Online Data Collection: Strategies for Research

Darcy Haag Granello and Joe E. Wheaton

Online data collection, through e-mail and Web-based surveys, is becoming an increasingly popular research methodology. In this article, the authors outline the benefits and limitations of this type of data collection to help researchers determine whether their data could be collected online in a way that retains the integrity of the data. A detailed procedure, including strategies to manage limitations, is given for researchers wishing to conduct their own online surveys.

In an age of ever-advancing technology, Americans are becoming progressively more computer literate. More and more people have access to the Internet, and the Internet is fast becoming the communication method of choice for many Americans (Duffy, 2000). Researchers from many disciplines are starting to see the benefits of collecting data using the Internet, and increasingly, journals are publishing data that have been collected online (Schleyer & Forrest, 2000).

Despite this increased use of the Internet for data collection, there is little published research on the process of data collection online. That is, discipline-specific studies publish the *results* of their Web-based surveys in discipline-specific journals, but little information is available on the specifics of how Internet-based data collection can be accomplished. Thus, it is difficult for a researcher wishing to use this data collection method to find resources to use as guides.

A review of recent issues of the flagship journal of the American Counseling Association and the journals of its largest divisions and affiliates *(Journal of Counseling & Development, The Family Journal, Counselor Education and Supervision, Journal of Mental Health Counseling,* and *Professional School Counseling)* found that the publication of data collected online has not yet become standard in the field of counseling. In fact, since 1998, no articles have been published in any of these journals that explicitly state that data were collected from an e-mail-based or Web-based survey method. (One study by Niles, Akos, & Cutler in a 2001 issue of *Counselor Education and Supervision* used e-mail correspondence to conduct qualitative interviews with participants, although participants originally were contacted via postal mail.)

There is, however, an advantage to using Web-based surveys that make it a particularly appealing method for the counseling profession. The advantages of reduced time, low-ered cost, ease of data entry, flexibility in format, and ability to capture additional response-set information are universal to Internet-based data collection in all fields. However, in the counseling profession, the additional benefit of access to populations that have not entered the mental health system (e.g., potential clients with mental health problems who have not yet accessed the system) may make online data collection particularly advantageous for counseling research.

It is possible that online data collection is not used in the field of counseling because of the limitations of this type of data collection, such as difficulties in obtaining a representative sample, low response rates, and problems with technology. These limitations are valid and pertinent to researchers in every field and are discussed in detail later in this article. It also is possible that online data collection has not been used in the counseling literature because of a lack of models and training for this type of research methodology. Therefore, this article is intended to help provide a structure for researchers wishing to engage in this type of data collection.

Based on a review of the literature and the findings from a large-scale online study that we previously conducted (Wheaton & Granello, 2001), this article outlines the advantages and limitations of online data collection for the field of counseling and then describes a step-by-step how-to method for researchers wishing to conduct their own Web-based surveys. Practical suggestions and lessons learned are included for researchers wishing to engage in their own online data collection.

AN OVERVIEW OF ONLINE DATA COLLECTION

There are several methods for collecting data online. The two most common are e-mail surveys and Web-based sur-

Darcy Haag Granello *and **Joe E. Wheaton,** Counselor Education, Rehabilitation Services, and School Psychology, The Ohio State University. Correspondence concerning this article should be addressed to Darcy Haag Granello, School of Physical Activity and Educational Services, 356 Arps Hall, 1945 North High Street, Columbus, OH 43210 (e-mail: granello.1@osu.edu).*

veys. With e-mail surveys, the participant receives an e-mail with a survey embedded in it, clicks on the "reply" button, fills the survey out, and clicks on the "send" button. The researcher then transfers the raw data into a database. Web-based surveys, on the other hand, require the instrument to be available on a Web site, and individuals are solicited—either by traditional mail, e-mail, telephone, or through other Web sites—to participate in the survey. Participants are given access information to enter the survey Web site; they complete the form online and then click on a "submit" button when they have completed it. Both of these methods have advantages and disadvantages.

