Issues Surrounding the Use of the Internet for Data Collection

Jennifer Klein

Response Factors

Response Rates

Most research on Internet response rates has examined the rates of e-mail versus postal surveys, with a wide variation in response rates among e-mail surveys. This variation is not surprising given the lack of consistency in numerous variables, such as the variety of sample groups, research topics, and method reported in the various studies. Sheehan and McMillan (1999) noted that researchers planning e-mail surveys will have difficulties determining sample size because of the minimal information on which to base estimations of response rates. The majority of researchers who examined response rates of Internet surveys found no difference in response rates among respondents of e-mail versus postal surveys (Bachmann, Elfrink, & Vazzana, 1996; Comley, 1996; Dillman, 1999; Mehta & Sivadas, 1995; Rosenfeld, Booth-Kewley, & Edwards, 1993; Schaefer & Dillman, 1998). It is important to note that some of these studies are somewhat dated. It is possible that response rates may be different today.

A couple of researchers reported lower rates by e-mail than by postal mail. Schuld and Totten (1994) had a response rate of 19% for e-mail versus 56.5% for postal when surveying university faculty members on their perception of computer shareware copying. The researchers suggested that the topic may have reduced the respondents’ willingness to associate their e-mail address with their response. Tse et al. (1995) conducted a survey by e-mail and postal mail in the Chinese University of Hong Kong that examined respondents’ attitudes toward business ethics and corruption. The results of the two different methods were 6% for e-mail versus 27% for postal. The researchers explained the lower e-mail response rate in terms of fear of the new e-mail technology, respondent difficulty completing e-mail surveys, and traceability of the respondent for a potentially sensitive subject.

Very little has been published regarding response rates from Web-based surveys. Jones and Pitt (1999) included Web response rates in their research. They sampled 500 university staff members using e-mail, e-mail notification plus a Web survey, or a postal survey. Response rates were highest with the postal survey (72%) compared with both the e-mail (34%) and the Web (19%) surveys. In comparison, Comley (1996) surveyed 3,564 U.K. Internet magazine readers with all three methods and obtained results similar to one another for each method (i.e., postal, 15.4%; vs. e-mail, 13.5%; vs. Web, 17%). Additional research is required to examine this issue further.

Speed of Response Times

The Internet as a form of data collection is unparalleled in terms of speed for survey responses. Batagelj and Vehovar (1998) reported that a period of 1 month has become widely accepted to carry out Web surveys where the sample is not preselected. Other research findings support this suggestion of quick responses. Dillman (1999) reported that in 3 days, nearly 25% of the respondents had returned their surveys through the Internet. Comley (1996) reported that nearly 66% of the responses were received within 3 days of original contact compared with 11 days for postal responses. Mehta and Sivadas (1995) stated similar findings, with 50% arriving in the...
first 3 days compared with 3 weeks for postal.

Drop-Off Rates

Oppermann (1995) noted that fast drop-off response rates occurred after each e-mail reminder because respondents decided quickly whether to cooperate. Use of the delete key makes disposing of a survey request easier and quicker than any other survey method, and the decision not to respond can be made before the entire request appears on the screen. When mailboxes are full, especially after a few days of absence, users often delete everything of lesser interest to them. Oppermann suggested that as time progresses and the novelty of Internet surveys wears off, surveys may be treated just like traditional mail and receive decreased response rates.

Response Quality

Several researchers reported that respondents who completed Internet surveys provided more information than they did for postal surveys (Comley, 1996; Dillman, 1999, 2000; Mehta & Sivadas, 1995; Schaefer & Dillman, 1998). Schaefer and Dillman (1998) and Dillman (1999) noted that respondents gave more detailed responses to open-ended questions on the e-mail version. Schaefer and Dillman hypothesized that the lengthier e-mail responses could have been that moving visually through the screens required more effort than filling out a paper questionnaire. More concentration is required because respondents can view only a few questions at a time (fewer than in the paper version); therefore, each question may be less likely overlooked and more thoroughly completed.

