Review: Web Surveys: A Review of Issues and Approaches

Author(s): Mick P. Couper

Source: *The Public Opinion Quarterly*, Winter, 2000, Vol. 64, No. 4 (Winter, 2000), pp. 464-494

Published by: Oxford University Press on behalf of the American Association for Public Opinion Research

Stable URL: https://www.jstor.org/stable/3078739

WEB SURVEYS
A REVIEW OF ISSUES AND APPROACHES

MICK P. COUPER

As we enter the twenty-first century, the Internet is having a profound effect on the survey research industry, as it is on almost every area of human enterprise. The rapid development of surveys on the World Wide Web (WWW) is leading some to argue that soon Internet (and, in particular, Web) surveys will replace traditional methods of survey data collection. Others are urging caution or even voicing skepticism about the future role Web surveys will play. Clearly, we stand at the threshold of a new era for survey research, but how this will play out is not yet clear. Whatever one's views about the likely future for survey research, the current impact of the Web on survey data collection is worthy of serious research attention.

Given the rapidly growing interest in Web surveys,[1] it is important to distinguish among different types of Web surveys. The rubric "Web survey" encompasses a wide variety of methods, with different purposes, populations, target audiences, etc. I present an overview of some of the key types of Web surveys currently being implemented and do so using the language of survey errors. In order to judge the quality of a particular survey (be it Web or any other type), one needs to do so within the context of its stated aims and the claims it makes. By offering a typology of Web survey designs, the intent of this article is to facilitate the task of evaluating and improving Web surveys.

Web surveys represent a double-edged sword for the survey industry. On the one hand, the power of Web surveys is that they make survey data collection (as opposed to survey participation) available to the masses. Not only can researchers get access to undreamed of numbers of respondents at dramatically lower costs than traditional methods, but members of the general

MICK P. COUPER is a senior associate research scientist in the Survey Research Center at the University of Michigan and a research associate professor in the Joint Program in Survey Methodology. The author would like to thank Fred Conrad, Don Dillman, Karol Krotki, George Terhanian, and Eleanor Singer for their helpful suggestions on earlier drafts of this article, and Peter Miller and Vincent Price for their valuable editorial contributions.

1. As evidence of this interest, see the comprehensive bibliography on Web survey methodology maintained by the RIS (Research on Internet in Slovenia) Project at the University of Ljubljana, at http://www.websm.org.

Public Opinion Quarterly Volume 64:464–494. © 2000 by the American Association for Public Opinion Research

population too can put survey questions on dedicated sites offering free services and collect data from potentially thousands of people. The ability to conduct large-scale data collection is no longer restricted to organizations at the center of power in society, such as governments and large corporations. The relatively low cost of conducting Web surveys essentially puts the tool in the hands of almost every person with access to the Internet, potentially fully democratizing the survey-taking process. Furthermore, Web surveys make feasible the delivery of multimedia survey content to respondents in a standardized way using self-administered methods. This opens up a whole new realm of survey possibilities that were heretofore impossible or extremely difficult to implement using more traditional methods.

On the other hand, the potential risk of Web surveys is that with the proliferation of such surveys, it will become increasingly difficult to distinguish the good from the bad. The value of surveys that could be done on the Web is limited—as with other approaches—by the willingness of people to do them. Thus, the whole enterprise may be brought down by its own weight if we get to a point where persons are so bombarded with survey (or other) requests that they either tune out completely or base their participation decisions on the content, topic, entertainment value, or other features of the survey. Well-designed, high-quality Web surveys may very well be overwhelmed by the mass of other data-gathering activities on the Web. We may already be seeing this oversurveying effect in telephone surveys, where the proliferation of telemarketing is threatening the use of this mode of data collection for representative surveys of the general population. In summary, then, while Web surveys in general may become increasingly easy to do (both cheaper and quicker), good Web surveys (as measured by accepted indicators of survey quality) may become increasingly hard to carry out.

The intent of the preceding discussion is not to portend gloom for the survey industry but, rather, to note that we simply do not yet know how things will turn out for this exciting new method and for the industry as a whole. While the methods of survey research may be changing in ways that we cannot predict, the fundamental criteria by which we evaluate surveys remain largely unchanged.

Survey Quality

Any discussion of Web surveys must be couched in the context of the type, form, and function of the survey. More so than any other mode of survey data collection, the Internet has led to a large number of different data-collection uses, varying widely on several dimensions of survey quality. Any critique of a particular Web survey approach must be done in the context of its intended purpose and the claims it makes. Glorifying or condemning an entire approach to survey data collection should not be done on the basis of a single implementation, nor should all Web surveys be treated as equal.

O'Muircheartaigh (1997, p. 1) offers as a definition of error in surveys: "work purporting to do what it does not do." He goes on to write, "Broadly speaking, every survey operation has an objective, an outcome, and a description of that outcome. Errors (quality failures) will be found in the mismatches among these elements." Survey quality is not an absolute but should be evaluated relative to other features of the design (such as accuracy, cost, timeliness, etc.). We also need to evaluate the quality of a particular approach in light of alternative designs aimed at similar goals (e.g., quota samples, mall intercepts, low response-rate random-digit dial [RDD] surveys, magazine insert surveys, customer satisfaction cards, etc.).

Several years ago, I predicted that the rapid spread of electronic data-collection methods such as the Internet would produce a bifurcation in the survey industry between high-quality surveys based on probability samples and using traditional data collection methods, on the one hand, and surveys focused more on low cost and rapid turnaround than on representativeness and accuracy on the other.[2] In hindsight, I was wrong, and I underestimated the impact of the Web on the survey industry. It has become much more of a fragmentation than a bifurcation (in terms of Web surveys at least), with vendors trying to find or create a niche for their particular approach or product. No longer is it just "quick and dirty" in one corner and "expensive but high quality" in the other; rather, there is a wide array of approaches representing varying levels of quality and cost.

The problem is that almost all of these are done under the general rubric of "surveys," making it more difficult for the lay person (and indeed many in the survey industry or informed users) to differentiate between the different offerings in terms of quality. It is becoming even more important that both the producers and consumers of survey data become educated in the elements of survey quality, so that the different survey designs are not treated as equal.

It is thus useful to evaluate the types of Web surveys currently available in terms of the traditional measures of quality and sources of errors in surveys. Several error typologies have been offered in recent years (see, e.g., Groves 1989). It is generally agreed that the major sources of error in surveys include sampling, coverage, nonresponse, and measurement error, all of which must be evaluated relative to costs. I first discuss the implications of these various sources of error for Web surveys and then present a typology of different Web survey types in light of the different error sources.

2. In a panel discussion on "CASIC: Brave New World or Death Knell for Survey Research: A Debate," at the 1997 annual conference of the American Association for Public Opinion Research (AAPOR).

COVERAGE AND SAMPLING ERROR

Coverage error is presently the biggest threat to inference from Web surveys, at least to groups beyond those defined by access to or use of the Web. The problems of sampling in many Web surveys also present a formidable barrier to probability-based sample surveys on the Web.

Coverage error is a function of the mismatch between the target population and the frame population. The target population can be viewed as the set of persons one wishes to study or the population to which one wants to make inference (e.g., the adult population of the United States). The frame population is defined by Groves (1989, p. 82) to be "the set of persons for whom some enumeration can be made prior to the selection of the sample." An alternative definition, proposed by Wright and Tsao (1983, p. 26), refers to "the materials or devices which delimit, identify, and allow access to the elements of the target population." Examples of frames include all residential telephone numbers (for a telephone survey) or all personal e-mail addresses (for Web surveys).

Sampling error arises from the fact that not all members of the frame population are measured. If the selection process were repeated, a slightly different set of sample persons would be obtained. Note that we are explicitly dealing with probability samples here, that is, samples where every member of the frame population has a known, nonzero chance of selection into the survey. Thus while coverage error refers to people missing from the frame (in this case, those without Internet or Web access), sampling error arises during the process of selecting a sample from the frame population, necessitating a means of identifying people on the frame.

One can immediately see two problems in Web surveys. One is that not everyone in the target population (unless this is narrowly defined as currently active Web users, for example) is in the frame population. The second problem is constructing the frame for Web surveys. Even if every person in the United States has access to the Web, the difficulties of constructing a frame to select a probability sample of such persons are daunting. Given that the process of selecting elements from a frame of Web users (or however the frame is defined) is highly dependent on the type of Web survey being conducted, we focus our attention here on the problems of coverage.

At present, coverage error represents the biggest threat to the representativeness of sample surveys conducted via the Internet. As Groves (1989, p. 85) notes, coverage error is a function of both the proportion of the target population that is not covered by the frame and the difference on the survey statistic between those covered and those not covered.

Let us first address the issue of the coverage rate, or the proportion of the target population that potentially can be reached via the Web. While the proponents of Web surveys point to the astonishing growth rate of the Internet as a reason for their optimism regarding Web surveys of the general public, there remain major concerns regarding the future penetration of this technology.

There are many different estimates of the penetration of the Web. While it is admittedly difficult to measure such a rapidly moving target, estimates of Internet access in American households have varied widely, partly because of definitional problems (definitions of access and use, household- vs. person-level estimates, home vs. office, etc.) and partly because of the varying quality of the surveys used to estimate Internet penetration. Some studies have even used surveys conducted via the Internet to measure the size of the Internet population.

