

This section includes shorter papers (e.g., 10-15 double-spaced manuscript pages or less) describing methods and techniques that can improve evaluation practice. Method notes may include reports of new evaluation tools, products, and/or services that are useful for practicing evaluators. Alternatively, they may describe new uses of existing tools. Also appropriate for this section are user-friendly guidelines for the proper use of conventional tools and methods, particularly for those that are commonly misused in practice.

Is a Web Survey as Effective as a Mail Survey? A Field Experiment Among Computer Users

Nancy Ellen Kiernan
Penn State University

Michaela Kiernan
Stanford University

Mary Ann Oyler
Pennsylvania State Cooperative Extension in Adams County

Carolyn Gilles
Pennsylvania State Cooperative Extension in Crawford County

Abstract: With the exponential increase in Web access, program evaluators need to understand the methodological benefits and barriers of using the Web to collect survey data from program participants. In this experimental study, the authors examined whether a Web survey can be as effective as the more established mail survey on three measures of survey effectiveness: response rate, question completion rate, and the lack of evaluative bias. Community- and university-based educators ($n = 274$) attending a 2-day program were randomly assigned to receive a Web or mail survey evaluating the program. Among those participants successfully solicited by e-mail, Web survey participants were more likely to respond (95%) than mail survey participants (79%). Web survey participants completed similarly high numbers of quantitative questions as mail survey participants, provided longer and more substantive responses to qualitative questions, and did not demonstrate evidence of evaluative bias. These results suggest that program evaluators could expand their use of Web surveys among computer users.

Nancy Ellen Kiernan, Program Evaluator, Pennsylvania State University, 438 Ag Admin Bldg., University Park, PA 16802; phone: (814) 863-7621; e-mail: nekiernan@psu.edu.

American Journal of Evaluation, Vol. 26 No. 2, June 2005 245-252
DOI: 10.1177/1098214005275826
© 2005 American Evaluation Association

Keywords: *Web survey; mail survey; experiment; quantitative questions; qualitative questions; response rate; question completion; program evaluation*

With the exponential increase in access to the Web, program evaluators need to understand the methodological benefits and barriers of using the Web to collect survey data from program participants (Morrel-Samuels, 2003; Watt, 1999). Until recently (Cobanoglu, Warde, & Moreo, 2001; Couper, Traugott, & Lamias, 2001; McCabe, Boyd, Couper, Crawford, & d'Arcy, 2002; Vehovar, Batagelj, Manfreda, & Zaletel, 2002), few systematic studies have investigated Web-based surveys, a research gap identified at the 1999 International Conference on Survey Nonresponse. This experimental study emerged from recognition of that gap.

To have confidence in survey data collected on the Web and in the ability to infer the results to a target audience, it is important to investigate whether Web surveys can be as effective as more established survey methods, such as mail surveys. Therefore, in the present study, we randomly assigned program participants to receive either a Web survey or a mail survey and compared the surveys on three key measures of survey effectiveness: response rate, question completion, and the lack of evaluative bias. A mail survey was chosen as the comparison survey method because it achieves high response rates and collects sensitive data without social desirability bias (Booth-Kewley, Rosenfeld, & Edwards, 1993; Goodfellow, Kiernan, Ahern, & Smyer, 1988; Rossi, Freeman, & Lipsey, 1999). Below, we outline a priori expectations about the two methods on the three measures of effectiveness on the basis of past research and our own pilot study.

