Whereas the telephone culture placed great emphasis on verbal exchange, the new culture places a greater emphasis on writing skills. Significant limitations are placed on an individual's ability to succeed in life by not becoming computer literate and lacking Internet access. Effective use of computers requires good keyboard skills and being able to compose and deliver prose quickly. Writing skills are important for being able to respond to self-administered questionnaires and this skill can be

expected to improve further as the use of computer-administered surveys goes up, at least in the short run.

A major societal trend also exists towards "self-administration." Many activities that once required individuals to interact with another person are now being shifted to a self-administration mode. Using ATMs to obtain banking information and money instead of going to a bank teller is an example. Others include using touchtone input to stop and restart newspaper delivery, renew library books, and register for college classes. Home diagnostic kits for all sorts of medical information from blood sugar levels to HIV diagnosis, ordering airline tickets, purchasing stocks, buying gasoline in service stations, getting information from the Web, and serving as one's secretary are other examples. Effective functioning in U.S. society now requires skills in following instructions and interacting with machines that were not generally anticipated a decade ago. Self-administration is required for more and more aspects of life.

The role of the telephone is changing also. It has evolved from being an interruption that demanded priority over other activities and controlled behavior, to potential intrusion that can be delayed and otherwise controlled effectively. Answering machines, active call-screening, caller identification, call forwarding, and other technologies are being used increasingly to select when and with whom we communicate. And, increasingly, it appears more and more interaction that normally would be accomplished by voice telephone is shifting to written electronic mail.

A related trend is the increase of multi-telephone and multi-line households. And, judging from the growth in cordless, cellular, and other devices (including personal pagers), the telephone is becoming a personal device rather than a household instrument, making it more difficult to reach a random sample of the entire population in this way.

Technology

Technological developments are providing new options--unthinkable a few years ago--for survey self-administration. The typical recent choices of either conducting a mail-out/mail-back survey versus conducting a telephone voice interview now seems strangely antiquated. E-mail, fax, the Web, touchtone data entry—all of which are self-administered—have become feasible survey

methods, especially business surveys (e.g., Clayton and Werking, In Press).

Many households now have multiple ways of screening calls: first by obtaining unlisted numbers (nearly 30 percent of all U.S. households), second by caller identification that displays the number from which a call is being placed, and third by active screening of all callers when they start to leave a voice message. Ironically, just as a person's ability to avoid unwanted calls has increased, the ability to keep one's names off widely accessible lists is decreasing. It is now possible for nearly anyone with Internet access to locate the address and telephone number for any person with a listed telephone number from a national directory and even to locate the household on a local map. Credit card information, magazine subscription lists, mail order purchase informa-tion, utility lists, and large numbers of other address information sources continually are compiled, unduplicated, and used to create national address lists. These lists, and the information that comes with them, can be used to identify segments of the population with unusual characteristics ranging from disabilities to ownership of specific types of recreation equipment.

Increased use of weighting for characteristics of nonrespondents to surveys, difficulty in getting past technological gatekeepers by telephone to obtain data that then have to be weighted, and multiple problems associated with screening households by phone to find eligible individuals for special interest surveys, all present important challenges for telephone surveys. When these problems are juxtaposed with the ability to compile larger, higher quality mail lists, one can envision greater interest in finding ways to design and implement reliable list-based surveys.

Costs and efficiency

During the 1980s and 1990s, the benefits of developments in computer technology accrued mostly to the telephone method. Now, however, substantial benefits are being realized by self-administered survey methods. Responding to e-mail now requires a level of effort no greater than responding to a telephone call, i.e., one can answer the message immediately and reply (by clicking on the send icon) instantly with no more effort than was required to hang up a telephone. Each recontact with a telephone survey provided an immediate opportunity for responding, whereas mail protocols typically used some reminders without sending a replacement questionnaire. Now, any recontacts by e-mail can include a replacement questionnaire as easily as

not. The CATI software that automatically skipped questions determined unapplicable by previous answers is emulated in Web surveys. In addition, Web surveys can be self-editing without additional processing steps, in much the same way as CATI reports have been processed.