Dillman, Tortora, Conradt, and Bowker (1998) examined the differences in responses of a “plain” questionnaire requiring 317 K of memory compared with a “fancy” questionnaire that required 959 K. The fancy version was constructed with HTML tables. Alternating pink and purple bands of background colors were used to align each question item with the response box for that question. In addition, a somewhat unusual method of formatting was used to place answer choices immediately next to the scroll bar. Labels were provided above the answer, using graphic images. Transmission times took 120 sec for the plain survey and 345 sec for the fancy one, meaning that respondents took a longer time downloading the bigger file. Response rates differed in that 93.1% of those who logged in to the plain version completed all of the survey, whereas 82.1% of those who logged in to the fancy version finished. The increased time and added aesthetics appeared to be more of a hindrance than an assistance. These results suggest that using a plain survey without color and HTML tables provides better results than a fancier version.

Advantages of Internet Surveys

In examining e-mail and Web-based survey methods, many mutual advantages are shared with phone and written surveys.

• Very useful among specific populations. Today, scientists, engineers, teachers, students, librarians, businesspersons, politicians, health professionals, and others are already connected to networks.
• Reduced time required for survey responses. Time to respond can be reduced from weeks to days or even hours (Dillman, 2000). As the pace of rehabilitation professions increases with technological advances, more and more management and research decisions must be made within shorter and shorter periods. Because of their long time delays, conventional paper and phone surveying techniques will become applicable to fewer and fewer projects.

• Easy and direct transfer of data analysis. Programs are available to eliminate manual data entry.
• Easy to perform follow-up communication (Batagelj & Vehovar, 1998). Respondents can be acknowledged seconds after they have completed a survey. Acknowledgments also can be tailored based on survey responses. For example, respondents who indicate that their overall satisfaction with a therapy program was excellent to good can be e-mailed thank-you messages; those who indicated that their satisfaction was fair or poor can be sent messages of apology with a promise for rapid, personal follow-up. Follow-up communication also can include the possibility of distributing the results of the survey.
• Reduced costs for deploying and tabulating responses. Many of the costs of conventional surveys are eliminated (i.e., postage and paper, telephone interviewer labor, phone charges, manual data entry).
• User convenience. Survey completion is done when suitable and is less time consuming for the respondent (e.g., middle of the night if they want vs. dinnertime calls for phone surveys). One does not need to take into account the different time zones when exchanging information with others.
• Wide geographic coverage.
• Relatively anonymous nature of the Internet allows the researcher and respondents to remain strangers (Thomas, Leeseberg, Lafreniere, & Dumula, 2000).
• No interviewer bias.
• Can adjust font and graphics for visual impairments.
• Flexibility in response medium (e-mail, fax, mailing).
• Already in written format. The administrator does not need to spend endless hours transcribing the material.
• Less likely to produce social desirability answers for questions about personal issues because there is no direct interaction with another person (Dillman, 2000).

Limitations of Internet Surveys

Although these benefits are impressive, the limitations of using the Internet as a tool for gathering data must also be addressed. Surveys on the Internet have developed in an extremely short period of time and are still in a phase of intensive development. The development of Internet surveys has presented measurement challenges not previously faced by researchers and clinicians, and little research has been able to provide solutions.

• Selection bias. Individuals with access to the Internet tend to be men, have a higher education and income level and are more likely younger than the general population, and tend to be in better general health (Bell & Kahn, 1996; Dillman, 2000). Fischbacher, Chappel, Edwards, and Summerton (2000) reviewed reports
on which the Internet had been used for data collection and found that 24 of the 43 reports (56%) mentioned that underrepresentativeness was a limitation of the survey; however, almost 60% generalized their findings to a source population.