Survey Response Rate

Web surveys typically have had poor response rates, suggesting resistance to this method (Archer, 2003; Cobanoglu et al., 2001; Dillman, 2000; Truell, Bartlett, & Alexander, 2002; Vehovar et al., 2002). With few exceptions (McCabe et al., 2002; Parker, 1992), response rates for Web surveys parallel the poor response rates for e-mail surveys, that is, surveys transmitted electronically by modem or direct wire to a computer, and they are typically self-selective (Dillman, 2000; Vehovar et al., 2002). To identify factors that may affect Web survey response rates, we conducted a pilot study (Kiernan, 1999) a year before the present study. Among a random sample of 101 community- and university-based educators attending a 2-day professional development program, 60% of pilot respondents indicated that they were moderately or very likely to respond to a Web survey. In the qualitative data, pilot respondents revealed the benefits of Web surveys. Such surveys were perceived as quicker and easier to fill out, easily accessed via the Web, and less likely to be misplaced. However, 40% of pilot respondents said that they were not very likely or only somewhat likely to respond to a Web survey. The qualitative pilot data revealed two dimensions of resistance to Web survey. First, these respondents perceived several barriers for responding to Web surveys, which they saw as taking longer, not being entirely and easily visible (unlike mail surveys), requiring reminders (unlike the presence of mail surveys), difficult to access with limited computer memory, and unfamiliar (e.g., to a "cybernetic dinosaur like myself"). Second, pilot respondents worried about confidentiality and feared that Web surveys might be reviewed on organizations' servers. Although mail surveys have successfully conveyed confidentiality to employees within organizations (Booth-Kewley et al., 1993; Hosseini & Armacost, 1993), concerns about confidentiality have led to lower response for other survey methods (Singer, Mathiowetz, & Couper, 1993). Thus, given past research and our pilot study, our a priori expectation for the present experimental study was that Web survey participants would have a lower response rate than mail survey participants.

Question Completion

Although quantitative and qualitative questions tend to evoke different completion rates in mail surveys (Dillman, 2000), little is known for Web surveys. Previous research comparing the quantitative question completion of e-mail survey participants with that of mail survey participants has yielded mixed results (Bachmann, Elfrink, & Vazzana, 1996; Couper, Blair, & Triplett, 1999; Mehta & Sivadas, 1995; Schaefer & Dillman, 1998). Anticipating no barriers to completion in the present study, we expected Web survey participants to be as likely as mail survey participants to complete quantitative questions. Previous research has found that e-mail survey participants were more likely than mail survey participants to complete qualitative questions and to write longer answers, with researchers speculating that typing answers takes less effort than handwriting them (Bachmann et al., 1996; Schaefer & Dillman, 1998). Thus, we expected Web survey participants to be more likely than mail survey participants to complete qualitative questions.

Views in Web Surveys

Limited research to date suggests that Web survey participants would evaluate programs as positively (or negatively) as mail survey participants (Wright, Aquilino, & Supple, 1998). However, our pilot study identified a factor that had the potential to bias the views of Web survey participants. If Web survey participants were worried about confidentiality, either because of their Internet connections or their employers' access to servers, they might present themselves more favorably and evaluate programs more positively. Thus, we expected Web survey participants to evaluate the program more positively than mail survey participants.

Method

Participants

In the present experimental study, participants ($n = 274$) consisted of community- and university-based educators who design and run educational programs in community settings and who were attending an annual professional development program. Each year, the 2-day program of plenary sessions, workshops, posters, and discussions focuses on a different topic, such as stakeholder involvement, marketing strategies, and pedagogical techniques. In the year the present experimental study was conducted, the topic focused on expanding educators' use of communication tools, including video conferencing, PowerPoint presentations, digital TV, and CD-ROMs.

The population consisted of 447 educators when registration closed. We randomly sampled 274 participants from the registration list and randomly assigned half to each survey method ($n = 137$). The sample size, chosen to minimize survey burden for the population, could detect a 15% difference in response rate between the two methods, with 80% statistical power using a two-tailed chi-square test at $p < .05$ (Hulley et al., 2001).

Survey Solicitation

To increase participation, on the basis of prior published research (Couper et al., 2001; Dillman, 2000; Mehta & Sivadas, 1995) and our pilot study (Kiernan, 1999), the solicitation

strategy used four contacts. A week before the program, registrants received a personalized e-mail request indicating that they had been randomly selected and asking them to evaluate the program. Participants randomly assigned to the Web survey learned that they would receive the Web site address and password at the program and had 15 days to post on the Web. Participants assigned to the mail survey learned that they would receive a postage-paid survey at the program and had 15 days to postmark the survey. Both groups were assured of confidentiality and voluntary participation. Within 10 days of the program's end, both groups received two nonpersonalized e-mail reminders. All registrants had Web access from office or laptop computers.