For regular mail-out/mail-back self-administered surveys, effective optical scanning has been an elusive goal. In the past, optical scanning has been limited to reading marks. In the 1990s, technology for imaging the entire questionnaire onto computer screens and reading characters as well as marks has become widely accessible. It appears that we will soon be able to switch virtually all questionnaires from keypunching to optical scanning, although systems to accomplish that remain quite expensive (Dillman and Miller, In Press). The electronic methods of reaching people for self-administered surveys are especially intriguing because they occur in an environment that increasingly expects people to be able to communicate in writing.

In essence, the elimination of mail-out/mail-back costs that electronic mail and Web-based surveys make possible, the increased speed of response that these technologies also offer, the near automatic retrieval and analysis of Web surveys, and optical imaging and scanning developments that will make large volume processing of questionnaires more efficient—all suggest that cost and efficiency barriers to the use of self-administered surveys are being mitigated.

Survey error structure

The Columbus Dispatch poll serves as a reminder that survey error is not an attribute directly associated with survey mode. Rather it is a four-dimensional concept defined by coverage, sampling, measurement, and nonresponse errors, each of which may differ for a particular mode as well as by specific populations and survey situations.

In the past, self-administered methods and mail-out/mail-back surveys in particular were viewed as inferior, mostly because of inadequacies with regard to coverage and nonresponse. Lists were not available for general public samples, and response rates tended to be lower than desired. For many specialized populations, neither of these is a problem. And, even for general public surveys, certain aspects of these concerns may be changing. As telephone answering machines have spread and cultural readiness to decline telephone surveys has grown, it appears that significantly

higher response rates are not likely to be obtained by telephone. In fact, lower telephone response rates are encouraging the use of mixed-mode surveys, whereby mail or fax are tried when an interview cannot be completed by telephone.

We are at an interesting juncture in the history of telephone service. If the era of wireless communication with inexpensive telephones come to fruition through development of low-orbital satellites (as some have predicted), the connection of telephones to households may be broken and the telephone truly becomes a personal device. As a result of not being able to sample from telephone numbers, we may be forced to develop survey methods that are list-dependent. The ability to develop, manipulate and maintain multi-source lists is developing rapidly just as our RDD paradigm is being threatened. Already, mailing lists are being used to augment RDD lists so that households can be contacted by either or both methods. If enhanced lists should become an acceptable method of sampling, then I expect that this change alone will foster increased interest in the use of self-administered methods. Developments in this area need to be watched closely.

A substantial amount of recent work on survey measurement has not produced evidence that interview methods are inherently superior for producing data. In fact, the reverse may be true. There is fairly strong evidence that interview methods, whether in-person or by telephone, yield more socially desirable answers and respondent acquiescence than do self-administered methods, including the direct use of computers to provide answers (e.g., de Leeuw, 1992). With the exception of item nonresponse and completeness of open-ended responses, both of which may be improved for self-administered methods, a topic addressed in the final section of this paper, there is little evidence that interviews produce better measurement. In any event, better measurement is not a compelling reason for avoiding self-administered methods (Dillman et al., 1996).

Where we are and whither we are tending

If the four forces of change--societal organization and culture, technology, costs and efficiency, and survey error considerations--are decisive ones in shaping our survey future, then I think a strong case can be made that greater use of self-administered survey methods is inevitable. Societal demand for small group/area surveys is moving us in that direction, and the shift towards a self-administration culture supports it. New technologies also are expanding the range of

methods for implementing self-administered surveys. In addition, the cost efficiencies that benefitted telephone interviewing by eliminating keypunching are poised to benefit self-administered surveys in the early 2000s through the application of CATI-like methods to Web surveys and optical scanning and imaging of paper questionnaires. Finally, survey error considerations do not represent major obstacles to the trend promulgated by this constellation of forces.

