What Have We Learned About Email Surveys?

Dulane Moran, Research Into Action, Portland, OR
April Armstrong, Research Into Action, Portland, OR
Jun Suzuki, Research Into Action, Portland, OR

ABSTRACT

This paper considers the advantages and disadvantages of email surveys versus phone surveys through a comparison of two mixed-mode projects. Research indicates that it has become increasingly difficult to maintain response rates with telephone surveys, causing the costs of data collection to rise. After conducting a literature review on the benefits and drawbacks of email surveys, the authors consider their own experience and review survey results for signs of bias in email and phone survey data. This experience is compiled from recent projects where they collected data using a random sampling technique for both phone and email simultaneously. In all referenced projects, the researchers had access to population lists. While we found that the data collected via email showed no significant differences from the data collected via phone, we found indications that email respondents become less engaged in the survey experience as it progressed and that the two modes may differ in demographic over- or under-representation. Finally, the tenants of quality survey research remain important, regardless of the mode employed.

Introduction

Program evaluation relies heavily on survey research, most commonly telephone surveys. Research indicates that it has become increasingly difficult to maintain response rates with telephone surveys, causing data collection costs to rise (Fricker et al. 2005). In response to pressure to complete surveys with a larger number of respondents in a shorter amount of time, at less cost, evaluators increasingly turn to email or web-based surveys. Often, an email survey emerges as an option when resources are limited or a large sample size is required—data are collected and analyzed and the report is written. The phone survey does not occur, making it difficult to compare results across methods, given the same population and set of questions.

Although email surveying presents many practical advantages, there are concerns that the resulting data may allow for increased nonresponse bias, self-selection bias, and other skewing factors compared to phone surveys (Peytchev, Baxter & Carley-Baxter 2009). This paper examines the current literature about email surveys and the expected or potential sources of bias.

Identifying potential response bias requires having the same population respond to the same set of questions simultaneously—varying only the survey method. In the past two years, we have had several opportunities to launch both email and phone surveys simultaneously using random sampling of the same population. These projects provided an opportunity to review the response patterns and respondent characteristics between data collected through email and phone methods.

To prepare this paper, we analyzed the responses for two projects in which email and phone surveys were used with the same population at the same time. We considered the results in light of existing literature on email versus phone survey methods and assessed whether or not there are signs that the data collected have a pattern of bias. Specifically, we reviewed our data in light of literature (Smyth et al. 2009) that indicates email respondents:

1. Are more likely to choose extreme values on scales;
2. May be less constrained by trying to please the person on the phone; and

3. May offer more extensive verbatim responses.

We also discuss the pre-existing conditions under which an email survey is more desirable than a phone survey: the existence of (valid) email addresses, an appropriate population, the source of the list, and the relationship between the sponsor and the respondent.

Literature Review: What Do We Know?

Groves and Kahn (1979) discuss the rapid evolution of data collection methods over the past 30 years. When survey research began, face-to-face interviews were the preferred method of data collection. Over time, this became prohibitively expensive, and most data collection shifted to mail and telephone. Random digit dialed (RDD) surveys gained popularity in the 1970s, when research suggested that telephone data was of comparable quality to data collected face-to-face (Groves and Kahn 1979), but at a lower cost. In recent years, there are signs that telephone interviewing has taken on many of the problems previously associated with face-to-face interviews, including high costs. The causes of increasing cost include the shift from landlines to cell phones and a growing "do not call" registry, both of which result in the necessity to make more calls to collect each complete response. The costs associated with achieving a similar response rate to studies conducted 20 years ago are now much higher (Holbrook et al. 2007).

Email surveys offer substantial financial advantages over traditional telephone surveys. In the past decade, Internet usage and email communication has increased. Fewer people are maintaining dedicated landline telephones in their homes (Nie, Hillygus & Erbring 2002), and both landline and wireless telephone lines commonly have caller identification that alerts the recipient of an unknown or blocked caller. This problem is aggravated with cell phones, which are almost universally equipped with caller identification and voice mail. According to the Nielsen Company, by 2008 almost 20 million households did not have a landline and relied solely on wireless telephone service (Nielsen Wire 2008). That number is likely higher today. According to Internet World Stats (2010), 77% of the North American population is using the Internet. The explosion of email access and smart phone/Wi-Fi devices means that people not only accept being contacted electronically, they expect it. Survey methodologies are evolving to catch up to the rapid advances in Internet connectivity. Email surveys are poised to become increasingly robust as digital communication becomes the preferred method of contact for many people.