Survey Design

To establish consistency between the survey methods in this experimental study, we designed the Web and mail surveys with many identical features. First, the Web and mail surveys had identical wordings and formats (e.g., vertical answer categories with no dropdown boxes on the Web survey and with the place to indicate the chosen answer in front of each answer). Second, all questions included directions for greater survey response (Couper et al., 2001; Dillman, 2000): "Click button on your answer below" in the Web survey or "Circle answer below" in the mail survey. Third, each survey opened with a thank-you for participation and the due date. The entire first question in the Web survey appeared on the first screen (Dillman, 2000). To reduce the Web survey barriers mentioned in our pilot study, Web survey participants could scroll through for an overview or change answers, and they had quick access given a simple graphic layout and restrained use of color that minimized demand for computer memory (Nielsen, 2000).

For both surveys, there was a demographic question about educator type (community based, university based, or other). There were 14 quantitative questions (2 on knowledge about program content, 7 about attitudes toward program components, 3 about attitudes toward program content, 1 on previous behavior, and 1 on intentions about future behavior) and 2 qualitative questions (1 on intentions and 1 on barriers, both about future behavior).

Analyses

The SPSS chi-square goodness-of-fit test for comparing proportions, *t* tests for comparing means, and the Kruskal-Wallis test for comparing medians for independent samples were used to test differences between Web and mail surveys ($p < .05$). A Bonferroni correction controlled for the number of analyses for more than one quantitative question of a particular type (e.g., seven attitude questions about program components), specifically, $.05/7$ ($p < .007$).

Results

Survey Solicitation Rate

There was no difference between the two survey methods in the percentage of registrants successfully solicited and given surveys, $\chi^2(1, N = 274) = 0.5, p > .05$. Of the 137 registrants asked to participate in a Web survey, 73.7% ($n = 101$) were successfully solicited and given the Web survey Web site. Of the 137 registrants asked to participate in a mail survey, 77.4% ($n = 106$) were successfully solicited and given the mail survey. Of the 67 participants not successfully

solicited across both methods, 56.7% ($n = 38$) did not reply to the e-mail request, 9.0% ($n = 6$) replied and did not want to participate, 19.4% ($n = 13$) replied that they would not be at the program long enough to justify participation, and 14.9% ($n = 10$) canceled registration.

Survey Response Rate

In contrast to our a priori expectation outlined in the introduction, Web survey participants did not have a lower survey response rate than mail survey participants (70.1%, $n = 96$, and 61.3%, $n = 84$, respectively), $\chi^2(1, N = 274) = 2.3, p > .05$. Indeed, among participants successfully solicited and given surveys, Web survey participants had a higher response rate than mail survey participants (95.0%, $n = 96$, and 79.2%, $n = 84$, respectively), $\chi^2(1, N = 207) = 11.5, p < .01$. On the basis of time-stamped data, 13% ($n = 12$) of Web survey participants posted surveys after traditional office hours (between 6 p.m. and 7 a.m. or on weekends).

Question Completion

Among survey responders ($n = 180$), Web survey participants were not more likely to be different types of educators than mail survey participants, $\chi^2(2, N = 178) = 0.4, p > .05$. Across both methods, more than 60% were community-based educators. Consistent with our a priori expectation, among survey responders ($n = 180$), Web survey participants were not more likely to complete the 14 quantitative questions than mail survey participants (all p values $> .05$). Across both methods, more than 92% completed the quantitative questions.

Although our a priori expectation was that Web survey participants would be more likely than mail survey participants to complete the qualitative questions, the results were not consistent across the two questions. Among survey responders, Web survey participants were less likely to complete the first qualitative question than mail survey participants (60.4%, $n = 58$, and 79.8%, $n = 67$, respectively), $\chi^2(1, N = 180) = 7.8, p < .01$. However, Web survey participants were not more likely to complete the second question than mail survey participants (88.5%, $n = 85$, and 91.7%, $n = 77$, respectively), $\chi^2(1, N = 180) = 0.5, p > .05$. Consistent with our a priori expectation, Web survey participants were more likely to have longer answers than mail survey participants for both qualitative questions. For the first question, Web survey participants typed an average of 17.1 words, whereas mail survey participants wrote an average of 11.0 words, $t(123) = 2.9, p < .01$. For the second question, Web survey participants typed an average of 21.7 words, whereas mail survey participants wrote an average of 9.6 words, $t(160) = 6.2, p < .01$.