Research for a Self-Administered Future

The forces that are creating a greater reliance on self-administered survey methods and less on interview methods are not the result of dramatic research findings about surveys. They are mostly the result of much more powerful forces of technological social, economic, and cultural change. However, appropriate research can contribute to an easier transition that protects survey quality.

In general, I believe that research typically proceeds along different paths. Some research is conceptualized within existing paradigms and is aimed more at refining current methods than developing new ones. Another path tends to note problems with existing paradigms and focuses on potential alternatives.

My concern that our telephone paradigm of interviewers and random digits may be breaking down makes me wonder if research aimed at improving these methods will help much with conditions to be faced in the early 21st century. And, if self-administered methods are to emerge as effective alternatives, what kind of research is needed.

In this regard I was struck more by some of the promising early investigations than by the revolutionary character of needed research. Here briefly are four research issues, each of which is important to determining the role of self-administered surveys early in the 21st century.

<u>How to design respondent-friendly self-administered surveys</u>

I cannot recall a research comparison in which the self-administered survey achieved a lower item nonresponse than did the interview method. Traditionally, we have thought of this mainly as a problem of not having an interviewer in control of questionnaire administration rather than of a

graphical questionnaire layout and design problem. At the same time I'm struck by the poor construction of many self-administered questionnaires for most surveys.

Over the years many survey methodologists, myself included, have attempted to articulate principles for questionnaire design, but we have done so mostly on the basis of subjective assessment (e.g., lining up answer boxes in a vertical way on the page seemed to look better than placing them horizontally wherever an answer category ends) (Dillman, 1978). The scientific principles underlying these suppositions for the most part have not been tested systematically and articulated as principles for design. I think this situation is changing.

Efforts are being made to draw from the visual design sciences using concepts such as ground-figure, law of Pragnanz, law of proximity, law of simplicity, etc., to distinguish bad design that allows, or even encourages, item nonresponse from good design that does not (Jenkins and Dillman, 1995). Using these concepts, questionnaires can be evaluated on the basis of two aspects of construction, clarity of the navigational path and information organization. Three critical characteristics of the graphical and word symbols on the page—location, brightness (including color) and shape—influence whether words are understood by respondents (Jenkins and Dillman, 1995, 1997). The theory that emerges leads to a myriad of hypotheses, virtually none of them empirically tested. For example:

> Use of a consistent ground-figure format (e.g., black letters on white paper) increases the accurate comprehension of survey questions on self-administered questionnaires, whereas changes in the ground-figure format (e.g., inserting a red word in a sequence of black on white paper words) decreases accurate comprehension.

> Compliance with skip directions is increased by placing them closer to the answer boxes and using larger fonts for them.

I am optimistic that we are on the verge of identifying principles of questionnaire construction that will produce far better data than ever thought possible for all types of self-administered

questionnaires. Research on these principles is fundamental to obtaining high quality data from self-administered surveys. It's a topic that also has been virtually ignored, as many self-administered forms resemble interview instruments more than they do something to facilitate a self-guided effort by respondents.

How to conduct high quality e-mail and web surveys

Numerous attempts have been made to conduct e-mail surveys in recent years. Many of these studies have treated the e-mail technology as if it were an entirely different survey technology. For example, questionnaires have been sent only once and, when a poor response was obtained, it was concluded that e-mail surveys were not yet possible. Just as multiple contacts are essential in personal, telephone, and mail surveys for achieving acceptable response rates, it's reasonable to assume that such is the case for e-mail surveys as well.

With this as a guiding principle, a recent attempt was made to survey Washington State University faculty, over 90 percent of whom had reported in an earlier survey that they used e-mail. A four-contact strategy was devised that included a prenotice and three contacts with personalized addressing of messages. Exactly the same response rates (58 percent) were obtained for the e-mail survey and a control panel contacted by mail. Item nonresponse rates were similar. However, as one might expect, respondents to the e-mail survey gave more detailed responses to open-ended questions. Another advantage of these e-mail surveys is that the bulk of the responses were returned more quickly. In fact, the survey was e-mailed on a Friday afternoon, and by Monday nearly 25 percent of the subjects had responded (Schaeffer and Dillman, 1997).