Advantages of Online Data Collection

Online surveys have several important advantages over paper-and-pencil surveys that make them particularly attractive to researchers. These include reduced response time, lower cost, ease of data entry, flexibility of and control over format, advances in technology, recipient acceptance of the format, and the ability to obtain additional response-set information. In the field of counseling, an additional advantage is the ability to access individuals not involved in the mental health system.

Reduced response time. One of the primary advantages of e-mail and Web-based surveys is that they dramatically decrease response times (Lazar & Preece, 1999). Farmer (1998) reported that typical turnaround time is 4 to 6 weeks with traditional mail surveys, 2 to 3 weeks for telephone surveys, and only 2 to 3 days for Web-based surveys. Franceschini (2000) also noted reduced turnaround time. In his study, half of the respondents were sent mail surveys, the other half were surveyed via the Internet. He reported that 21 of the 29 Web-based responses were received before there were any responses to the traditional mail survey. In fact, Bauman, Airey, and Atak (1998) found that the majority of recipients of an e-mail survey either responded within 1 to 2 days of receiving the initial solicitation or not at all.

Lowered cost. Costs for e-mail and Web-based surveys can be substantially lower than for traditional mail surveys because there are no printing, postage, or stationery costs (Bauman et al., 1998). There are, however, some costs associated with Web-based surveys, primarily for programming, using space on a server, and some limited data entry and/or manipulation. Farmer (1998) argued that Web-based surveys are 50% less expensive to implement than telephone surveys, and 20% less expensive than mail surveys. In their study, Schleyer and Forrest (2000) found the Web-based survey to be 38% less expensive than mail surveys. If there are technical problems with the survey, however, costs associated with "help desk" technicians can rise dramatically. In one study, the "help desk" for an e-mail survey was overwhelmed and "handled over 900 incoming toll free calls regarding the survey, most of them . . . technical questions about e-mail" (Couper, Blair, & Triplett, 1997, p. 9). The authors of this study concluded that when the cost of maintaining the help desk was added to the start-up cost for the e-mail survey, there were no cost savings associated with the e-mail sur-

vey. It seems that initial piloting of electronic surveys is essential to reduce the number of unforeseen technical problems that can increase costs.

Ease of data entry. In traditional paper-and-pencil surveys, data entry can be extremely expensive and time-consuming. An electronic survey can be configured to send data to a database or spreadsheet, eliminating the need for manual data entry. This also eliminates potential errors in rekeying data. Automatic data entry is typically an advantage only for Web-based surveys. For e-mail surveys, data still need to be manually transferred to a database.

Flexibility and control over format. Using the Web allows researchers to use flexible design formats such as color, graphics, innovative question displays, split screens, embedded programs (applets), animation, and sound (Dillman, Tortora, & Bowker, 1999). Moreover, researchers can control the order in which respondents answer the questions more easily than with paper-and-pencil surveys, which allow respondents to flip back and forth among the pages and change answers (Wyatt, 2000). Other rules, such as "select one answer only" or "do not leave this question blank," can be enforced with radio buttons (Lazar & Preece, 1999). With electronic surveys, the order and formatting of questions can be easily altered, which is particularly useful for Delphi studies (Wyatt, 2000). One study found that dropout rates were significantly lower when demographic information was collected at the beginning of the survey (dropout rate 10.3%) rather than at the end (dropout rate 17.5%; Bosnjak & Tuten, 2001). Although flexibility of format can be extremely useful, there are no definitive answers as to the psychometric effects of the various Web-based formatting options (Arnau, Thompson, & Cook, 2001).

Advances in technology. Most Internet surveys are now done using HTML format, with the potential respondent contacted via an e-mail cover letter. HTML editors are becoming increasingly more sophisticated and easy to use, and data can be captured by a program on the server called a Common Gateway Interface (CGI) script. Several products exist that provide both the editing capacity for HTML and the necessary CGI scripts for capturing data. The most common of these are Microsoft's FrontPage and Macromedia's ColdFusion (Solomon, 2001). There are also some software programs designed specifically for Web-based surveying that offer some additional features, such as managing the distribution of e-mail cover letters, built-in statistical analysis, the ability to generate reports, and automatic tracking of respondents. Examples of these programs include Perseus's Survey Solutions for the Web, Creative Research System's The Survey System, and Survey Said™ Survey Software (Solomon, 2001).