In a post hoc analysis, we examined whether the longer answers of the Web survey participants resulted from a tendency to type complete sentences or from an increase in useful information. In general, only about a third or fewer of the participants across both methods had any complete sentences in their answers for either question. For the first question, Web survey participants were not more likely to have any complete sentences than mail survey participants (22.4%, $n = 13$, and 11.9%, $n = 8$, respectively), $\chi^2(1, N = 125) = 2.5, p > .05$. For the second question, Web survey participants were more likely to have any complete sentences than mail survey participants (36.5%, $n = 31$, and 11.7%, $n = 9$, respectively), $\chi^2(1, N = 162) = 13.4, p < .01$. We then examined the number of independent ideas for the two qualitative questions. On the basis of two raters, one of whom was blind, the interrater reliability for the number of independent ideas was excellent for the two questions for both the Web and mail surveys ($r = .82$ to $.96, p$ values $< .0001$). For the first question, Web survey participants had as many independent ideas as mail survey participants (medians of 2.0 and 2.0, respectively), $\chi^2(1, N = 125) = 3.1, p >$

.05. For the second question, Web survey participants had a greater number of independent ideas than mail survey participants (medians of 2.0 and 3.0, respectively), $\chi^2(1, N = 162) = 29.7$, $p < .01$.

Views in the Survey

Contrary to our a priori expectation, among survey responders ($n = 180$), Web survey participants did not view the program more positively (or negatively) than mail survey participants for 13 of 14 quantitative knowledge, attitude, behavior and intention questions (p values $> .05$), except for an attitude question with only a moderate effect size, Cohen's $d = .44$.

Discussion

In this experimental study, we examined whether a Web survey was as effective as a mail survey among computer users on three key measures of survey effectiveness: response rate, question completion, and the lack of evaluative bias. Contrary to our expectation that the Web survey would have a lower response rate than the mail survey, we found that the Web survey was as effective (70%) as the mail survey (61%). Indeed, among those Web survey participants who were successfully solicited by e-mail, the response rate (95%) was significantly better than that of the mail survey (79%). The higher than expected response rate for the Web survey could be because Web survey participants may face fewer barriers than mail survey participants. For instance, a Web survey may require less effort because participants can simply click on a Web address in an e-mail to locate the Web survey and, after completing it, merely click to send it. In contrast, mail survey participants must physically retrieve a survey, complete it, and get it to a mailbox to be postmarked by a due date. Additionally, Web survey participants may have more opportunities and stimuli to respond than mail survey participants. For example, routine access to e-mail after office hours from a laptop or home computer can create opportunities to respond to a Web survey. Indeed, about 40% of Americans say that they log on after 5 p.m. (Horrigan & Rainie, 2002). Consistent with this speculation, we found that 13% of Web survey participants posted their surveys after 6 p.m. or on weekends on the basis of the date-and-time stamp automatically recorded when the Web survey was returned. This percentage is similar to the difference in response rates between the Web and mail surveys, (i.e., 16%); however, more research is needed on patterns and timing of Web survey response.

We also found that a Web survey appears to be as effective as a mail survey in the completion of quantitative questions that measure knowledge, attitudes, behaviors, and intentions. This finding is important in light of the frequent use by program evaluators of this wide array of quantitative questions (Rossi et al., 1999). However, completion rates of the qualitative questions were equivocal. One reason for the lower completion rate by Web survey participants on the first qualitative question may be that participants may not have initially understood that they had to switch from using a mouse to using a keyboard to type out answers (Dillman, 2000). However, confirming our a priori expectation based on related research on e-mail surveys (Dillman, Eltinge, Groves, & Little, 2002), we found that once Web survey participants engaged in the qualitative questions, their answers were longer than the mail survey participants for both questions. In fact, Web survey participants had more independent ideas in their answers than mail survey participants, a component in Web survey research that has not previously been studied. Evaluators could consider using Web surveys in search of extended qualitative answers to garner experiences that explain and validate program participants' attitudes, experiences important to the participants but unknown to the evaluators, and the language and

metaphors that participants use to describe what is important (Patton, 2002). Further research is needed in experimental studies using a larger array of qualitative questions to determine how evaluators can take advantage of the longer and potentially more substantive answers in qualitative questions in Web surveys.