Surveying by e-mail has to overcome, in the short run, coverage problems (paper questionnaires were sent to faculty not having e-mail addresses) and adaptations to the technology (a postcard reminder by mail without a questionnaire makes much sense, but in an e-mail survey sending a reminder without another questionnaire does not). In the longer run, coverage is unlikely to be a problem for some kinds of surveys, e.g., those of many types of professionals and employees of most organizations. This type of self-administered survey also must deal with different e-mail formats, e.g., line widths. As computer and transmission technologies continue their rapid growth in capacity, this seems to be a transitional problem.

How to conduct high quality World Wide Web surveys

Web surveys offer far more long-term surveying potential than do stand-alone e-mail surveys. E-mail systems do not allow standard formatting, whereas web surveys do. Web survey instruments can be built using the same principles of design as mentioned above for constructing respondent-friendly control routes through questionnaires—just as the best available CATI systems do at present. In fact, advanced programming routines (such as Java) are opening new design dimensions, encompassing instrument usability and functionality features such as data self-editing, that are impossible to simulate in other survey mechanisms. However, the existence of different computers, screen configurations, and web browsers makes it difficult to use these advanced web survey methods.

Another current problem with web surveys is reaching people. Web use is not as prevalent as e-mail use. Further, if people are asked to complete a web survey, they face a daunting task of keying in a typically long web address. This problem has an easy solution. If people are running both a web access program and e-mail, and a particular type of address is placed in e-mail, then they can do something as simple as double-clicking to move quickly and effortlessly to the survey. This task appears beyond the literacy level of many current e-mail users, but that seems likely to change as computer literacy becomes even more prevalent in U.S. society.

International surveys have been difficult to do by physical mail. E-mail and web surveys seem ideal technologies for reaching rapidly across boundaries and great distances, with the result that their use can be expected to increase as computer capabilities and literacy continues to expand worldwide.

Substitution of optical scanning and imaging for keying of self-administered mail survey answers

The development of optical scanning systems has been especially troublesome for survey methodologists (Dillman and Miller, In Press). Traditionally, making marks on surveys has required filling in circles or ovals with special pencils or pens. If handwritten answers were

required, it was necessary to run these answers through separate coding and keypunching processes. Also troubling was the difficulty of correcting mismarked answers with white-out or other time-consuming efforts.

The development of electronic imaging systems that incorporate optical character recognition technologies has improved conversion of handwritten numbers and letters. It seems unlikely, however, that this technology will be able to read cursive writing without frequent errors; the variations are too great.

A new approach to optical scanning and imaging may be overcoming these problems (Dillman and Miller, In Press). It involves adhering to principles for respondent-friendly design to achieve more readable answers. In addition, it involves imaging the entire questionnaire into a computer and reconceptualizing the data entry process so that it occurs independent of the physical questionnaire. People are asked to mark boxes instead of fill in ovals, a process that has less error. Printing of white boxes on color-tone backgrounds help people confine their answers to the space that is scanned for answers, so that marks for one box do not interfere with answers for another. Thus, answers are easier for the optical device to read.

Next, an image of the questionnaire is projected onto a computer terminal. The programming that supports this system makes a "confidence" assessment of the response interpretations; those that cannot be read confidently are directed toward correction files. Then, these potential errors are processed by individual items (rather than by questionnaire) for an operator to make a judgment. If the operator cannot make a judgment by reviewing only the answer categories on the screen, then the entire questionnaire can be projected to reach a context judgment. No attempt is made to read narrative answers, the length of which is difficult to predict (e.g., "Why

did you choose to purchase this computer?"). These answers are brought to the computer screen and an operator keys a file of standardized responses.