A fall 1998 report by Mediamark Research (cited in www.thestandard.com, November 16, 1998) estimated that 53 million Americans (or 27 percent of U.S. adults) were online, having used the Internet in the past 30 days, whether from home, work, or elsewhere. The estimates were based on 20,000 in-person interviews conducted between September 1997 and August 1998.

Hoffman, Novak, and Schlosser (2000) report data from the CommerceNet/Nielsen Internet Demographic Study (IDS) conducted in May/June 1998 and based on an RDD survey of 4,042 persons 16+ in the United States. This study yielded an estimate of 69.6 million recent (6 months preceding the survey) Internet users, representing 34.4 percent of the U.S. adult population.

In a November 1999 press release, the Strategis Group (www.strategisgroup.com) reported that 101.7 million adult (18+) Americans, or 49.7 percent, use the Internet, up from 65 million in mid-1998, and 84 million at the end of 1998. This number had increased to 106.1 million, or 52.4 percent, of adult Americans, by December 1999. These estimates are based on the question "Do you use the Internet or online services at home or at work?" posed in an RDD telephone survey.

National Public Radio, Kaiser, and the Kennedy School of Government conducted a telephone survey of 1,506 adults 18 and older in November/December 1999. Estimates from this survey were that 53 percent of all adults currently use the Internet, whether at home or work or elsewhere (64 percent report ever having used the Internet). The survey also found that 60 percent of U.S. adults have a computer at home.

Data from CyberDialogue's American Internet User Survey conducted in the third quarter of 1999 (www.cyberdialogue.com) found that 35 percent of the U.S. adult population (or 69.3 million persons) were active online.

Using an RDD telephone survey, IntelliQuest (www.intelliquest.com, press release, April 19, 1999) estimated that 83 million adult U.S. users, or 40 percent of the U.S. population 16 or older, are accessing the Internet. This is an increase of 10 million over the 73 million users (27.8 percent of the U.S. adult population) estimated in October 1998 (see www.nua.ie/surveys/how_many_online). The Nua Internet Surveys site also shows July 1999 estimates from Nielsen//NetRatings of 106.3 million U.S. Internet users, or 39.4 percent of the population. The Nielsen//NetRatings Internet universe is defined as all members (2 years of age or older) of U.S. households that currently have access to the Internet.

Data from the February 1999 Harris Poll (www.harrisinteractive.com), an RDD survey of 2,015 adults surveyed by telephone, showed that 63 percent of all adult Americans use a computer, whether at home, work, or elsewhere. Forty-two percent of the adult population reportedly use a computer at home. The poll report (Taylor 1999) did not include estimates of Internet users. Terhanian (2000), however, reported figures from the Harris Poll (month not known) of 42 percent of U.S. adults who used the Internet in 1999.

While all these figures are now dated, the point is to show the divergent estimates of Web usage from a variety of sources during the same approximate time period. In 1998–99 we see estimates ranging from 27 percent to 53 percent of U.S. adults who use the Internet. Data from the Census Bureau suggest that the lower numbers may be more accurate.

Probably the most accurate measures of Internet access and use come from periodic supplements to the Current Population Survey (CPS), conducted by the U.S. Bureau of the Census. These are based on a face-to-face survey of almost 50,000 U.S. households each month, the latest of which was conducted in December 1998.[3] The CPS has a response rate around 93 percent and is designed to cover all households in the United States.

Data from this study suggest that 42.3 percent of U.S. households (or about 44 million households) have one or more computers in the home. A total of 25.6 percent of all households have access to the Internet from home (either using a personal computer or WebTV). This represents almost 27 million households out of a total of about 104 million households in the United States. Turning to person-level estimates for adults (18 or older), the CPS data suggest the distribution of Internet use, as of December 1998, shown in table 1.

Thus, just over 65 million U.S. adults, representing about one-third of the adult population, use the Internet whether from home or elsewhere. While this represents enormous growth in Internet use in recent years, access or use is still far from universal. Whether Internet or Web surveys can be representative of the U.S. population at some future time depends on the continued growth in Web use. The current debate is essentially about how far such growth will go and how fast it will occur.

Much of the optimism regarding the potential of Web surveys is based on the predicted trajectory of future penetration, extrapolating from the tremendous growth in WWW use seen in recent years. But this remains speculation. Inevitably, the growth curve must slow down and eventually plateau, but at what point will it do so (see Mack 1999)? Will Web access ever be universal?

One assumption is that because the pace of adoption of the Web is pro-

3. Previous Internet and computer use supplements were conducted in 1994 and 1997. The supplement is again being fielded in August 2000, and plans are underway to introduce an annual supplement starting in September 2001.

Table 1. Internet Use among U.S. Adults, December 1998

	%	No. of Users (in Millions)
Use Internet:		
Both at home and outside home	7.2	14.1
At home only	15.9	30.9
Outside home only	10.3	20.1
Total	33.4	65.1
Do not use Internet	66.6	129.5
Total	100.0	194.6

ceeding faster than other technologies such as the telephone, the eventual level of adoption will exceed that of the telephone. Telephone penetration in the United States has remained relatively constant at about 95 percent for the last several decades. It remains to be seen whether the Web will necessarily reach, let alone exceed, that level of penetration.

Part of the debate depends on how one views the World Wide Web. If we view it as an information medium (akin to newspapers or perhaps watching news on TV), the penetration of the Web may be constrained by the literacy of the population and interest in such information sources. As an information source, the need for information and the ability to find and make use of the information on the Web may not be universal. For example, what proportion of the population read newspapers? What proportion of the U.S. population possesses sufficient literacy skills to make use of the Web as a source of information? Data from the National Adult Literacy Survey (Kirsch et al. 1993) suggest that this is not a trivial problem in the United States. The literacy problem is likely to affect Web surveys of all types, whether probability based or not (see also Barnes 1996), in much the same way as it affects mail surveys.

If, on the other hand, the Web is seen primarily as a communication medium, its success may depend on replacing the telephone as the preferred medium of communication (and literacy issues may still hinder full use as long as it remains primarily a written rather than oral medium). The typically asynchronous and relatively impersonal communication common on the Internet (relative to the telephone) may limit its growth as a communication medium, at least in its present form.

It is possibly as an entertainment medium that the Web has the greatest potential to reach a broad sweep of the U.S. population, if the penetration levels of TVs and VCRs are anything to go by. Television is primarily a visual medium, whereas the Internet (and to a lesser extent the Web) is currently a verbal medium (i.e., predominantly text based).

There are some who claim that the Web is all of these things and more. But it is really how it is viewed by potential users that will determine how widespread the adoption of the Web or its eventual successors will be. It is fair to say that whatever the Web is now, it may be very different several years from now. At the time of writing, access to or use of the Web in the United States is far from ubiquitous and is likely to remain so for some time to come. The fact that a large proportion of the population is currently not covered by typical Web surveys may be a serious threat for attempts to develop probability-based samples of the general population for Web surveys.

The problem is not just one of how many people or what proportion of the population can be reached with a particular method but also one of how different those who are covered by the technology are from those who are not. As noted earlier, the other key issue for survey coverage relates to the differences between the covered (those who can potentially be reached with the technology) and those not covered (those without access to, or use, of the technology) on the variables of interest. As Eaton (1997) notes, "If . . . growth follows the pattern of other technologies (phones, TV, etc.) it will be at least a generation before Internet surveys are reasonably representative and they may never be fully representative, as the final group of nonadopters may remain significantly large due to cost and/or inability to adopt [*sic*] to the new technology."

The demographic differences between those with Web access and those without are already well documented. Probably the best evidence comes again from the CPS data. Several reports by the Department of Commerce have focused on the "digital divide" between the haves and the have-nots. The latest of these was released by the National Telecommunications and Information Administration (NTIA) in July 1999.

The report notes (p. xiii) that even while overall penetration of the Internet has increased, "For many groups, the digital divide has *widened* as the information 'haves' outpace the 'have nots' in gaining access to electronic resources." For example:

- Households with incomes of $75,000 and higher are more than *twenty times* more likely to have access to the Internet than those at the lowest income levels. (P. xiii, emphasis in original)
- Black and Hispanic households are roughly two-fifths as likely as White households to have home Internet access. (P. xiii)
- Those with college degrees or higher are nearly sixteen times as likely to have home Internet access. This disparity is even greater in rural areas. (P. 2)
- Even controlling for race, household composition still has a significant impact on Internet access. (P. 7)

There are also significant differences in access by age, rural/urban status, and region of the country (NTIA 1999).

While results from other surveys, such as the GVU surveys of Internet users (http://www.gvu.gatech.edu/user_surveys/), suggest that the Internet population is gradually beginning to resemble the overall population of the United States, these data are based on self-selected or volunteer samples of Internet users, whereas the CPS data are from a probability-based sample of the U.S. household population.

Even if demographic characteristics from a Web survey appear to match those of the general population, as some claim they do (e.g., Gonier 1999), this is only part of the story. The key question is whether the two populations are similar on the substantive variables of interest, whether these are attitudes, behaviors, purchase or voting intentions, or whatever. Here there is much less research evidence, as this generally requires parallel surveys using different methods.