Email surveys have several practical advantages: they can be distributed easily, they require no telephone charges or interviewer time, and respondents can complete the survey at a time convenient to them. After the survey is designed and distributed, there is no additional incremental cost for each additional complete. The responses collected through online survey tools are already formatted into a database and can be analyzed easily, making it appealing to the research firms responsible for analyzing the data (Medin, Roy & Ann 1999). Like telephone surveys, online surveys can also present response categories in a random order and use skip patterns to present respondents with only the questions that are relevant to them. Similar to mail surveys, online and email surveys allow for presentation of visual material, photos, logos, and equipment, and control over whether or not a participant is exposed to an image, and when (Tannenbaum & Feldman 2001).

Email surveys may also offer advantages in terms of convenience by providing another path for reaching populations who may be difficult to contact by telephone. Difficult to reach groups include younger people who are less likely to be contacted via RDD survey approaches, busy people who work long hours and/or have families, and people who work non-traditional schedules, such as night shifts. Studies suggest that busy people and households where all adults are in the labor force are more difficult to reach for surveys (Abraham et al. 2006; Groves & Couper 1998). Reaching these populations by phone can be very difficult and sometimes impossible. Survey administrators may have to call in the evenings or on weekends, which adds to project costs and extends project timelines. When email addresses are available, an option to

complete the survey on-line could allow these interviewees to respond to a survey on their own schedule—late at night or early in the morning, when the telephone interviewer is unavailable.

Response Rate

The response rate for a survey is calculated as the ratio of the number of sample units that cooperate in a survey to the total number eligible to complete the survey (American Association for Public Opinion). Researchers place a great deal of emphasis on achieving a sufficient response rate, as response rates are an indicator of data quality and representativeness. According to Weisberg, Krosnick and Bowen (1989), it was not uncommon for survey researchers to obtain response rates of 90% in the 1950s. Now, with increased marketing, caller identification technology, and respondent fatigue with telephone survey efforts, people are less willing to respond to phone calls from people they don't know and response rates are typically much lower: usually under 50% and sometimes closer to 10%. A higher response rate improves the ability to infer findings about a population because a higher response rate should compensate for nonresponse bias. Nonresponse bias occurs when the answers given by respondents are different from the answers that would have been given by those people who chose not to respond. Groves et al. (2006) conducted a review of recent literature and concluded that nonresponse rates do not *necessarily* alter survey estimates, suggesting that in some cases, there is no nonresponse bias, even if there is low response.

In telephone surveying, a quota, or minimum number of completes, is established based on the desired statistical confidence and precision of the results. Once that number is achieved, the survey effort stops because of the additional cost for each completed survey beyond the minimum requirement. Email surveys, on the other hand, allow a research team to distribute access to the survey to an entire population of potential respondents, with no additional time or effort for any contact above the minimum number of required completes. This increased distribution allows more people a chance to participate and can increase the total amount of responses. Although this potentially increases the amount of total completes, it does not address the same voluntary response bias experienced with phone surveys—respondents still choose to participate.

Level of Effort

Because phone surveys require an interviewer, they are more expensive and therefore are directed at a sample drawn from the population. If response rates are low, researchers may elevate the level of effort to collect more data points by calling each sample point multiple times—often calling at different times—expanding the sample drawn from the population, or offering an incentive for participation. Because these additional data points are collected under different circumstances, they may vary from those brought into the sample under the original conditions (Peytchev, Baxter & Carley-Baxter 2009) and thus mitigate potential sources of bias. Email surveys can be easily distributed to an entire population. Email surveys of a known population present each person in the population a relatively equal opportunity to respond. Each potential participant receives the same number of reminders about the survey and the same set of information. Email surveys provide a relatively level playing field, since each respondent is expected to have access to the Internet and the computer skills required to complete an email survey, though those that do not respond may reflect a less computer-literate group. A detailed examination of the cost of the electronic survey method compared to other survey methods found that the typical cost per 10,000 respondents is $0.65 for a web survey and $1.64 for an email survey—with both methods costing less than a surface mail survey (Watt 1999).