Observation of an advanced and fairly expensive imaging system of this nature has convinced me that—by foregoing the effort to scan all answers accurately, expecting to subject difficult ones to an operator's interpretation, and developing a correction process that is very fast (because of not having to go item-by-item through one questionnaire at a time, as well as not having to retrieve the actual questionnaire)—keypunching in its traditional form may soon be a thing of the past. Adding to this belief is research that suggests questionnaires designed for electronic imaging do not inherently get lower response rates (Dillman and Miller, In Press). A barrier to the use of an advanced imaging system such as that described above is cost, a factor that seems likely to diminish as computer capabilities continue to expand with larger memories and faster operations.

In sum, promising advances are being made in three areas important to obtaining quality responses to self-administered surveys: respondent-friendly design, conducting electronic surveys by e-mail and the web that make it possible to parallel the CATI achievements of the 1980s and 1990s, and optical scanning. In all of these areas further development is essential, thus defining an important research and development agenda for survey methodologists.

Conclusion

It has been some 135 years since Abraham's Lincoln observation , "If we could know where we are and whither we are tending we could better figure how to get there." Yet, it seems strikingly modern. The creation of institutional structures and mechanisms, and certainly surveying is an important one at this time in history, inevitably reflect forces of societal change. The last decades of the 20th century have unleashed technological, social, cultural and economic forces that seem likely to change significantly, and once again how data on our current situations are collected and used. These are forces that cannot be ignored by survey methodologists.

I believe that the nature of these forces are such that the use of self-administered surveys--by traditional mail, electronic mail, fax, the Web, touchtone data entry, and eventually voice recognition. To some extent this greater use is of interest because of changes in the societal

structure that has pushed the random digit telephone interviewing paradigm to its current prominence. To a greater extent though, its the result of beginning to apply computer technologies for writing, designing, administering and processing self-administered questionnaires in a way that is now bringing the same efficiencies to these types of surveys that came to telephone interview surveys in the 1980's and 90's. These efficiencies suggest that interviewing costs will eventually become unaffordable for most surveys just as they have in the delivery of banking services, dispensing of gasoline, receptionist transfers, and a host of other societal activities.

Certainly, these changes will not occur over night. The successful Columbus Dispatch election survey by mail, or the victory of Zogby's unconventional methods over those of other pollsters are more suggestive of chinks in the armor of current telephone methods, than gaping holes that portend immediate disaster. But, most significant change does not occur dramatically in an instant. Anomalies frequently happen, and give rise to repair efforts, just as more dents are being discovered. Unfortunately many of the deficiencies occurring in telephone survey methods are less the elements subject to the control of survey methodologists, than they are the cultural change that is enveloping the telephone device, e.g. detachment from households, and shifting from the role of controlling human behavior, to being controlled through multiple layers of protection from list access to voice mail.

The attention to maintenance and repair has led to the growing capability of alternative possibilities being ignored by many, and particularly those with a heavy investment in the current methods. These self-administered possibilities, though far from being useable as an immediate replacement in most situations, are getting better. Bringing attention to such possibilities, i.e. the replacement of interview methods by a variety of self-administered methods, has been my purpose in this paper. It's also been my purpose to contemplate needed lines of research that will produce an understanding of the potential and limitations of these methods and if successful facilitate their widespread adoption.

I expect that transition to greater use of self-administered methods will occur first through mixed-mode surveys as we use them to compensate for the deficiencies of interview methods, and evaluate whether complete replacement is possible. Indeed that is already the case. This mixing