A recent study by the Pew Research Center (Flemming and Sonner 1999) compared respondents from three groups: an RDD telephone survey, a volunteer sample of Web users, and a selected sample of Web users. They found that both online samples overrepresented males, college graduates, and the young. But, perhaps more important, they found important differences between the online samples and the telephone sample on a variety of opinion items. In several cases, the deviations from the telephone sample exceeded 20 percentage points. They conclude: "More important, there were no predictable patterns to the success or failure of the Internet surveys. Respondents who took the surveys online were not consistently more conservative or liberal than those in nationwide telephone surveys, nor were they more optimistic or pessimistic. The lack of any predictable patterns to the differences raises important questions about the utility of Internet polls to replace traditional telephone survey practices" (Flemming and Sonner 1999, p. 13). In contrast, Taylor (2000, p. 61) reports on several items from online surveys that closely parallel results from telephone surveys. However, he notes that other items from the online surveys were "substantially different" from the telephone survey results, with some of the differences diminishing with propensity weighting while other differences remained unaffected by the weighting.

In summary, these findings suggest that the "Internet population" (if such could be defined) is different from the general population of the United States in many respects. In fact we should remember that it is the *World Wide* Web, and as such its reach extends far beyond the boundaries of the United States. Even though the Web is in a state of massive growth and flux, these population differences are likely to persist for some time. The challenge for Web survey researchers is to find creative ways to make such surveys more inclusive of the desired target populations (e.g., all adults in the United States) or to limit generalizations to more restricted populations (e.g., Internet users).

Finally, to return to sampling error for a moment, there is a misguided assumption behind many Web surveys that large samples necessarily mean more valid responses, or that sample size (or, more correctly, number of respondents) is the only element in sampling error. Statistical inference is possible only with probability-based sample designs. With nonprobability designs, any efforts to generalize to a population, to present sampling errors or confidence intervals, are misleading. A key distinction must be made between scientific surveys designed to permit inference to a population, and data collection efforts where the emphasis is simply on numbers of respondents rather than representation. We should not confuse these two approaches.

NONRESPONSE ERROR

Some Web survey types, as we shall see below, have found solutions to the coverage problem. One approach is to limit the study to those with access to or use of the Web. Another approach is to overcome the limitations of restricted technology access by making the access available to all those included in the sample. Even if we could successfully identify a frame population, and that population is one of interest to clients or analysts, the problems of nonresponse may still threaten the utility of Web surveys. Nonresponse error arises through the fact that not all people included in the sample are willing or able to complete the survey. As with coverage error, nonresponse error is a function of both the rate of nonresponse and of the differences between respondents and nonrespondents on the variables of interest (Groves and Couper 1998).

For surveys where the frame cannot be identified, the problem of nonresponse is hard to define. For example, if an open invitation is issued on a Web portal to participate in a survey, the denominator of those eligible to participate is typically not known, and therefore the nonresponse rate is unknowable. This means that the measurement or evaluation of nonresponse error is tractable only in cases where the frame and the chance of selection are known (in other words, probability-based surveys).

Given this, there is currently little information available on nonresponse in Web surveys. We must rely primarily on e-mail surveys to give us a handle on the nonresponse problem. Several recent studies have compared response rates from e-mail studies to those from mail surveys of the same populations. These studies are summarized in Couper, Blair, and Triplett (1999) and Schaefer and Dillman (1998). For all but one study, the e-mail surveys failed to reach the response rate levels of the mail surveys. Several explanations may account for this difference.

One is that tried and tested motivating tools used in mail surveys (e.g., advance letters, personalized signatures, letterhead, incentives, etc.) cannot be implemented in the same way in Web surveys, and functional equivalents are yet to be developed and tested. There is at present little experimental literature on what works and what does not, in terms of increasing response rates to Web surveys (but see Crawford, Couper, and Lamias 2000). Many of the techniques developed and tested over time to increase response rates in mail surveys (see Dillman 1978, 2000) may not work the same way in fully electronic Web surveys. Finding electronic equivalents of response-stimulating efforts is work that remains to be done.

A second possible reason for lower response rates could be that technical difficulties interacting with an Internet survey (whether e-mail or Web) may discourage some from completing (or even starting) the survey, relative to the ease of completing a paper-and-pencil mail survey. Slow modem speeds, unreliable connections, low-end browsers, etc. may inhibit web survey completion from home. In some countries (and for some Internet providers), connect-time costs may deter people from doing so. In such cases, there may be real costs of completing a Web survey. Furthermore, not all browsers and hardware platforms are equal, and not all users have equal familiarity with the Web. One of the attractions of the Web over other forms of self-administered surveys is the potential for delivering rich audiovisual stimuli. But the use of such materials (requiring high-end browsers, plug-ins, cookies, etc.) may well prevent a segment of the population of interest from completing the survey. The promise of delivering full-screen streaming video over the Web is still limited by bandwidth. We do not know how many people fail to complete Web surveys for technical reasons.

Yet another explanation may relate to confidentiality concerns with respect to electronic mail. Some organizations keep records of all incoming and outgoing messages, and if the topic of the survey is particularly sensitive, this may discourage employees from completing company-related surveys at the office. Concerns about privacy and/or confidentiality may be a key factor affecting Web surveys in general. One of the distinct advantages of self-administered surveys is the ability to collect sensitive information with less social desirability bias. But concerns about the security of the Web may negate this benefit, potentially producing higher nonresponse (or less honest reporting) on surveys of sensitive topics.

All of these issues may well be temporary factors that are resolved as our knowledge of how to implement Internet surveys increases. But for now they suggest that low response rates (relative to mail) may be a big concern in Web surveys (see, e.g., Crawford, Couper, and Lamias 2000; Dommeyer and Moriarty 2000). Additional evidence for this comes from surveys of the University of Michigan student population discussed later (Couper, Traugott, and Lamias 1999) and from the Pew Center study of RDD-recruited Web survey respondents (Flemming and Sonner 1999).

As the coverage problems are overcome in Web surveys, the problems of nonresponse will likely become increasingly prominent. Furthermore, we may be already starting at a disadvantage relative to other modes of survey data collection. We do not have the buffer that allowed telephone surveys to mature before telemarketing reached troublesome levels. On the Web, direct marketers already appear to be well ahead of the survey industry. Even before the introduction of legislation to restrict unsolicited e-mail, strong norms operate on the Web to limit spamming and other types of mass mailing (see, e.g., Sheehan and Hoy 2000; Wang, Lee, and Wang 1998). These legal and ethical strictures all militate against mass e-mail as a method for soliciting participation in a survey, unless sample persons have previously agreed to receive such e-mail. There is no random generation of e-mail addresses as in RDD telephone surveys; hence the prevalence of Internet panels, whether recruited using some other medium (phone, mail) or through banner ads and other open solicitations.

In summary, sampling from a list or using RDD recruitment to identify and enlist potential respondents permits measurement of nonresponse rates, facilitates strategies targeted at the reduction of nonresponse, and provides auxiliary variables for improvement of postsurvey nonresponse adjustment. For such surveys, much work remains to be done, both on understanding why some people agree to do a Web survey while others do not and on developing procedures for minimizing the potential bias due to nonresponse. For nonprobability surveys the issue of nonresponse has little meaning.

MEASUREMENT ERROR

Given that problems of nonobservation (sampling, coverage, and nonresponse error) have swamped the discussion of survey quality in Web surveys, relatively little attention has been paid to the problem of measurement error to date (this will certainly change as the method matures).

Measurement error simply stated is the deviation of the answers of respondents from their true values on the measure. Measurement errors in self-administered surveys could arise from the respondent (lack of motivation, comprehension problems, deliberate distortion, etc.) or from the instrument (poor wording or design, technical flaws, etc.). In interviewer-administered surveys, well-trained interviewers can often explain unclear terms to respondents, keep them motivated, reassure them of the confidentiality of their answers, probe incomplete or inadequate responses, and so on. In self-administered surveys there is no interviewer to intermediate. In order to minimize respondent error, the survey instrument must be easy to understand and to complete, must be designed to keep respondents motivated to provide optimal answers, and must serve to reassure respondents regarding the confidentiality of their responses.

While the importance of question wording in influencing respondent answers is well-recognized, there is a growing literature that suggests that the design of the survey instrument (such as the placement of questions, flow of instrument, typographical features, etc.) also plays an important role, in both self-administered and interviewer-administered surveys. This has been demonstrated experimentally in self-administered paper surveys by Schwarz and Hippler (cited in Sudman, Bradburn, and Schwarz 1996, p. 123), and in Web surveys by Couper, Traugott and Lamias (1999) and Dillman et al. (1998). Smith (1995) has also shown that unintended design or layout changes can affect the responses obtained both in interviewer-administered and in self-administered surveys.

On the Web, unlike on paper, the appearance of a survey can vary from respondent to respondent because of different browser settings, user preferences, variations in hardware, and so on. Design may thus be much more important for Web surveys, both because there are more tools available to the designer (color, sound, images, animation, etc.) and because of variation in how these may be seen by respondents.