Self-Selection Bias

Self-selection bias occurs when individuals self-select themselves into a group. Both phone and email surveys can suffer from self-selection bias if the respondents who choose to respond differ from the overall population by some relevant trait. If characteristics of a population are not known, a random sampling method is used to attempt to collect a representative sample. Email surveys can also be sent to a random sample or to an entire population. Since the entire population approach adds no additional cost or time to the survey effort, it seems logically preferable. However, surveying an entire population is not a method of probability sampling and therefore there is a risk of self-selection bias. This is of more concern when there is unidentified heterogeneity: that is, differences within the population that are unknown or unexpected. If the survey population is homogenous or similar enough, self-selection could be relatively harmless.

Social Desirability

The influence of social desirability is another issue present in both phone and email survey research. Social desirability bias refers to the phenomenon of people giving responses they expect to be perceived positively by others. Phone surveys present at situation where the respondent may want to respond how he/she thinks the interviewer wants them to respond. Theoretically, email surveys remove the oversight of the interviewer, allowing respondents to answer without fear of judgment. However, a recent study comparing four data collection methods (face-to-face interviews, telephone interviews, mail surveys, and online surveys) found no evidence of social desirability bias across the four methods (Rossiter 2009).

Population vs. Pre-selected List of Participants

A key difference between the majority of studies comparing email and phone surveys and our research is that our research typically uses lists of known program participants, not a random sample from the general population. This list provides complete contact information, including phone and email listings. Many previous studies compare results achieved through attempts to sample an entire population of respondents. Our studies use lists of participants in a program, meaning all potential respondents have program participation in common. Many of the issues identified in previous research do not apply to research done with pre-screened lists of known participants.

Our Investigation: What Do We See?

In order to identify potential mode effects, we analyzed two surveys, both of which employed a mixed-mode technique using telephone and email for data collection. We considered the results in light of the literature and assessed whether or not there are signs that the data collected have any patterns suggesting mode bias. Although these studies were not particularly designed for comparative study, these empirical examples provide some important insights.

Exploratory Case #1: NYSERDA Environmental Monitoring, Evaluation, and Protection (EMEP) Constituent Survey

We conducted this survey as part of the process evaluation of New York State Energy Research and Development Authority's (NYSERDA) EMEP program. EMEP maintains a general mailing list for thousands of science, policy, and academic professionals—all of whom receive periodic electronic information about program-related activities. We selected the population of this survey from the program's

general mailing list. All members of the list had been contacted by NYSERDA previously. We surveyed this population to assess: the type of organization; their professional role; their familiarity with and use of EMEP products and services; and issues regarding information content or distribution.

Exploratory Case #2: Clean Energy Works Portland (CEWP), Applicant Survey

Clean Energy Works Portland is a pilot project to provide homeowners access to low-cost financing for comprehensive energy efficiency home improvements. Participants in the program received an audit, assistance from an Energy Advocate, and guidance for retrofit and weatherization projects. We conducted this survey to learn about applicants who had applied to the program, but did not participate, sometimes because they had been screened out.

Survey Methods

For both surveys, we randomly drew a minimum sample size necessary to complete the required number of interviews from a sampling frame and then randomly assigned each sample group to a phone or email survey. To minimize nonresponse bias to the email survey, we sent initial invitations and followed up with four reminder emails for those respondents who had not already completed the survey. To minimize nonresponse to the phone survey, we made up to six call attempts to each respondent in the phone sample. These surveys had almost identical questions between the two modes. The response rates ranged between 30% to 33% (Table 1).

Table 1. Survey Completions and Response Rate

Sample	**NYSERDA EMEP**		**Clean Energy Works**	
	Phone	**Email**	**Phone**	**Email**
Total population	1,522		1,100	
Total sample	250	500	150	150
Completed surveys	76	164	45	49
Response rate	30%	33%	33%	30%

Note: We employed the AAPOR's response rate calculation method (RR1=minimum response rate) defined in Standard Definitions (AAPOR 2011).

Analysis and Results

We compared all responses provided by phone and email respondents of both survey efforts to test for measurement bias due to mode difference. We employed Chi Square analysis for categorical or nominal variables, T test for ratio variables, and Mann-Whitney U nonparametric test for ordinal scale variables. NYSERDA EMEP data were primarily nominal data points about respondents' organization and their past activities, whereas CEWP data consist of many rating scale batteries that attempted to measure respondents' opinions in addition to nominal data points.

The statistical tests generally found no response patterns unique to each survey mode. We concluded that the data are comparable, whether phone or email survey; as a result, we merged the data and reported combined results.

Although the overall response patterns were not compromised by using two different modes, we found several interesting differences between the two modes. None of them suggests which mode is superior, yet they provide some insight into question design decisions, mixed-mode approach, and other considerations when online survey is used to gain improved reliability.