The results of the present experimental study demonstrate that overall, no systematic differences emerged between the Web and mail surveys in how participants evaluated the program. Concern that Web survey participants may have about confidentiality in using computers on organizational servers, which might lead them to present themselves more favorably by evaluating a program more positively, did not appear to be a factor.

One limitation of this study is that the sample of computer users had e-mail and Web access and thus may be less representative of participants in other program evaluations. However, the use of e-mail and the Web continues to increase rapidly in the United States among lower educational and income groups, and e-mail and the Web are used as tools for work among a wide range of occupations (Horrigan & Rainie, 2002). For instance, producers on small eastern farms monitor regional insect counts on the Web to calibrate the timing and degree of pesticide use (Pest Watch, 2004). Another limitation of this study is that the sample consisted of educators who attended a professional development program and may be less representative of other educational program participants. However, many private industries, educational institutions, branches of the military, and government agencies invest heavily in educational programs for their staff members who have access to the Web and are using the Web to evaluate programs (Morrel-Samuels, 2003; The Rodale Institute Farm, 2004). Future research is needed to determine the influence of survey components on Web survey response, such as the combination of quantitative and qualitative questions; the number of questions; survey length and content (Crawford, Couper, & Lamias, 2001); the sensitivity of question content (McCabe et al., 2002); layout, such as drop-down boxes and skip questions (Couper, Tourangeau, Conrad, & Crawford, 2004); and the type of contact (Porter & Whitcomb, 2003). Future research needs to characterize high and low responders among program participants on variables such as gender, race, age, and education (Goodfellow et al., 1988; McCabe et al., 2002).

Despite the study limitations, this study has important strengths. It is one of the few studies to use an experimental framework randomly assigning program participants to a Web survey and a comparison survey method to evaluate survey effectiveness. This study also pitted the Web survey against a challenging comparison, the mail survey, which is an established method frequently used in organizational contexts and known for its effectiveness to evoke high response rates and reduce bias. Finally, this study analyzed the completion of both quantitative and qualitative questions, using a range of frequently used measures in program evaluation.

Given that a Web survey can achieve as effective a response rate as a mail survey; be as effective in the completion of quantitative questions; elicit longer, more substantive qualitative answers than a mail survey; and evoke the same evaluative views of a program among computer users, these results complement other documented benefits of the Web for program evaluators, including high data quality, less time, and low costs (Archer, 2003; Booth-Kewley et al., 1993; Couper et al., 2001; Dillman, 2000; Morrel-Samuels, 2003; Watt, 1999). In conclusion, Web surveys hold promise for program evaluators. We believe that this experimental study will provoke further innovative research into the benefits and appropriate contexts for using Web surveys.

References

- Archer, T. M. (2003). Web-based surveys. *Journal of Extension*, 41(4), 1-5.