There is growing interest in the design of Web surveys (see, e.g., Dillman et al. 1998), but to date little empirical work on optimal design has been published. It is not my intention to review this issue in depth here; however, it is important to note that design may interact with the type of Web survey being conducted and the population at which the survey is targeted. In other words, the appropriateness of a particular design must be evaluated in the context of its intended goal and audience. The design of a Web poll intended primarily as entertainment might be quite different than one designed for scientific purposes. Similarly, the design of a survey on a Web site targeted at teenage girls (e.g., www.smartgirl.com) would likely have different design requirements than one aimed at older persons (e.g., www.aarp.org). The notion of a one-size-fits-all approach to Web survey design is premature. Furthermore, the Web is a fundamentally different medium than paper. The range of design options, the visual features, and the required respondent actions all differ. We have much to learn about what design knowledge and practice translates across media and what does not. There is much work to be done to determine optimal designs for different groups of respondents and types of surveys.

Another source of measurement error that is unique to panel or longitudinal surveys (often employed in Web surveys) is that of panel conditioning (or time-in-sample bias). Panel conditioning occurs through the ongoing participation of members in a panel (see Kalton and Citro 1993; Kalton, Kasprzyk, and McMillen 1989). Given their experience with the survey over time, their responses may increasingly begin to differ from the responses given by people answering the same survey for the first time. Given the current lack of a suitable sampling frame, many survey organizations are creating panels of Web survey respondents, and panel effects remain a concern for such surveys. Even though the surveys may vary over time, the mere act of participating in an ongoing panel may change respondent behavior and attitudes. While statistical adjustments for panel effects are already being used, there is much still to be learned about the nature and extent of the effects.

The discussion thus far should not imply that high-quality Web surveys are not possible. On the contrary, Web surveys offer the research community enormous opportunities for low-cost self-administered surveys using a wide variety of stimulus material (sound, images, video, etc.) that has heretofore

Table 2. Types of Web Surveys

Nonprobability Methods	Probability-Based Methods
1. Polls as entertainment	4. Intercept surveys
2. Unrestricted self-selected surveys	5. List-based samples
3. Volunteer opt-in panels	6. Web option in mixed-mode surveys
	7. Pre-recruited panels of Internet users
	8. Pre-recruited panels of full population

simply not been available or has been too costly to implement widely in interviewer-administered surveys. The use of computer-assisted methods permits the inclusion of design features (e.g., randomization, customization of wording, real-time editing, etc.) not easily implemented in paper surveys. The Web survey potential is indeed vast. However, the advent of the Web as a tool for survey data collection does not obviate the traditional concerns of representativeness and replicability.

Types of Web Surveys

With the review of various sources of error to serve as context, we proceed to a discussion of the major types of Web surveys prevalent today. There are many variations on these major themes, but the types presented below (and summarized in table 2) represent the key classes of Web survey in operation today. These approaches must be evaluated relative to their stated goals and in the context of the various sources of survey error described above. I draw a key distinction between nonprobability surveys and probability-based surveys and identify several subcategories within each. In nonprobability surveys, members of the target population do not have known nonzero probabilities of selection. Hence, inference or generalizations to that population are based on leaps of faith rather than established statistical principles.

NONPROBABILITY APPROACHES

Type 1: Web surveys as entertainment. This type of Web survey may not be considered a survey in the scientific sense of the word, but because of their popularity and the possibility that these may be confused with real surveys in the minds of some people, I mention them briefly. These surveys are intended primarily for entertainment purposes and usually state this explicitly.

Several Web sites are dedicated to the posting and completion of polls. There is often no control over what questions are posed, or who responds. These sites include misterpoll.com and survey.net. (O'Connell [1998] mentions Survey Central, Open Debate, and the Internet Voice as three sites that seek to "centralize and formalize user-created polls.") There is generally no pretense at science or representativeness, and the primary goal of these sites is as a forum for exchanging opinions. The polls often produce running tallies of results as they are collected. For example, the description on the misterpoll.com site reads as follows: "Mister Poll is interested in what you think about anything and everything. He maintains a directory of the most interesting and topical polls created by his staff and independent pollsters for your general entertainment. None of these polls are 'scientific', but do represent the collective opinion of everyone who participates. You can vote on any open poll. Results from closed polls remain in the directory for reference" (http://www.misterpoll.com).

Survey.net's claims are a little less modest. The FAQ prepared by the site's developer states the following:

> Traditional polls rely on a "human factor" and a supposed "random sample" from which to generate results. It has been argued that traditional polls could be more accurate than the SURVEY.NET system . . . someone once commented that my system was "biased towards those who are willing to participate." That's funny. I challenge anyone to show me any survey which endeavors to collect accurate information which isn't biased. The scenario doesn't exist. Unlike other surveys, SURVEY.NET is incapable of prejudice. Anyone able to participate is welcome to. . . . The ONLY bias or slant in SURVEY.NET demographics is that our respondents have access to the Internet's World Wide Web. (http://www.survey.net/sv-faq.htm)

Another type of Web survey as entertainment are the "question of the day" polls popular on many media Websites. One example is CNN Quickvote, which states: "This poll is not scientific and reflects the opinions of only those Internet users who have chosen to participate. The results cannot be assumed to represent the opinions of Internet users in general, nor the general public as a whole" (http://www.cnn.com). These instant polls are found on numerous high-traffic Web sites.

In summary, while these types of Web polls or surveys are ubiquitous, they typically do not lead to generalizations beyond reflecting the views of those who chose to respond, and they tend to have fleeting value. As such they present no real threat to legitimate (scientific) surveys as long as the data they generate are not misused. As Beniger (cited in O'Connell 1998) notes, "While amateur Web polls are completely unscientific, they are no worse than the so-called call-in polls perpetrated by many television and radio stations or the polls by printed ballots published in many newspapers or magazines."

Type 2: Self-selected Web surveys. This approach to Web surveys uses open invitations on portals, frequently visited Web sites, or (in some cases) dedicated "survey" sites. This is probably the most prevalent form of Web survey today and potentially one of the most threatening to legitimate survey enterprises because of the claims for scientific validity that are sometimes made.

Often these surveys have no access restrictions and little or no control over multiple completions. These can be viewed as the digital-age equivalents of 1-900 polls or magazine insert surveys. Generally the distinction between this type of survey and the first is that the first type usually makes few claims to generalizability, while the second type does. Furthermore, type 2 surveys often claim legitimization through the support of respectable scientific institutions.

A prominent example of this type of survey is the series of WWW User Surveys conducted by the Georgia Tech University's Graphic, Visualization, and Usability Center (GVU). The GVU recently completed its tenth Web user survey in late 1998, with over 5,000 participants. An early report on the fifth survey implies broad representation:

> For the third and fourth surveys, we were able to collect data from approximately 1 out of every 1000 Web users (based on current estimates of the number of people with Web access). For random sample surveys, having a large sample size does not increase the degree of accuracy of the results. Instead, the accuracy depends on how well the sample was chosen and other factors [Fowler 1993]. Since we use nonrandom sampling and do not explicitly choose a sample, having a large sample size makes it less likely that we are systematically excluding large segments of the population. Oversampling is a fairly inexpensive way to add more credibility to a nonrandom Web-based survey. (Kehoe and Pitkow 1996)

As noted in the methodology report on the tenth survey (http://www.gvu.gatech.edu), the GVU surveys employ "nonprobabilistic sampling." Participants are solicited using announcements on a variety of newsgroups, banner ads on high traffic sites, announcements to a mailing list maintained by GVU, and so on. The report notes that "these sites are specifically targeted to increase the likelihood that the majority of WWW users will have been given an equal opportunity to participate in the surveys." The GVU report goes on to note that compared to "other WWW user data published that utilize random techniques," the GVU data "show a bias" in the experience, intensity of usage, and skill set of the users, but not the core demographics of users. This supports the argument made earlier that matching on demographics does not guarantee an absence of bias on the variables of interest.

Analysis of the December 1998 CPS supplement shows a different picture. Comparing adults who have used the Web in the past 12 months (whether at work or home or both) to the total population of the United States, we still find differences with respect to many demographic variables. For example, 24.2 percent of the Web users have an education of high school or less, compared to 50.2 percent of the total population. On the other hand, 14.8 percent of Web users have an advanced degree, compared to 7.5 percent of the total population. Similarly, 6.9 percent of Web users are African American compared to 11.4 percent of the population; 5.1 percent of Web users are Hispanic, while 10 percent of the total population are Hispanic. Only 3.1 percent of Web users are over 65 years old, compared to 15.5 percent of the full population. As far as family income is concerned, 25.8 percent of the total population and 11.9 percent of Web users have incomes under $25,000, while 16.1 percent of the population and 29.7 percent of Web users have family incomes of $75,000 or more. Thus, claims that the Web population resembles the general population demographically are overly optimistic. The fact that respondents to self-selected Web surveys may more closely resemble the general population raises more questions than answers about possible selection bias.

Another example of a self-selected Web survey is the National Geographic Society's "Survey 2000," launched in the fall of 1998, as part of the society's efforts to mark the millennium. Invitations to complete the survey were posted on its own site and the URL was published in the society's magazines. The survey focused on geographical mobility, community, and cultural identity. The survey consisted of three main instruments (the Canadian and U.S. adult survey, the youth survey, and the international survey), and respondents were asked to choose the appropriate form on entering the survey. Over 80,000 visits to the survey site were recorded in the period the survey was open, and over 50,000 questionnaires were completed.[4] Results were published in the December 1999 issue of *National Geographic* (see May 1999).