methods is being propelled in part by the fact that computer technologies make it easier to combine methods into one survey and blend the results, than was possible in the past. Predicting the future, particularly on matters that depend upon human creativity, technological innovation, and cultural change is dicey at best. The practical implications of individual technologies inevitably depend upon their intersections with other technologies, and make predicting specific changes difficult. In 1985, I attempted to predict the implications of information technologies for rural America (Dillman, 1985). At the time the paper was described to me by a friend as long on speculation and short on compelling evidence, and maybe a little unrealistic. A recent rereading of that paper suggested to me that not only did the writing seem almost archaic in its terminology, but that I had greatly underestimated (rather than overestimated) the impacts of such technologies. I was unable to imagine how powerful and accessible computers would become, how much the speed of transmission would improve, and the almost immediate necessity of fax machines in every business. More importantly, I didn't anticipate the rapid development of e-mail becoming essential to virtually all businesses. The form taken by the web revolution of the 1990's was something I had not been able to imagine at all. I will be surprised if I am not also surprised by some of the new technological forms relevant to surveying that are created in the next 13 years and the magnitude of change that accompanies them.

References

Bell, D. 1965. **The End of Ideology: On the Exhaustion of Political Ideas in the Fifties** (2nd Edition). New York Free Press: New York.

Blankenship, A. B. 1977. **Professional Telephone Surveys.** McGraw Hill.

Clayton, Richard and George Werking. ??? "Business Surveys of the Future: The World Wide Web as a Data Collection Methodology." Chapter 27 in Mike Couper et al. (eds.), **Computer-Assisted Survey Information Collection**. Wiley & Sons: New York.

de Leeuw, E. 1992. **Data Quality in Mail, Telephone, and Face-to-Face Surveys.** TT Publications: Amsterdam.

Dillman, Don A. 1978. **Mail and Telephone Surveys: The Total Design Method.** Wiley & Sons: New York.

Dillman, Don A. and Kent Miller. In Press. "Optical Scanning Technologies, Response Rates and Data Qualityi in Mail Surveys by Small Survey Organizations." In Mick Couper, et al. (eds.), **Computer-Assisted Survey Information Collection.** Wiley and Sons: New York.

Dillman, Don A., Roberta L. Sangster, John Tarnai, and Todd Rockwood. 1996. "Understanding Differences in People's Answers to Telephone and Mail Surveys." In Braverman, Marc T. and Jana Kay Slaters (eds.), **Current Issues in Survey Research, New Directions for Program Evaluation Series**, Chapter 4:45-62. Jossey-Bass: San Francisco.

Groves, Robert. M. And Robert L. Kahn. 1979. **Surveys by Telephone: A National Comparison with Personal Interviews.** New York Academic Press.

Groves, Robert M. 1989. **Survey Errors and Survey Costs**. New York: Wiley and Sons.

Heberlein, Thomas A. And Robert Baumgartner. 1978. "Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature." **American Sociological Review** 43:447-62.

Hox, Joop and Edith D. de Leeuw. 1994. "A Comparison of Nonresponse in Mail, Telephone and Face-to-face Surveys." **Quality and Quantity** 28:329-344.

Jenkins (also Redline), Cleo and Don A. Dillman. 1995. "The Language of Self-Administered Questionnaires as Seen Through the Eyes of Respondents." **Volume 3, Statistical Policy Working Paper 23: New Directions in Statistical Methodology,** pp.470-516. U.S. Office of Management and Budget.

Jenkins (also Redline), Cleo and Don A. Dillman. 1997. "Towards a Theory of Self-Administered Questionnaire Design." In Lars Lyberg et al., **Survey Measurement and Process Quality**. Wiley & Sons: New York.

Morin, Rich. 1996. "A Bad Night at the Exits." **Washington Post Weekly.** November.

Moore, David. 1996. "Perils of Polling '96: Myth and Fallacy." **The Polling Report** 12(23): December.

Schaefer, David and Don A. Dillman. 1997. "Development of a Standard E-mail Survey Methodology: Results of an Experiment." Unpublished paper presented at annual meeting of the American Association for Public Opinion Research.

Visser, Peggy, Jon Krosnick, Jesse Marquette, and Michael Curtin. 1996. "Mail Surveys for Election Forecasting?" **Public Opinion Quarterly** 60(2):181-227.