- Bachmann, D., Elfrink, J., & Vazzana, G. (1996). Tracking the progress of e-mail vs. snail-mail. *Marketing Research*, 8(2), 30-35.
- Booth-Kewley, S., Rosenfeld, P., & Edwards, J. E. (1993). Computer-administered surveys in organizational settings: Alternatives, advantages, and applications. In P. Rosenfeld, J. E. Edwards, & M. D. Thomas (Eds.), *Improving organizational surveys: New directions, methods, and applications* (pp. 73-101). Newbury Park, CA: Sage.
- Cobanoglu, C., Warde, B., & Moreo, P. J. (2001). A comparison of mail, fax and Web-based survey methods. *International Journal of Market Research*, 43(4), 441-452.
- Couper, M. P., Blair, J., & Triplett, T. (1999). A comparison of mail and e-mail for a survey of employees in U. S. statistical agencies. *Journal of Official Statistics*, 15(1), 39-56.
- Couper, M. P., Tourangeau, R., Conrad, F., & Crawford, S. D. (2004). What they see is what we get: Response options for Web surveys. *Social Science Computer Review*, 22(1), 111-127.
- Couper, M. P., Traugott, M. W., & Lamias, M. J. (2001). Web survey design & administration. *Public Opinion Quarterly*, 65, 230-253.
- Crawford, S. D., Couper, M. P., & Lamias, J. J. (2001). Web surveys: Perceptions of burden. *Social Science Computer Review*, 19(2), 146-162.
- Dillman, D. A. (2000). *Mail and Internet surveys: The tailored design method* (2nd ed.). New York: John Wiley.
- Dillman, D. A., Eltinge, J. L., Groves, R. M., & Little, R. J. A. (2002). Survey nonresponse in design, data collection and analysis. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 3-26). New York: John Wiley.
- Goodfellow, M., Kiernan, N. E., Ahern, F., & Smyer, M. A. (1988). Response bias using two-stage data collection: A study of elderly participants in a program. *Evaluation Review*, 12(6), 638-654.
- Hosseini, J. C., & Armacost, R. L. (1993). Gathering sensitive data in organizations. In P. Rosenfeld, J. E. Edwards, & M. D. Thomas (Eds.), *Improving organizational surveys: New directions, methods, and application* (pp. 29-57). Newbury Park, CA: Sage.
- Horrigan, J. B., & Rainie, L. (2002). *Getting serious online*. Available at http://www.pewinternet.org/pdfs/PIP_Getting_Serious_Online3ng.pdf
- Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D., Hearst, N., & Newman, T. B. (2001). *Designing clinical research*. Philadelphia: Lippincott Williams & Wilkins.
- Kiernan, N. E. (1999). [Pilot study: Using the Web for program evaluation]. Unpublished data.
- McCabe, S. E., Boyd, C., Couper, M. P., Crawford, S., & d'Arcy, H. (2002). Mode effects for collecting alcohol & other drug use data: Web & U. S. mail. *Journal of Studies on Alcohol*, 63(6), 755-761.
- Mehta, R., & Sivasdas, E. (1995). Comparing response rates and response content in mail versus electronic mail surveys. *Journal of the Market Research Society*, 37, 429-439.
- Morrel-Samuels, P. (2003). Web surveys' hidden hazards: Companies replacing paper surveys with Web-based versions. *Harvard Business Review*, 81(7), 16-17.
- Nielsen, J. (2000). *Designing Web usability*. Indianapolis, IN: New Riders.
- Parker, L. (1992, July). Collecting data the e-mail way. *Training and Development*, pp. 52-54.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Pest Watch. (2004). Home page. Retrieved October 22, 2004, from <http://www.pestwatch.psu.edu>
- Porter S. R., & Whitcomb, J. E. (2003). The impact of contact type on Web survey response rates. *Public Opinion Quarterly*, 67(4), 579-588.
- The Rodale Institute Farm. (2004). *2004 field day and workshop survey*. Retrieved October 22, 2004, from <http://www.newfarm.org/surveys/ifafs2004a.htm>
- Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: A systematic approach* (6th ed.). Thousand Oaks, CA: Sage.
- Schaefer, D. R., & Dillman, D. A. (1998). Development of a standard e-mail methodology: Results of an experiment. *Public Opinion Quarterly*, 62, 378-397.
- Singer, E., Mathiowetz, N. A., & Couper, M. P. (1993). The impact of privacy and confidentiality concerns on survey participation. *Public Opinion Quarterly*, 57, 465-482.
- Truell, A. D., Bartlett, J. E., & Alexander, M. W. (2002). Response rate, speed, & completeness: A comparison of Internet-based and mail surveys. *Behavior Research Methods, Instruments, & Computers*, 34(1), 50-54.
- Vehovar, V., Batagelj, Z., Manfreda, K. L., & Zaletel, M. (2002). Nonresponse in Web surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 229-242). New York: John Wiley.
- Watt, J. H. (1999). Internet systems for evaluation research. In G. Gay & T. L. Bennington (Eds.), *Information technologies in evaluation: Social, moral, epistemological and practical implications* (pp. 23-43). San Francisco: Jossey-Bass.
- Wright, D. L., Aquilino, W. S., & Supple, A. J. (1998). A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol, and drug use. *Public Opinion Quarterly*, 62, 331-353.