At the completion of the survey period, the site was closed, and the following message appeared (http://survey2000.nationalgeographic.com/): "Survey 2000 has ended. We received more than 50,000 responses—*twice the minimum required for scientific validity*—and we thank everyone who contributed to this pioneering project. The information you provided will help our team of scholars answers a key question: How does where you live shape who you are? We'll look at how mobility has affected—or hasn't—respondents' sense of identity and community as well as their tastes in music, food, and reading" (emphasis added).[5]

In their analysis of the survey results, Witte and Howard (1999, p. 12) note that while the survey did not yield a random sample and the selection probabilities are unknown, "this *does not mean that the survey cannot yield representative social science data*" (emphasis in original). They claim that the selection probabilities can be "estimated" by comparing the distributions on standard demographic variables to official government statistics and applying weighting.[6] This assertion is based on the assumption that matching two "sam-

4. Multiple completions by the same person were possible.
5. Following adverse reactions on AAPORNET, these statements were subsequently removed from the SURVEY2000 site.
6. Such weighting has not yet been applied in any of the papers based on the SURVEY2000 data (e.g., Bainbridge 2000; May 1999; Witte, Amoroso, and Howard 2000).

ples" on a variety of demographic characteristics will ensure that they also match on the survey variables of interest (see also Bainbridge 1999, 2000; Witte, Amoroso, and Howard 2000).

Not surprisingly, despite the large number of respondents who decided to complete the survey after visiting the National Geographic Society's Web site, the respondents do not resemble the U.S. population (or even the Internet population insofar as it can be described) on a number of key indicators. For example, Bonnie Erickson (personal communication, 1999) compared the Canadian NGS respondents (over 5,000 respondents) to data from the 1992 Canadian General Social Survey (GSS) and concluded that the "NGS respondents are clearly a cultural as well as electronic elite." For example, while 88 percent of Canadian NGS respondents reported going to the movies, the comparable figure from the GSS was 54 percent. Similarly, 73 percent of NGS respondents and 34 percent of GSS respondents reported visiting a museum or art gallery; 97 percent of NGS and 68 percent of GSS reported reading books; 60 percent of NGS and 22 percent of GSS respondents reported attending a theater or stage performance.

Similar differences are found when comparing the results for U.S. adult respondents to the 1997 Survey of Public Participation in the Arts (SPPA), based on a national probability sample (National Endowment for the Arts 1998).[7] For example, 59.8 percent of NGS respondents reported seeing live theater in the last 12 months, compared to 24.5 percent for musical theater and 15.8 percent for nonmusical theater in the SPPA. Similarly, 77.1 percent of NGS respondents and 34.9 percent of SPPA respondents reported visiting an art museum or gallery. The self-selected nature of the NGS survey, coupled with its placement on the National Geographic Society's Web site, is likely to yield respondents who are more interested in and more likely to participate in cultural events and who are presumably also more widely traveled than the general population.

Another example of this approach to Web surveys comes from an article in the August 23, 1999, edition of the *Boston Herald*. The headline claimed that an estimated 11 million people worldwide are addicted to porn, chat, and e-mail. The article cites a study based on 17,251 responses to a questionnaire that psychologist David Greenfield posted on ABCNEWS.com. A total of 990 or 5.7 percent answered "yes" to five or more questions focusing on whether they used the computer to escape their problems and feel anxiety when they couldn't go on line. Based on these responses of self-selected respondents, the estimates were extrapolated to the world population.

There are many more surveys of this type, often done by organizations with established scientific credibility, making validity and representation

7. The 1997 SPPA was conducted as a stand-alone RDD telephone survey by Westat, Inc. A 55 percent response rate was obtained. The SPPA data are adjusted for nonresponse and weighted to CPS control totals. This comparison does not imply that the SPPA data are without flaws.

claims that go far beyond the data. It seems likely that repetitions of the *Literary Digest* debacle (see Squire 1988) will occur, that controversies such as those surrounding Shere Hite's magazine-insert surveys of sexual behavior (e.g., Hite 1979, 1981, 1987; see also Smith 1989) will proliferate. The potential fallout of such events on carefully designed and executed Web surveys is likely not to be positive.

Type 3: Volunteer panels of Internet users. This approach creates a volunteer panel by wide appeals on well-traveled sites and Internet portals. Basic demographic information is collected from these volunteers at the time of registration, creating a large database of potential respondents for later surveys. Access to these later surveys is typically by invitation only and is controlled through e-mail identifiers and passwords (the first two types of Web surveys do not restrict access). Ballot stuffing or passing the survey along to others to complete is prevented. It is this type of Web survey that has received most attention in the media and within the survey industry of late. This appears to be the fastest-growing segment of the Web survey industry, with dozens of such panels already in existence.

Selection of panel members for a particular survey may be based on quota sampling or probability sampling methods. There may be more control on the selection of respondents for a particular sample, there may be more demographic characteristics available for selection, and there may be a larger pool of potential respondents from which to draw, but these features do not change the fundamental character of this approach. As with the other types of Web surveys discussed so far, the initial panel is a self-selected sample of volunteers.

Arguably the best-known example of this approach is the Harris Poll Online. According to Harris Interactive's Web site (accessed in August 2000), their online research panel has over 6.5 million members. The site goes on to claim, "We are able to survey much larger numbers of consumers than could be done cost-effectively using telephone or mail techniques, *producing results that are much more reliable and projectable*" (emphasis added).

In a presentation at the AAPOR annual conference in May 1999, George Terhanian (vice president of Internet Research) stated that "in principle, we at Harris Interactive believe that there should be no difference aside from sampling error between survey response elicited through Harris Poll telephone research and Harris Poll Online Research as long as both surveys: ask the same questions; occur at the same time, and draw samples at random from the exact same population *or are thoughtfully and appropriately weighted*" (emphasis in original).

Gordon Black, chairman and CEO of Harris Interactive, stated that "Internet research is a 'replacement technology'—by this I mean any breakthrough invention where the advantages of the new technology are so dramatic as to all but eliminate the traditional technologies it replaces: like the automobile did to the horse and buggy. Market research began with door-to-door household surveys which gave way to telephone polling in the mid-1960s and is now making a quantum leap forward with new Internet research techniques" (Harris Interactive press release, August 1, 1999). In responding to critics of the online poll, Black stated, "It's a funny thing about scientific revolutions. People who are defenders of the old paradigm generally don't change. They are just replaced by people who embrace new ideas" (*Wall Street Journal*, April 13, 1999; see also Mitofsky 1999).

A key feature of Harris Interactive's approach is the use of propensity weighting or propensity score adjustment, designed to compensate for "biases in online samples" (Taylor 2000). The method, generally attributed to Rosenbaum and Rubin (1983), was developed to reduce selection bias in analytic statistics (e.g., the relationship between smoking and lung cancer) obtained from observational studies and was not intended to permit generalization to the full population in descriptive studies. Using parallel telephone surveys, Harris Interactive estimates the propensity of being in the online sample, based on a vector of covariates measured in both modes. The success of this approach depends on the choice of variables used for the adjustment and on the quality of the benchmark measures (the telephone surveys). Regarding the former, Harris Interactive claims that their adjustment procedures are proprietary, so except for knowing that they include both demographic and nondemographic items and use five to six items for adjustment, the details of the approach are unknown. The use of telephone surveys as benchmarks for weighting is interesting in light of arguments that low response rates for the latter is one reason Web surveys will supercede telephone surveys (see Black 1998). In model-based adjustment approaches, such as used by Harris Interactive, correct specification of the model (choice of appropriate variables for adjustment) is critical. Design-based approaches (probability samples), if executed correctly, lessen the risk from model misspecification.[8]

Greenfield Online is another example of an online panel. Their site (www.greenfieldonline.com; visited in August 2000) makes claims of being "the world's largest Internet-based marketing research panel," with over 500,000 registered respondents representing households containing over 1.7 million individuals. According to a news report from the *Boston Globe* (cited in worldopinion.com/latenews/, August 13, 1999), Harris Interactive was suing Greenfield Online for defamation after Greenfield claimed to have the largest survey base and accused Harris of "spamming" to get its survey participants.

Another example of a large online panel of volunteers is maintained by NFO Research, Inc. In fact, the NFO Interactive's Web site (www.nfow.com, visited in August 2000) had the following claim: "We at NFO Interactive are leading the market research industry into this brave new world. . . . Today, we have the largest, representative interactive panel in the world, NFO// net.source."

8. See Groves (1989, chap. 6) discussion of model- versus design-based approaches.

The site goes on to note that "NFO//net.source is our *fully representative panel* of nearly 260,000 (and growing) interactive households and over 750,000 interactive consumers" (emphasis added).

Unlike Harris Interactive, neither Greenfield nor NFO appear to use telephone surveys to evaluate or adjust the results obtained from their Web panels. Furthermore, even with a base of cooperative respondents, response rates to Web surveys on opt-in panels are not high: Harris Interactive report response rates under 10 percent for single invitation surveys of the general database, but around 20–25 percent for those directly registered or recruited through banner ads (Terhanian 2000), while Greenfield Online report response rates ranging from 20 to 60 percent (Schmidt 2000).

These are but a few examples of the many online panels that already exist or are currently being formed. A Web site, www.money4surveys.com, lists (in August 2000) over 80 links to "market research companies on the Web willing to pay for your opinions." While many of these are Internet startups, the list includes many established market research companies. Clearly, these companies are positioning themselves to be the leaders in this burgeoning field. Already we see claims and counterclaims of who was first and who is biggest, along with mention of proprietary systems and techniques.