Recipient acceptance of the format. There is some evidence that the Internet is becoming more acceptable to respondents as a method of collecting data, particularly for men (Dillman et al., 2001) and for individuals who are college educated (Cartwright, Thompson, Poole, & Kester, 1999; Franceschini, 2000). E-mail surveys tend to be very easily understood and completed by recipients, although there are cautions about reduced self-disclosure for e-mail surveys

because the recipient's e-mail address is attached to the response (Harris & Dersch, 1997). Several authors have noted that self-disclosure is increased when people communicate using Web-based surveys as compared with traditional paper-and-pencil surveys (e.g., Joinson, 1999, 2001; Moon, 2000). Conboy, Domar, and O'Connell (2001) found that an Internet survey offered the necessary assurances of anonymity to allow respondents to give accurate data surrounding very sensitive health issues, which is particularly relevant for researchers in the field of mental health where sensitive data is also collected. Participants seem to accept claims of confidentiality and anonymity, even though there is a possibility that Internet password and encryption codes could be broken (Harris & Dersch, 1997) or the Internet protocol (IP) address of the respondent could be identified.

Obtain additional response-set information. With traditional paper-and-pencil surveys, researchers can only see the *results* of the participants' responses. Using the Web, researchers can learn about the respondents' answering *process* (Bosnjak & Tuten, 2001). For example, with Web-based surveys, researchers can identify the number of people who viewed the survey compared with those who completed it or, if the software will allow, the number of people who started the survey but did not complete it (Bosnjak & Tuten, 2001). With both e-mail and Web-based surveys, information such as time of day or day of the week of the response can also be tracked. Participants in online discussion groups often log on in the evening and during the night. One study of participants in an electronic support group found that 31% of postings (i.e., messages or responses placed online) were between 11 p.m. and 7 a.m. (Winzelberg, 1997). It is possible that this is also when respondents complete electronic surveys, although this has not yet been studied.

Ability to access participants outside of the mental health system. Most data that are currently available on mental health are based on surveys of individuals who have already accessed the behavioral health care system and professional counselors (or other mental health care providers). Other research on mental health focuses on surveys completed by broad populations to which researchers have access (e.g., college undergraduate, workplace). However, Internet-based surveys provide access to a pool of participants who have an identified problem or are at risk and have not sought assistance. For example, if a researcher wished to study the psychological correlates of eating disorders, participants could be solicited to participate using a Web site that focuses on eating disorders or through an online support group for people with eating disorders. This eliminates the typical research limitation of poor generalizability to nonclinical populations. In other words, the survey would now be available to individuals at all stages of the disorder and would include all variations of beliefs about treatment. However, a limitation is that it would be possible for individuals without eating disorders to complete the instrument.

Limitations of Online Data Collection

Despite the many advantages of Web-based surveys, concerns about their use have been raised by researchers in many fields in which this methodology has been used. These concerns focus on the following limitations: representativeness of the sample, response rates, measurement errors, and technical difficulties.

Representativeness of the sample. Internet use in the United States is growing at a rate of 2 million new Internet users each month. In 1 month alone (September 2001), 143 million Americans (54%) used the Internet, representing a 26% increase over August 2000 (ClickZ Network, 2002). Despite this growth, access to the Internet remains unequally distributed throughout the U.S. population. Most Web users are White (87.2%), male (66.4%), married (47.6%), and highly educated, with almost 88% having some college education and more than 59% having obtained at least one degree. In addition, 48% of Internet users are 35 years old or younger (Graphics, Visualization, and Usability Center, 1999). There are recent reports indicating that the demographics of Internet users are becoming more inclusive. Annual growth rates for 1999 and 2000 were 25% for Internet use by individuals in the lowest income households (less than $15,000 per household per year). Internet use among African Americans increased at an annual rate of 33% during 2000 and by 30% for Hispanics during this same year (ClickZ Network, 2002). Nevertheless, it remains the responsibility of the researcher to ensure that all members of a defined population have equal access to the technology needed to complete the survey (Dillman, Tortora, Conradt, & Bowker, 1998). To the extent that certain portions of a target population are systematically eliminated from a sample, the generalizability of the survey results is compromised. In such cases, the reliance on Internet-based methodologies would be inappropriate.

