

WEB SURVEY METHODS

INTRODUCTION

MICK P. COUPER
PETER V. MILLER

A key characteristic of Web surveys is their diversity. Unlike other modes of data collection, where the method tells us something about both the sampling process and the method of data collection, the term “Web survey” is too broad to give us much useful information about how the study was carried out. For example, referring to an RDD telephone survey describes both the method of sampling (in part) and the mode of data collection. But there are so many different ways to identify sampling frames for Web surveys, to invite people to complete such surveys, and to administer surveys over the Internet (see Couper 2000) that the term “Web survey” conveys little evaluative information. The implications of this diversity are twofold. First, broad generalizations or claims about Web surveys relative to other methods of data collection are ill-advised. Second, much more detail about the process is needed in order for the reader to make judgments about the quality of the process itself or about the resulting data. The papers in this special issue reflect some of the many ways that the Internet can be used—whether alone or in combination with other methods—to conduct surveys.

Despite their relatively short history, Web surveys have already had a profound effect on survey research. The first graphic browser (NCSA Mosaic) was released in 1992, with Netscape Navigator following in 1994 and Internet Explorer in 1995. The first published papers on Web surveys appeared in 1996. Since then, there has been a virtual explosion of interest in the Internet generally, and World Wide Web specifically, as a tool for survey data collection (see www.WebSM.org for a detailed bibliography). This is not to say that the early claims that Web surveys will make all other methods of data collection obsolete have come to pass. But it is fair to say that the methodological attention that Web surveys have received has exceeded other modes in a similar time period. In part, this is because the relative cost of Web surveys makes them a more accessible method of data collection than telephone or face-to-face surveys. In addition, the computerized nature of Web surveys facilitates conducting

MICK P. COUPER is a research professor in the Survey Research Center, Institute for Social Research, University of Michigan and in the Joint Program in Survey Methodology (JPSM). He is the author of *Designing Effective Web Surveys* (Cambridge, 2008). PETER V. MILLER is associate professor of Communication Studies at Northwestern University (p-miller@northwestern.edu). He is a former editor of *Public Opinion Quarterly*. Address correspondence to Mick P. Couper; e-mail: mcouper@isr.umich.edu

experiments and speeds the process of collecting data, leading—for better or worse—to more research papers in a shorter time frame.

The survey enterprise is becoming more “democratized” on the Web. Individuals with little or no financial support and training now routinely seek survey cooperation from other Web users. This trend is in line with the more general capability and culture of Web communication—users of the medium are frequently content producers rather than just passive recipients of information. Prominent Web facilities are structured to encourage interaction. Blogs and social networking sites depend on Web users’ interest in sharing and soliciting information from (even) a world-wide set of partners. Survey capability is a common feature of software designed for Web communication. Because of the Web’s egalitarian ethos and the relative ease with which Web surveys can be conducted, control over the content and method of these efforts is now shared by survey experts and lay enthusiasts. Just as Internet sites challenge traditional news and scholarly publication outlets, variegated amateur Web surveys now help to constitute the contemporary information environment along with more established data sources.

In some respects, this development is salutary. Survey research has long been enabled and constrained by its sponsorship: the topics and methods investigated have been shaped by the interests of funding sources. To the degree that survey facilities are inexpensive and widely available, the research agenda can expand to include important topics that do not attract support from the usual sources. Even for the established sponsors of surveys, particularly in the commercial sector, the Web offers the potential of fast and cheap investigation of issues that would have been deemed too costly in other survey modes. But, despite the popularity of Web surveys, and their obvious value for a wide range of research endeavors, there remain fundamental concerns with the method. The key challenges are errors of nonobservation (see Groves 1989) or issues of representation, especially with regard to more general populations of interest.

While we have seen dramatic growth in Internet access and use in both the developed and developing world, the proportion of those with Internet access appears to have reached a plateau in the USA in recent years. Further, regardless of the actual level of penetration, substantial proportions of the population remain without access to the Internet, and the differences between those with and without access show no signs of diminishing. In addition, the nature of Internet use is fundamentally different from that of, say, the telephone (although we are seeing changes in the latter with the introduction of cell phones and Internet-based telephony), meaning that frames of Internet users in a form suitable for sampling do not—and likely will not—exist.

Two broad approaches to dealing with these inferential problems can be identified. The first—the “design based” approach—attempts to build probability-based Internet panels by using other methods (such as RDD samples) for sampling and recruitment and—where necessary—providing Internet access to those without. This approach is typified by Knowledge Networks in the USA

and CentERdata's MESS panel in the Netherlands, although several other similar efforts are underway. The second—the “model-based” approach—begins with a volunteer or opt-in panel of Internet users, and attempts to correct for representational biases using propensity score adjustment or some other weighting method. See the papers by Callegaro and DiSogra and Toepoel, Das and Van Soest in this issue.

The debate over alternative ways to use Web surveys to make general population inferences is healthy and extends beyond the Web survey realm. With increasing threats to earlier design-based approaches, coming from the rapidly increasing costs associated with maintaining high response rates in face-to-face surveys, or the challenges to telephone surveys posed by the rise of cell phone only households, increased use of Internet telephony and declining response rates (see the POQ special issue on cell phones, Volume 71, number 5, 2007), the profession should welcome increased attention to model-based alternatives that is driven in part by the Web survey challenge.