As noted earlier, it is not the fact that a very large panel of volunteers is being used to collect systematic information on a variety of topics that is of concern, but the fact that the proponents of this approach are making claims that these panels are equal to or better than other forms of survey data collection based on probability sampling methods (especially RDD surveys). The claims go beyond saying that these panels are representative of the Internet population to claiming that they are representative of the general population of the United States. These assertions rest on the efficacy of weighting methods to correct deficiencies in sampling frames constituted by volunteers. We need thorough, open, empirical evaluation of these methods to establish their validity.

PROBABILITY-BASED METHODS

In contrast to the previous types of Web surveys, these approaches begin with probability samples of various forms. Doing so does not guarantee representativeness, as nonresponse may threaten the inferential value of these surveys. But, unlike nonprobability designs, with knowledge of the universe or frame and with information on the process of recruitment, these approaches permit measurement of the sources of nonresponse, which could be used to better inform design-based (as opposed to poststratification-only) adjustment approaches.

Given that Web access is not universal and no frame of Web users exists, there are essentially two approaches to achieving probability-based Web samples. One is to restrict the sample to those with Web access, thereby restricting the population of interest. The other is to use alternative methods (e.g., RDD, mixed mode) to identify and reach a broader sample of the population. We examine various approaches below.

Type 4: Intercept surveys. First, there are intercept-based approaches that target visitors to a Web site. In a fashion similar to that of exit polls, these approaches generally used systematic sampling to invite every *n*th visitor to a site to participate in a survey. The frame is narrowly defined as visitors to the site, eliminating coverage problems. Cookies are typically used to prevent multiple invitations to the same person. While limiting for generalizations to broader populations, this approach is very useful for customer satisfaction surveys, site evaluation, and the like (see, e.g., Feinberg and Johnson 1997).

Two key problems related to this approach are timing and nonresponse. The timing issue is one of identifying the optimal time to invite the visitor to complete the survey. If one does so on arrival at the site, one is more likely to include both those who successfully completed their task on the site and those who abandoned the site before making a purchase, finding the information they needed, etc.

CyberDialogue is one company that does intercept surveys of online visitors to clients' Web sites (see McLaughlin 2000). They use a JavaScript pop-up to invite participation and direct the visitor to the survey Web site. McLaughlin (2000) reports response rates averaging 15 percent to these invitations. The browsing behavior of both those who agree to the survey request and those who decline are tracked (using cookies) for 30 days to provide data for weighting, raising concerns about informed consent. The low response rates raise concerns about nonresponse bias. One can speculate that those who choose to compete the pop-up survey may have very different views about the Web site being evaluated than those who ignore the request.

Type 5: List-based samples of high-coverage populations. Given that Internet access or use is far from ubiquitous, should we abandon the idea of Web surveys altogether? This approach suggests not and argues that Web surveys are useful for a subset of the population with very high or complete coverage. While this limits the broad utility of Web surveys, there are still many groups for which such surveys are uniquely suited. Furthermore, as penetration increases, such uses are likely to grow and spread.

The basic approach to this type of Web survey is to begin with a frame or list of those with Web access. Invitations are sent by e-mail to participate, and access is controlled to prevent multiple completions by the same respondent or passing the survey along to others to complete. Intra-organizational surveys and those directed at users of the Internet were among the first to adopt this new survey technology. These restricted populations typically have no coverage problems (by definition), or very high rates of coverage.

Student surveys are a particular example of this approach and are growing in popularity. In a recent study on affirmative action, students at the University of Michigan were surveyed via the Web (Couper, Traugott, and Lamias 1999).

Lists of student e-mail addresses were obtained from the Registrar's office and used to invite participation in the access-controlled Web survey. All incoming students are assigned an e-mail account, and it was verified during the study that all but 5 percent of the sample had actually checked their e-mail in the period following the invitation (included in the 5 percent could be some who forwarded mail to another system or used a system that did not allow detection of e-mail use). This level of coverage exceeds that of telephone surveys of the general population and appears to be quite common in the U.S. college population.

Despite the low noncoverage, nonresponse remains a big concern in these surveys. In the Michigan survey, for example, a response rate of between 41.5 percent (excluding partial completions) and 47.1 percent (including partial questionnaires) was obtained. Similar results were obtained for another Web survey of Michigan students done at approximately the same time by Market Strategies, Inc. (Reg Baker, 1999, personal communication; see also Guterbock et al. 2000; Kennedy et al. 2000; Kwak and Radler 2000).

In summary, while coverage is less of a concern in this type of Web survey, and the population of inference, while restricted, is known, nonresponse remains a key concern. As noted above, the research on e-mail surveys suggests that much work remains to be done to bring Internet survey participation rates up to the levels of mail surveys of similar populations (Couper, Blair, and Triplett 1999; Schaefer and Dillman 1998).

Type 6: Mixed-mode designs with choice of completion method. This approach views the Web as one alternative among many that might be offered to a respondent in a mixed-mode design. These approaches are popular in panel surveys of establishments (firms, businesses, schools, farms, etc.), where repeated contacts with respondents over a long period of time are likely. Minimizing respondent burden and costs are key concerns, while the nature of the questions being asked typically leads to the conclusion that the measurement error effects of varying mode are not large.

As an example of this type, the Current Employment Statistics (CES) program at the Bureau of Labor Statistics (BLS) has been testing alternative approaches for several years, including a Web reporting option (see Clayton and Werking 1998). In 1996, BLS did a survey of reporters in three industry groups (computer and data processing services; other service industries; and state and local government). Across all three, only 10.7 percent of units contacted had a compatible browser, e-mail, and Web access on their desktop. As of March 1998, only about 11 percent of businesses are reporting by facsimile, electronic data interchange (EDI), or Web (Rosen, Manning, and Harrell 1999). This proportion is likely to increase over time.

Another example of this approach is the Census Bureau's Library Media Center (LMC) survey (see Tedesco, Zuckerberg, and Nichols 1999; Zuckerberg, Nichols, and Tedesco 1999). Citing a 1999 U.S. Department of Education report that 89 percent of public schools have access to the Internet, they mailed a questionnaire to 474 public schools and 450 private schools. The letter informed schools of a Web-based reporting option and included the URL. Because of Census Bureau security provisions, a separate letter containing the password was mailed to schools the following day. The total completion rate was 47 percent for public schools and 37 percent for private schools. The Web-based survey accounted for only 1.4 percent of the public school returns and less than 1 percent of the private school returns (Tedesco, Zuckerberg, and Nichols 1999). Given the option, respondents overwhelmingly chose the paper survey over the Web alternative.

A similar experiment was run as part of the Detroit Area Study conducted by the University of Michigan in early 1999 (see Crawford 1999). A sample of 1,500 persons in the Detroit area was sent a mail questionnaire, with an invitation to do a Web option. While the overall response rate to the survey was 60.3 percent, only 72 of the 832 respondents availed themselves of the Web option. This represents 8.7 percent of the respondents and 4.8 percent of the total sample. See Collins and Tsapogas (2000), Olson et al. (2000), and Ramirez, Sharp, and Foster (2000) for other examples of Web surveys as part of a mixed-mode strategy.

Given the proportions of returns using the Web option, this is again suggestive that nonresponse may be a problem in Web surveys (relative to mail). Furthermore, this approach is not likely to yield much cost saving over a mail-only survey, other than for very large sample sizes. The fixed costs of the e-mail development cannot be amortized over many cases, and duplicate systems are still needed to accommodate both paper and electronic returns. The approach also raises questions about the equivalence of measurement across the two media. Similar concerns would arise if Web-based responding were combined with telephone interviewing in a mixed-mode environment. Nonetheless, there may be a role for such mixed-mode approaches in repeated (panel) surveys and in business surveys, especially as Web penetration in the commercial sector increases.

Type 7: Pre-recruited panels of Internet users. This approach resembles the nonprobability approach to Web-panel creation. The key difference is that, whereas the earlier type is based on a panel of volunteers, this type of survey recruits panel members using probability sampling methods such as RDD telephone surveys. Telephone interviews are used to collect background information, identify those with Internet access, and recruit eligible persons into the Internet panel. In this way, the goal is to obtain a probability sample of Internet users or those with access to the World Wide Web (however the population of interest is defined). Following agreement to participate, panel members are sent an e-mail request to participate in the Web survey. Access is controlled through IDs embedded in URLs, personal identification numbers, and/or passwords to ensure that only those who are invited to do so complete the survey and do so only once.

If the population of interest is current users of the Internet, then coverage is not a key concern with this type of survey. Nonresponse is likely to be the biggest concern and can occur at many stages of the process. Initial nonresponse to the RDD survey yields little information on the eligibility of sample persons, their sociodemographic characteristics, and their reasons for not participating. On the other hand, if the goal is to compare results of Web surveys to those of RDD telephone surveys, the *relative* effect of nonresponse at this stage should be the same. Further sample losses may occur during the telephone interview, where respondents (deliberately or otherwise) claim not to have Internet access or fail to provide a valid e-mail address. Finally, even among those who have Web access and agree to do the Web survey, many may fail to do so when sent the invitation. However, for this latter group at least, measures can be obtained during the telephone recruitment that could inform an examination of nonresponse bias, by comparing those who did the survey to those who did not, on the variables of interest from the phone survey. This approach can thus be useful for exploring the nonresponse bias associated with Web surveys.