Signs of "Satisficing"

Although the available data do not allow assessing hypothesis of specific response bias discussed in the literature (such as social desirability), we found some indications of *satisficing*. Satisficing refers to survey takers' mental shortcut during survey participation. Though many respondents are well intentioned when they start a survey, commitment to the task begins to wane as fatigue or disinterest increases. This lack of commitment results in respondents taking mental shortcuts to get through each question. Such reduced mental effort is called *satisficing* and can result in weakened data quality (Krosnick 1991). We observed satisficing effects in questions that asked the respondent to rate a series of items in scale batteries and multiple-response questions.

The CEWP Applicant Survey questionnaire included several groups of questions that employed rating scales. For each group of questions, we examined the proportion of respondents who selected the same rating scale for all the items within each battery. For example, if someone responded "5" for every question in a set asking the respondent to rate a concept on a one-to-five scale, we interpreted this as a sign of satisficing. Table 2 shows the percent of respondents who offered the same number rating for each question in a set in each survey mode, the differences between the two modes, and results of significance tests.

Table 2. Non-Differentiation in Rating Scale Batteries by Mode, CEWP Study

Question Batteries	**Phone Survey (n=45)**	**Email Survey (n=49)**	**Difference Between Modes**	**X^2, Significance**
Q8 (Six-question battery: Importance of each of the 6 factors to participation decision)	20%	27%	7%	NS
Q11 (Four-question battery: Importance of each of the 4 program features)	9%	22%	13%	NS
Q13 (Seven-question battery: Importance of each of the 7 financial benefits from program participation)	13%	33%	20%	$p < 0.05$

Although some respondents in both modes chose the same rating in each battery, in every case, the email survey respondents did this more often. The email survey results not only had a higher proportion of non-differentiation respondents for ratings of multiple items on the same response scale, but the difference between the two modes increased as the respondents proceeded further into the survey. In the last battery, the difference is statistically significant (X^2=4.88, p<0.05). This may indicate that, as the survey progressed, email survey respondents became increasingly inattentive when compared with phone survey respondents.

Multiple-Response Question: Interviewer's Effect and Special Question Design Consideration

The second type of question in which differences emerged was for multiple-response questions. EMEP's Constituents Survey included a few multiple-response type questions that allow for cross-mode

comparison. In the phone surveys, the question was asked in an open-ended format and interviewers recorded the responses using pre-coded response categories. The email survey presented the set of categories used as pre-codes for the phone survey and the respondents were asked to check all categories that apply. Table 3 shows mean number of response categories recorded by respondents in each survey mode, as well as significance test results. In both multiple-response questions, the phone survey resulted in significantly more mentions or checked items than the email survey ($T=10.9_{(q3)}$ $T=8.2_{(q10)}$, $p<0.001$).

Table 3. Mode Effects in Multiple-Response Questions, EMEP Study

Question Batteries	**Mean Number of Responses**		**T Test, Significance**
	Phone Survey (n=76)	**Email Survey (n=169)**	
Q3 (Scientific areas of interest, 9 items including "other" category)	5.0	2.4	$p < 0.0001$
Q10 (Dissemination methods of scientific findings, 8 items including "other" category)	4.5	2.8	$p < 0.0001$

These results have two implications about rating scales in the two survey modes. First, email survey requires special question design consideration that discourages satisficing behavior. The same multiple-response question could be presented with a forced-choice question format that uses *yes/no* options. An experimental study that compared results using multiple-response (check-all) and forced-choice (yes/no) formats in web survey not only found forced-choice format results in more number of items selection, but it entices deeper cognitive process and reduces satisficing behavior (Smyth et al. 2006).

Second, phone interviewers might be playing an active role in creating different results. Since the phone survey required the phone interviewer to code the responses from an open-ended question without reading each option, there is a possibility that the phone respondents would have chosen a different or more limited number of responses if allowed to consider all of the options to which their email counterparts had access.

Signs of Nonresponse Bias

It is important to be able to distinguish between mode effects and differences in the eventual sample composition due to nonresponse. Nonresponse bias occurs if the answers of respondents differ from the potential answers of those who did not respond to the survey, due to differences in the composition of the respondents. In the CEWP Applicant Survey, the data allowed some demographic comparisons between the two survey modes and the population. The original list included applicants' primary phone numbers, as well as email addresses for all the applicants, from which we randomly drew the same number of samples for each mode. The phone survey was conducted regardless of whether the number was a landline or cell phone.