However, transparency is critically needed to facilitate careful scientific evaluation of the alternative approaches. Proprietary claims and failures to provide details of the methods used will hinder the objective and open evaluations of alternatives that have been a hallmark of the survey profession for most of its existence. Now more than ever, we believe that the standards of disclosure represented in the AAPOR code are needed to encourage such evaluation, not only with respect to Web surveys, but also with respect to all modes of data collection. The paper by Toepoel, Das and Van Soest in this issue examines one of the salient concerns attendant to Web panels—the extent to which duration of panel participation affects the quality of survey participation. Callegaro and DiSogra propose a set of disclosure standards for response rates in probability-based and opt-in Web panels. Their work should hasten the development of metrics for assessing Web panel quality, at least with respect to response rates.

Web surveys do not appear to offer an antidote for the declining response rates affecting all modes of survey data collection. In fact, Web surveys appear to be at a disadvantage relative to other modes as a recent meta-analysis suggests (Lozar Manfreda et al. 2008). The response rate challenge affects all types of Web surveys, from list-based samples to pre-recruited probability-based panels and opt-in or volunteer panels. Regarding the latter, the early promise of an endless supply of willing respondents has not come to pass. Instead, the situation (in the USA at least) is one in which an increasing number of panel vendors and surveys appear to be chasing an ever-dwindling group of willing respondents, raising further concerns about the selectivity of respondents in such surveys. Completion rates for surveys of members of opt-in panels have declined to the low single digits. Cumulative response rates for probability-based Web panels are low as well, though final stage completion rates may be high (see Callegaro and DiSogra, Malhotra, and Bailenson et al. in this issue). Low response rates are a warning of potential trouble, but the hard business of identifying actual nonresponse bias in Web surveys is in its infancy. Much

work remains in identifying and mitigating the effects of nonobservational or selection biases.

Mixed-mode surveys involving the Web have received a lot of attention in recent years. If the goal is to increase overall response rates, the recent evidence is not very encouraging. Providing sample persons a choice of mode does not appear to increase overall response rates although the way in which the mode choices are sequenced does appear to affect the proportion of respondents completing the survey online. In other words, if the goal is one of cost reduction through adding Web surveys, or increased coverage through adding an alternative to the Web, mixed-mode designs show promise (see Rookey et al. in this issue). But they do not appear to be a solution to the problem of declining response rates. Mode comparisons, sometimes performed in connection with mixed-mode survey designs, are represented in this issue. Kreuter et al. and Heerwegh and Loosveldt explore the quality of responses obtained from Web surveys and other modes (IVR, CATI, and face-to-face), finding different advantages and disadvantages of the Web mode.

Continuing the focus on measurement error or data quality, considerable research attention has been paid to the issue of instrument design through experimentation in Web surveys. Indeed, this is one of the clear advantages of the method, with implications beyond the design of Web surveys. As a tool for research on questionnaire design and response processes, the Web has many strengths. The papers by Malhotra, Delavande and Rohwedder, and Galesic et al. in this issue illustrate some of the many ways in which Web survey experimentation can inform us about fundamental response processes. The research suggests that there are many exciting ways to improve the question asking and answering process and the quality of the resultant data using Web surveys. For applications where the representational challenges are of less concern, Web surveys offer opportunities for new ways to ask survey questions. The graphical and interactive nature of the Web, along with the ease of randomization and control of stimulus presentation means that researchers can go well beyond the traditional approaches to eliciting information from survey respondents. The papers by Delevande and Rohwedder and Bailenson et al. exemplify some provocative possibilities.

Have Web surveys lived up the hope and expectations of some, or the fears of others? Some claimed that Web surveys would replace other modes of data collection (especially telephone surveys); others saw Web surveys contributing to the disintegration or dilution—if not total demise—of the survey enterprise. Neither of these extremes has come to pass. Web surveys, like other methods of survey data collection, have strengths and weaknesses. Much of the research over the past several years has focused on identifying these strengths and weaknesses and finding ways to overcome the former and exploit the latter. The papers in this special issue continue this trend.

There are several areas where we are likely to see more research attention in the next few years. One, as mentioned already, is increased attention to the

inferential challenges facing all survey methods in a changing society. Another is that the Web itself is changing. With talk of Web 2.0 and the growth in social networking sites (see Couper 2007), we need to figure out what this will mean for traditional surveys and how we can use the increased interactivity of the Web to improve measurement. Related to this is the increased move toward mobile Web applications. The distinction between a telephone and Web browser is fast disappearing. Much work remains not only on the challenges such devices present for instrument design, but also on the opportunities these new tools present for data collection that is not restricted in time and space as traditional methods often are. The ability to do electronic diary data collection, to conduct surveys in response to actions taken by a respondent and the like, present an exciting area of future research for Web surveys.

Our sincere thanks go to the American Association for Public Opinion Research for sponsoring this special issue of *Public Opinion Quarterly* and distributing it free of charge to all interested researchers. In addition, we thank all the scholars who submitted papers, all of our many dedicated referees and the authors whose papers were selected for the issue (who endured our editing). Finally, we greatly appreciate the patience and routinely excellent work of Phyllis Silverstein, POQ Managing Editor, and Melanie Ortega, Production Editor at Oxford University Press.

References

- Couper, Mick P. 2000. "Web Surveys: A Review of Issues and Approaches." *Public Opinion Quarterly* 64(4):464–94.
- Couper, Mick P. 2007. "Whither the Web: Web 2.0 and the Changing World of Web Surveys." In *The Challenges of a Changing World: Proceedings of the Fifth International Conference of the Association for Survey Computing*, eds. M. Trotman et al., pp. 7–16. Berkeley, UK: ASC.
- Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: Wiley.
- Lozar Manfreda, Katja, Michael Bosnjak, Jernej Berzelak, Iris Haas, and Vasja Vehovar. 2008. "Web Surveys versus Other Survey Modes: A Meta-Analysis Comparing Response Rates." *International Journal of Market Research* 50(1):79–104.