Nonresponse may occur at many different stages of the process, but unlike the case in volunteer panels, the nonresponse rate is measurable. Furthermore, data can be collected at the earlier stages to examine bias from nonresponse at later stages of the process. In other words, one can collect demographics and other data on Internet users and nonusers, and on volunteers and nonvolunteers to help understand the nature of the coverage and nonresponse biases. This approach also permits the measurement of mode effects (telephone vs. Internet) and direct comparisons of results to comparable RDD surveys.

The Pew Research Center study mentioned earlier is one example of this type (for another, see Tortora 2000). Flemming and Sonner (1999) report that, on average, 36 percent of Internet users contacted in telephone surveys provide an e-mail address. Of these, about one-third actually participated in a subsequent Internet poll. This suggests a fairly dismal overall response rate, considering the product of the initial RDD response rate, the provision of an e-mail address among eligible Internet users, and the proportion who actually completed the Web survey.

Despite the rapid attrition at each step of the process, the Pew Center's recruited sample still differed significantly from a volunteer Web sample on a number of political items. Furthermore, both Web samples differed in many respects from an RDD telephone survey conducted at the same time. Much remains to be done to understand the dynamics of nonresponse between the telephone and Internet modes. But, in theory at least, this approach begins with a probability sample of the full (telephone) population, and assuming no nonresponse error permits inference to the population of Internet users.

Type 8: Probability samples of full population. The last type of Web survey is unique in that it is the only method that has the *potential* for obtaining a probability sample of the full population, not just those who currently have Web access. In some respects this approach is similar to type 7 Web surveys, in that one starts with a probability sample of the target population and uses non-Internet approaches to elicit initial cooperation (e.g., using RDD telephone surveys). Whereas type 7 surveys continue only with those who report having Web access, this approach provides the necessary equipment and tools to potential respondents in exchange for their participation in subsequent Web surveys. This is the only approach that allows generalization beyond the current population of Internet users. Because of the high cost of recruitment, this approach invariably employs a panel design.

The approach has its origins in attempts in the 1980s to use Videotex, Minitel, and other television-based text information systems in Europe to conduct surveys. The Dutch Telepanel, for example, was begun in 1986 and involved placement of a low-end computer in selected respondents' homes in exchange for completion of regular (weekly) surveys by members of the household (see Saris 1998). The media audience measurement devices placed in homes by A. C. Nielsen, Arbitron, and others are in a similar spirit.

A key problem with these approaches has been the low initial response rate to the recruitment interview and the low number of those interviewed who subsequently agree to participate in the panel. While generally not reported, it is estimated that the initial response rates for the Dutch Telepanel may be on the order of 20–30 percent. Once a household has agreed to accept the unit, attrition from the panel is generally low (Felix and Sikkel 1998; Saris 1998), and response rates to each individual survey sent to cooperating households are high.

One company, InterSurvey (www.intersurvey.com), has recently adopted this approach for the Web, recruiting panel members using RDD telephone surveys, and providing panel members with Web TV units and free Internet access in exchange for their participation in the panel.[9] The company recently announced having recruited its 100,000th member, with a goal of 250,000 panel members by 2001 (InterSurvey press release, July 21, 2000).

Full details about the success of InterSurvey's approach are yet to emerge. Rivers (2000) reports an overall response rate of around 56 percent for the RDD recruitment effort (80 percent contact rate × 70 percent cooperation rate). The installation rate (proportion of households agreeing to receive the unit who actually install it) is over 80 percent, while response rates to the initial profile surveys are averaging 93 percent. A key advantage of this approach is that information about the nonrespondents can be obtained at each stage, permitting detailed examination of likely nonresponse bias and panel attrition effects. The data are weighted to compensate for errors due to sampling, coverage, and nonresponse (Krotki 2000). Unlike the Harris Interactive approach, the probabilities of selection are known at each stage, as is the target population, permitting both standard weighting class adjustment and

9. The company recently changed its name to Knowledge Networks.

poststratification. Several recent reports present results from InterSurvey panels (e.g., Frankovic 2000; Greenberg 1999; Nie and Erbring 2000).

This approach potentially solves two of the major problems of Web surveys: (1) coverage and (2) browser compatibility problems or standardization. Coverage is solved by providing Web access in exchange for participation. Compatibility problems are circumvented by providing every panel member with identical equipment. This permits the delivery of survey instruments (including audiovisual material) in a consistent and reliable way to all panel members.

Nonresponse remains a concern for this approach. However, the initial RDD recruitment attempt may potentially provide information on both respondents and nonrespondents, permitting estimation of the extent of nonresponse bias. Concerns about panel conditioning can be addressed if a rotating panel design is used (as proposed by InterSurvey). Such a design permits one to measure the severity of the panel conditioning effect, and, if necessary, statistical adjustments can be used to control for the effects of panel membership on responses. Given that this approach is based on probability sampling methods, estimates of reliability can be produced and direct comparisons made to equivalent surveys using more traditional methods (such as telephone surveys). Despite the numerous advantages of this approach, it is likely to be the most expensive form of Web survey, requiring resources both for recruitment and panel maintenance. Whether these expenses can be justified by the improved quality obtained from a true probability-based survey is unknown at this stage. Nonetheless, this approach shows great promise for replacing high-quality probability-based surveys using more traditional methods.

Summary and Conclusion

As already noted, Web surveys are proliferating at an almost incomprehensible rate. The Internet has truly democratized the survey-taking process. However, one outcome of this process is that the quality of surveys on the Internet varies widely, from a simple set of questions intended to entertain to full probability-based designs intended to describe the general population. The need to educate consumers of surveys (whether sponsors/clients or the general public) regarding quality indicators is already apparent. To dismiss all Web surveys because of the overenthusiastic claims of the few is a mistake. Similarly, to assume that no major embarrassments will occur as the method matures is unrealistic.

Web surveys already offer enormous potential for survey researchers, and this is likely only to improve with time. The challenge for the survey industry is to conduct research on the coverage, nonresponse, and measurement error properties of the various approaches to Web-based data collection. We need to learn when the restricted population of the Web does not matter, under which conditions low response rates on the Web may still yield useful information, or how to find ways to improve response rates to Web surveys. While the sampling problem presents enormous challenges for Web surveys of the general population, the problems of noncoverage and nonresponse are not unique to this method, and statistical adjustments are commonly employed in survey research as an effort to compensate for these deficiencies. However, the extent to which weighting the results from volunteer panels can reliably produce reasonable estimates is unknown. For example, the *relative* quality of low response rate RDD surveys and volunteer Web panels must be evaluated. We must also learn how to optimally design Web surveys and maximize the benefits of the rich audiovisual and interactive self-administered medium we now have at our disposal. Only by fully understanding both the benefits and the drawbacks of this new method can we fully exploit the potential of Web surveys. We are faced with exciting opportunities to explore a new method of data collection to take the survey industry into the twenty-first century. Solid research and open sharing of research methods and results are needed to ensure that we do so in a responsible and informed manner.

References

Bainbridge, W. S. 1999. "Cyberspace: Sociology's Natural Domain." *Contemporary Sociology* 28(6):664–67.

———. 2000. "Validity of Web-Based Surveys: Explorations with Data from 2,382 Teenagers." Unpublished manuscript, National Science Foundation.

Barnes, S. B. 1996. "Literacy Skills in the Age of Graphical Interfaces and New Media." *Interpersonal Computing and Technology* 4 (3–4):7–26.

Black, G. S. 1998. "The Advent of Internet Research: A Replacement Technology." Paper presented at the annual meeting of the American Association for Public Opinion Research, St. Louis, May.

Clayton, R. L., and G. S. Werking. 1998. "Business Surveys of the Future: The World Wide Web as a Data Collection Methodology." 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. O'Reilly. New York: Wiley.

Collins, M. A., and J. Tsapogas. 2000. "An Experiment in Web-Based Data Collection." Paper presented at the annual meeting of the American Association for Public Opinion Research, Portland, OR, May.

Couper, M. P., J. Blair, and T. Triplett. 1999. "A Comparison of Mail and E-Mail for a Survey of Employees in Federal Statistical Agencies." *Journal of Official Statistics* 15(1):39–56.

Couper, M. P., M. Traugott, and M. Lamias. 1999. "Effective Survey Administration on the Web." Paper presented at the Midwest Association for Public Opinion Research, November.

Crawford, S. 1999. "The Web Survey Choice in a Mixed Mode Data Collection." Unpublished manuscript, University of Michigan.

Crawford, S., M. P. Couper, and M. Lamias. 2000. "Web Surveys: Perceptions of Burden." Paper presented at the annual meeting of the American Association for Public Opinion Research, Portland, OR, May.

Dillman, D. A. 1978. *Mail and Telephone Surveys; The Total Design Method*. New York: Wiley.

———. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York: Wiley.

Dillman, D. A., R. D. Tortora, J. Conradt, and D. Bowker. 1998. "Influence of Plain vs. Fancy Design on Response Rates for Web Surveys." Paper presented at the joint statistical meetings of the American Statistical Association, Dallas, August.