Table 4 shows the distribution of annual household income, reported by survey mode, compared with the population's distribution. The results indicate that the distributions of household income in both modes are significantly different from that of the population ($p<0.001$)—between the two survey modes, $X^2=9.9$, $p<0.001$—a much higher percent of applicants in a lower income bracket responded by phone (45%) compared to those who responded by email (15%), and vice-versa for middle to higher income households. In fact, the distribution of combined approach of the two modes was the closest to that of the population. It is likely that nonresponse in either mode systematically altered the distribution of the entire sampling frame.

Table 4. Nonresponse Bias by Household Income CEWP Study

Annual Household Income	Phone Survey (n=45)	Email Survey (n=41)	Total (n=86)	Population (N=1,100)
Less than $50,000	45%	15%	30%	30%
$50,000 to less than $110,000	33%	61%	47%	54%
More than $110,000	22%	24%	23%	16%
Total	100%	100%	100%	100%

Note: Eight email survey respondents refused to provide responses.

Conclusions and Recommendations: What Have We Learned?

This paper explored current knowledge of methodological advantages and disadvantages of web surveys and looked for specific areas of concern that survey researchers need to consider when choosing between survey modes. We then analyzed data from prior projects and retrospectively compared data quality between two phone and email surveys. We observed that, when compared with phone survey responses, the email survey respondents exhibited stronger satisficing behaviors. They appeared more likely to select the same scale in rating scale batteries and were more likely to select fewer items in multiple-response questions, especially as they progressed through the survey and fatigue increased. In one case, this evidence of satisficing could be due to question design choices: telephone interviewers asked one question as an open-ended question and coded responses, while email respondents were given numerous options and asked to choose all that applied. In another case, however, the question designs were nearly identical—providing stronger evidence of satisficing on the part of the email respondents. Limiting the number of items in rating scale batteries for email surveys and placing such questions at an earlier part of surveys could mitigate this effect. Similarly, using *yes/no* questions for multiple-response could force engagement with the items on email surveys and reduce satisficing behaviors.

We also examined evidence of nonresponse bias in reported household income between the two modes. Both web and phone survey respondents differed significantly from the population in terms of household income, which suggests that one cannot rely on one mode or the other without recognizing the potential for significant nonresponse bias. Our finding is based on one study and will need to be confirmed before we can say there is conclusive evidence of this phenomenon. However, our results suggest the use of mixed-modes could result in a more representative sample than employing only one mode.

Despite the limitations of comparability inherent in the two studies and limited data points available to test a wide range of mode effect hypotheses, this study suggests that there may be differences in data quality between an email and phone survey, and that researchers need to be aware of risks when using email surveys or mix-mode designs. There are, however, steps researchers can take to reduce these issues. These steps are consistent with the tenets of quality survey research and include:

- **Access to a quality contact list for the population of interest**: Phone and email contact information should be equally available, the population accessible, and no major variables are expected to drive differences in response
- **Random sampling**: Selecting a simple random sample from the known population is best for controlling for unknown bias—even when the additional incremental cost to complete

an email survey is zero. Unless the objective is simply to obtain as many data points as possible, survey researchers should continue to select a random sample for email survey efforts.

- **Effective question design**: Question design is always an important component of survey research; with email surveys it is paramount. There is no telephone interviewer to clarify terms or guide a confused respondent through a poorly written question. There is no friendly rapport to keep the email respondent on topic or engaged in the survey experience. The questions must be clear and concise and the survey must be attractive and engaging.
- **At least three reminders for email**: Our approach was to send three reminders for the email survey and six follow-up calls for the phone survey. The comparable response rates demonstrate that these are comparable approaches.

Considering the increasing costs and challenges in access for phone surveys and expanding digital access, electronic survey methods are likely to be an increasingly attractive strategy for survey researchers. In light of this, evaluators should conduct further investigation, documentation, and refinement of methodological standards for email surveys. Understanding the lessons being learned and operating with shared survey standards will allow survey researchers to produce more reliable data collected through web surveys.

References

Abraham, K., A. Maitland, and S. Bianchi. 2006. "Nonresponse in the American Time Use Survey: Who Is Missing from the Data and How Much Does It Matter?." *Public Opinion Quarterly* (5): 676-703.