- No universal e-mail address directory listing all known addresses. E-mail addresses have not been formulated in a standardized way that would enable researchers to access them randomly as is done with random digit dialing for telephone numbers.
- Users not very familiar with their computer system’s capabilities.
- No probing by researcher.
- Respondents traceable, which may reduce confidentiality and desire to cooperate.
- Respondents pay for Internet access and, thus, pay for filling out surveys.
- Technology difficulties, including nonstandardized program equipment, incompatibility between systems, and varying transmission capabilities.
- May be considered aggressive by respondents and not in harmony with the Internet culture or “netiquette.” Mehta and Sivadas (1995) noted that sending unsolicited e-mail surveys was unacceptable.
- Difficulty in including monetary or tangible incentives.
- Potential to respond to the survey several times. This limitation, however, can be discouraged through requiring respondents to generate an entirely new health and demographic profile for each submission, as suggested by Thomas et al. (2000). It is important to note that this issue is not specific to Internet surveys because multiple copies of paper surveys can be completed and returned by the same respondent.
- Persons with various disabilities often at a disadvantage with respect to acquiring new technology. Multiple reasons exist for this limitation, including, but not limited to, lower discretionary funds because of lower incomes and high costs of health care, decreased mobility to go into the community to purchase computer items, and decreased energy to discover the appropriate sources of technology.
- Client interaction less personal than with phone surveys.

**Threats: Legal and Ethical Issues**

The increasing use of the Internet for data collection is an area that has ethical and legal implications. The issue of respondent anonymity is important. When an individual replies to an e-mail survey, his or her address appears on the reply e-mail. One method to ensure respondent anonymity is to send the e-mail surveys to a neutral party who would erase the addresses and then forward them to the researcher. However, participants must trust that this step is done. Web surveys provide a manner in which respondents can complete surveys with increased anonymity. Although the risk of interception within cyberspace is uncommon, it cannot be eliminated. A copy of what was sent resides on an exchange server (not hard drive), and nothing ensures that the administrator of the Internet server or a computer hacker does not access the information. In addition, deleted information continues to exist in cyberspace. In fact, computer specialists state that messages believed to be no longer in existence make for some of the best evidence in criminal charges (Fiesta, 1996).

Researchers also need to be aware of the issue of wrongful disclosure of information to an individual who is unauthorized to view documents. Just as important private memos are locked in a desk or office, Internet data must be properly secured within the computer. If the information is being retained in an active computer file, the computer user must be careful to secure the document. One of the biggest dangers in this regard for health professionals is walking away from the computer without signing off, which breaches the security system and allows unauthorized access. In some facilities, this breach of security is considered the basis for disciplinary action against individuals who do not follow security policy and the individual who accesses the information without permission (Fiesta, 1996). In an attempt to prevent access to extremely sensitive information, experimentation is ongoing with biometric devices. These systems recognize specific human traits, such as voice, blood vessel patterns in the retina, hand shape, and fingerprints, and are expected to be in use later in the 21st century (Fiesta, 1996).

Jones and Pitt (1999) reported using an address list from the Web for sampling their respondents directly, without contacting the site’s Web master (a person who is responsible for the creation of the Web site). The use of addresses in this manner raises ethical issues about whether Web masters should act as gatekeepers in determining who has the right to use their address lists.

Researchers need to ensure that respondents provide informed consent and are aware of the risks and benefits of electronic records. Implied in this given consent are respondents’ rights to refuse having their information handled electronically and that in the future they may have their information withdrawn from the electronic environment. Currently, consent to completing an Internet survey is usually taken for granted, but this assumption may become less acceptable in the future. Soon, requirements may make it mandatory for respondents to fill out a consent form before completing an Internet survey. Two copies then will be sent automatically to the respondents e-mail address and to the researcher to ensure that all involved individuals have proof and confirmation of informed consent.

**Conclusion**

Use of the Web and e-mail surveys provide enormous opportunities but not without their challenges. According to all projections, these data collection methods will increase among the health professions because of the multitude of benefits, including international access to populations, convenience for the respondents and researcher, reduced costs, and increasing numbers of Internet users globally. The potential advancements in this realm are yet untapped. Development of more reliable and valid research designs and better standards will continue. These developments are viewed as crucial to ensuring that Internet data collection is an effective method of conducting high-quality research. ▲

**References**

snail-mail. *Marketing Research, 8*(2), 30–35.