Dommeyer, C. J., and E. Moriarty. 2000. "Increasing the Response Rate to E-Mail Surveys." Paper presented at the annual meeting of the American Association for Public Opinion Research, Portland, OR, May.

Eaton, B. 1997. "Internet Surveys: Does WWW Stand for 'Why Waste the Work?'" *Quirk's Marketing Research Review* (www.quirks.com), June.

Feinberg, S. G., and P. Y. Johnson. 1997. "Designing and Testing Customer Satisfaction Surveys on WWW Sites." In *Proceedings of the Society for Technical Communication's 44th Annual Conference*, pp. 298–301. Arlington, VA: STC..

Felix, J., and D. Sikkel. 1998. "Attrition Bias in Telepanel Research." Paper presented at the International Workshop on Household Survey Nonresponse, Bled, Slovenia, September.

Flemming, G., and M. Sonner. 1999. "Can Internet Polling Work? Strategies for Conducting Public Opinion Surveys Online." Paper presented at the annual meeting of the American Association for Public Opinion Research, St. Petersburg Beach, FL, May.

Fowler, F. J. 1993. *Survey Research Methods*. 2d ed. Newbury Park, CA: Sage.

Frankovic, K. 2000. "Internet Panel Response to the 'State of the Union' Address: An Experiment." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.

Gonier, D. E. 1999. "The Emperor Gets New Clothes." Paper presented at the Advertising Research Foundation's Online Research Day, Los Angeles, January.

Greenberg, A. 1999. "Women and the Economy." iVillage Election 2000, Women's Electorate Project Polls (http://ivillage.com).

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

Groves, R. M., and M.P. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: Wiley.

Guterbock, T. M., B. J. Meekins, A. C. Weaver, and J. C. Fries. 2000. "Web versus Paper: A Mode Experiment in a Survey of University Computing." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.

Hite, S. 1979. *The Hite Report on Female Sexuality*. New York: Knopf.

———. 1981. *The Hite Report on Male Sexuality*. New York: Knopf.

———. 1987. *Women and Love: A Cultural Revolution in Progress*. New York: Knopf.

Hoffman, D. L., T. P. Novak, and A. E. Schlosser. 2000. "The Evolution of the Digital Divide: How Gaps in Internet Access May Impact Electronic Commerce." *Journal of Computer-Mediated Communication* 5(3):1–55 (http://www.ascusc.org/jcmc/vol5/issue3/hoffman.html).

Kalton, G., and C. F. Citro. 1993. "Panel Surveys: Adding the Fourth Dimension." *Survey Methodology* 19(2):205–15.

Kalton, G., D. Kasprzyk, and D. B. McMillen. 1989. "Nonsampling Errors in Panel Surveys." In *Panel Surveys*, ed. D. Kasprzyk, G. Duncan, G. Kalton, and M.P. Singh, pp. 249–70. New York: Wiley.

Kehoe, C. M., and J. Pitkow. 1996. "Surveying the Territory: GVU's Five WWW User Surveys." *World Wide Web Journal* 1(3):77–84.

Kennedy, J. M., G. Kuh, S. Li, J. Hayek, J. Inghram, N. Bannister, and K. Segar. 2000. "Web and Mail Survey: Comparison on a Large-Scale Project." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.

Kirsch, I. W., A. Jungeblut, L. Jenkins, and A. Kostad. 1993. *Adult Literacy in America*. Washington, DC: National Center for Education Statistics.

Krotki, K. P. 2000. "Sampling and Weighting for Web Surveys." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.

Kwak, N., and B. T. Radler. 2000. "Using the Web for Public Opinion Research: A Comparative Analysis between Data Collected via Mail and the Web." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.

Mack, J. 1999. "Is Web Growth Tapering Off?" *ZDNET News* (www.zdnet.com/zdnn/), October 20.

May, V. A. 1999. "Survey 2000: Charting Communities and Change." *National Geographic* 196(6):130–33.

McLaughlin, T. 2000. "Customer Database Research: Guidelines for Complete and Ethical Data Collection." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.

Mitofsky, W. J. 1999. "Pollsters.com." *Public Perspective* (June/July): 24–26.

National Endowment for the Arts (NEA). 1998. *Survey of Public Participation in the Arts: Summary Report*. Washington, DC: NEA.

National Telecommunications and Information Administration (NTIA). 1999. *Falling through the Net: Defining the Digital Divide.* Washington, DC: U.S. Department of Commerce.
Nie, N. H., and L. Erbring. 2000. *Internet and Society: A Preliminary Report.* Palo Alto, CA: Stanford University, Stanford Institute for the Quantitative Study of Society.
O'Connell, P. L. 1998. "Personal Polls Help the Nosy Sate Curiosity." *New York Times*, June 18.
Olson, L., K. P. Srinath, M. C. Burich, and C. Klabunde. 2000. "Use of a Website Questionnaire as One Method of Participation in a Physician Survey." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.
O'Muircheartaigh, C. 1997. "Measurement Errors in Surveys: A Historical Perspective." In *Survey Measurement and Process Quality*, ed. L. E. Lyberg, P. P. Biemer, M. Collins, E. D. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin. New York: Wiley.
Ramirez, C., K. Sharp, and L. Foster. 2000. "Mode Effects in an Internet/Paper Survey of Employees." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.
Rivers, D. 2000. "Probability-Based Web Surveying: An Overview." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.
Rosen, R., C. Manning, and L. Harrell. 1999. "Controlling Nonresponse in the Current Employment Statistics Survey." Paper presented at the International Conference on Survey Nonresponse, Portland, OR, October.
Rosenbaum, P. R., and D. B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1):41–55.
Saris, W. E. 1998. "Ten Years of Interviewing without Interviewers: The Telepanel." 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. O'Reilly. New York: Wiley.
Schaefer, D. R., and D. A. Dillman. 1998. "Development of a Standard E-Mail Methodology: Results of an Experiment." *Public Opinion Quarterly* 62(3):378–97.
Schmidt, J. 2000. "Bill of Rights of the Digital Consumer: The Importance of Protecting the Consumer's Right to Online Privacy." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.
Sheehan, K. B., and M. G. Hoy. 2000. "Dimensions of Privacy Concern among Online Consumers." *Journal of Public Policy and Marketing* 19(1):62–73.
Smith, T. W. 1989. "Sex Counts: A Methodological Critique of Hite's *Women and Love*." In *AIDS, Sexual Behavior and Intravenous Drug Use*, ed. C. F. Turner, H. G. Miller, and L. E. Moses, pp. 537–47. Washington, DC: National Academies Press.
———. 1995. "Little Things Matter: A Sampler of How Differences in Questionnaire Format Can Affect Survey Responses." *Proceedings of the American Statistical Association, Survey Research Methods Section*, pp. 1046–51.
Squire, P. 1988. "Why the 1936 *Literary Digest* Poll Failed." *Public Opinion Quarterly* 52(1): 125–33.
Sudman, S., N. M. Bradburn, and N. Schwarz. 1996. *Thinking about Answers; The Application of Cognitive Processes to Survey Methodology.* San Francisco: Jossey-Bass.
Taylor, H. 1999. "On-Line Population Spends an Average of Six Hours on the Internet or Web per Week." Press release, Harris Poll no. 18, March 24.
———. 2000. "Does Internet Research Work? Comparing Online Survey Results with Telephone Survey." *International Journal of Market Research* 42(1):51–63.
Tedesco, H., R. L. Zuckerberg, and E. Nichols. 1999. "Designing Surveys for the Next Millennium: Web-Based Questionnaire Design Issues." *Proceedings of the Third ASC International Conference*, Edinburgh, September, pp. 103–12.
Terhanian, G. 1999. "Understanding Online Research: Lessons from the Harris Poll Online." Paper presented at the annual conference of the American Association for Public Opinion Research, St. Petersburg Beach, FL, May.
———. 2000. "How to Produce Credible, Trustworthy Information through Internet-Based Survey Research." Paper presented at the annual conference of the American Association for Public Opinion Research, Portland, OR, May.
Tortora, R. 2000. "A Comparison of Incentive Levels for a Panel of Internet Users: Some Preliminary Results." Paper presented at the Nebraska Symposium on Survey Research, Lincoln, April.

Wang, H., M. K. O. Lee, and C. Wang. 1998. "Consumer Privacy Concerns about Internet Marketing." *Communications of the ACM* 41(3):63–70.

Witte, J. C., and P. E. N. Howard. 1999. "Digital Citizens and Digital Consumers: Demographic Transition on the Internet." Unpublished manuscript, Northwestern University.

Witte, J. C., L. M. Amoroso, and P. E. N. Howard. 2000. "Method and Representation in Internet-Based Survey Tools: Mobility, Community, and Cultural Identity in Survey2000." *Social Science Computer Review* 18(2):179–95.

Wright, T., and H. J. Tsao. 1983. "A Frame on Frames: An Annotated Bibliography." In *Statistical Methods and the Improvement of Data Quality*, ed. T. Wright, pp. 25–72. New York: Academic Press.

Zuckerberg, A., E. Nichols, and H. Tedesco. 1999. "Designing Surveys for the Next Millennium: Internet Questionnaire Design Issues." Paper presented at the annual conference of the American Association for Public Opinion Research, St. Petersburg Beach, FL, May.