American Association for Public Opinion Research. 2011. "Standard Definitions: Final dispositions of Case Codes and outcome Rates for Surveys, 7th edition, revised 2011." http://www.aapor.org/AM/Template.cfm?Section=Standard_Definitions2&Template=/CM/ContentDisplay.cfm&ContentID=3156

Fricker, S., M. Galesic, R. Tourangeau, and T. Yan. 2005. "An Experimental Comparison of Web and Phone Surveys." *Public Opinion Quarterly* 69 (3): 370-392.

Groves, R. 2006. "Nonresponse Rates and Nonresponse bias in Household Surveys". *Public Opinion Quarterly* 70 (5): 646-675.

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

Groves, R., M. Couper, S. Presser, E. Singer, R. Tourangeau, G. Acosta, and L. Nelson. 2006. "Experiments in Producing Nonresponse Bias." *Public Opinion Quarterly* 70 (5): 720-736.

Groves, R. and R. Kahn. 1979. Surveys by Telephone: A National Comparison with Personal Interviews. New York: Academic Press.

Holbrook, A., M. Green, and J. Krosnick. 2003. "Telephone versus Face-To-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias." *Public Opinion Quarterly* 67 (1): 79-125.

Holbrook, A., M. Green, J. Krosnick, and A. Pfent. 2007. "The Causes and Consequences of Response Rates in Surveys by the News Media and Government Contractor Survey Research Firms". *In Telephone Survey Methodology*. 23:499–528. Hoboken, N.J.: John Wiley & Sons, Inc.

Internet World Stats. 2010. "World Internet Usage and Population Statistics." http://www.Internetworldstats.com/stats.htm. Bogota, Colombia: Miniwatts Marketing Group.

Krosnick, J. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5: 213-36.

Medin, C., S. Roy, and T. Chai. 1999. "World Wide Web versus Mail Surveys: A Comparison and Report. Paper presentation at ANZMAC99 Conference, Marketing in the Third Millennium, Sydney, Australia, November 28 – December 1. http://smib.vuw.ac.nz:8081/www/ANZMAC1999/Site/M/Medlin.pdf.

Nie, N. and L. Erbring. 2000. *Internet and Society: A Preliminary Report.* Palo Alto, Calif.: Stanford Institute for the Quantitative Study of Society Press.

Nie, N., D. Hillygus, and L. Erbring. 2002. "Internet Use, Interpersonal Relations and Sociability." In: Wellman, B. and C. Haythornthwaite (eds.), *The Internet In Everyday Life,* 215-243. Malden, Mass.: Blackwell Publishers, Ltd.

Nielsen Wire. "Call my Cell: Wireless Substitution in the United States." *Nielsen Wire.* New York, N.Y.: The Nielsen Company. Retrieved March 7, 2011 [No longer available].

Peytchev, A., R. Baxter, and L. Carley-Baxter. 2009 *"Not All Survey Effort is Equal: Reduction of Nonresponse Bias and Nonresponse Error." Public Opinion Quarterly* 73 (4): 785-806.

Rossiter, J. 2009. *A Comparison of Social Desirability Bias among Four Widely Used Methods of Data Collection as Measured by the Impression Management Subscale of the Balance Inventory of Desirable Responding PHD.* Ann Arbor, Mich.: ProQuest for Kent State University, College of Education, Health, and Human Services / Department of Educational Foundations and Special Services.

Smyth, J., D. Dillman, L. Christian, and M. McBride. 2009. "Open-ended Questions in Web Surveys: Can Increasing the Size of Answer Boxes and Providing Extra Verbal Instructions Improve Response Quality?" *Public Opinion Quarterly,* 73 (2): 325-337.

Smyth, J., D. Dillman, L. Christian, and M. Stern. 2006. "Comparing Check-All and Forced-Choice Question Formats in Web Survey." *Public Opinion Quarterly*, 70 (1):66-77.

Tannenbaum, B. and S. Feldman. 2001. "Energy Star Awareness as a Function of Survey Method." *Proceedings of the 2001 International Energy Program Evaluation Conference*. Salt Lake City: Utah: International Energy Program Evaluation Conference. August.

Watt, J. 1999. "Internet Systems for Evaluation Research." In G. Gay and T.L.Bennington (Eds.), *Information Technologies in Evaluation: Social, Moral, Epistemological, and Practical Implications*. San Francisco, Calif.: Jossey-Bass.

Weissberg, H., J. Krosnick, and B. Bowen. 1989. *An Introduction to Survey Research and Data Analysis*. Glenview, Ill.: Scott Foresman.