Response rates. Several studies show that e-mail surveys produce a significantly lower response rate than traditional mail surveys (Bachmann, Elfrink, & Vazzana, 1996; Couper et al., 1997; Crawford, Couper, & Lamias, 2001; Tse et al., 1995). One study indicated no significant differences between the two methodologies. However, the population from which this sample was derived was faculty at a large state university, almost all of whom had access to electronic mail (Schaefer & Dillman, 1998). The only published study showing a higher response rate for the e-mail survey than for the traditional mail survey was published in 1992, when the novelty of e-mail may have been a significant factor (Parker, 1992). The lack of anonymity in e-mail surveys may contribute to the lower response rates.

Researchers doing studies using Web-based surveys have also found lower response rates than for traditional mail surveys (Medin, Roy, & Ann, 1999; Nichols & Sedivi, 1998). However, unless the Web-based survey uses a sampling method that allows only certain individuals to access to survey, it is impossible to know the response rates. For example, when participants for electronic surveys are recruited

via newsgroups, search engines, or electronic mailing lists, researchers are not able to pinpoint the number of individuals who received the information, and therefore they cannot determine response rates nor speak to the representativeness of the sample (Schleyer & Forrest, 2000). To circumvent this difficulty, many Web-based surveys make use of an initial e-mail to a targeted group that contains a specific URL to access the survey. This e-mail can also include an access code, password, or PIN to ensure that only those who have been targeted can complete the survey and to prevent any individual from completing the survey more than once (Wyatt, 2000). However, researchers are cautioned not to make the Web-based survey too difficult to access by requiring too many codes and passwords, because this added complexity can lower response rates (Cartwright et al., 1999).

To reduce the problem of lowered response rates, several researchers have advocated a system of multiple reminders. This can be done easily if the original solicitation was done via e-mail. Kittleson (1997) found that it was possible to double response rates by sending e-mail follow-up reminders, but others have claimed more modest success (Solomon, 2001). Crawford et al. (2001) found that response rates increased when participants were told in their initial e-mail solicitation how much time the survey would take, when they received an automated (embedded) password, and when they received frequent e-mail reminders. They also found that the use of open-ended questions (particularly early in the survey or in blocks of more than one of this type of question) contributed to high abandonment rates.

Measurement errors. Very little is known about the psychometric implications of changing a survey from traditional paper-and-pencil to an electronic format (Arnau et al., 2001). Wyatt (2000) cautioned that the effects of design choices must be investigated, noting that in translating the survey from paper-and-pencil to electronic formats, items could be perceived differently by participants, thus affecting the validity of the survey. For example, if a respondent did not scroll down to see an entire list of options in a list box or did not understand how to correct a mistaken response, this could adversely affect survey results. Others argue that as long as the electronic survey format is similar to paper-and-pencil surveys, traditional surveys can appropriately be transferred to the Web (Lazar & Preece, 1999). There is, however, no empirical evidence on either side of this argument, and research needs to be done to test the psychometric effects of moving paper-and-pencil surveys to the Web.

Technical difficulties. Not everyone who completes a Web-based survey will be extremely computer-literate, nor will everyone have access to the most up-to-date technology. Dial-up access is still the most popular method to access the Internet (80%; ClickZ Network, 2002), and 66.5% of Internet users have a connection speed of 56k or slower (Graphics, Visualization, and Usability Center, 1999). As of 1999, most Internet users had as their primary computing platform either Windows 95 (44%) or Windows 98 (18%), and the most commonly used Internet browsers were Navi-

gator, as part of Communicator (Netscape, 45%); Navigator, stand alone (Netscape, 18%); and Internet Explorer (Microsoft, 34%; Graphics, Visualization, and Usability Center, 1999). Researchers using Web-based surveys must ensure that their pages are easily downloaded and maintain their formatting in all types of software and hardware environments. In addition, formatting issues, such as the use of open-ended questions or questions arranged in tables, can lead to higher dropout rates, as can the absence of clear navigational aids on the Web site (Bosnjak & Tuten, 2001). Bowker and Dillman (2000) found that the placement of the electronic survey on the page (left-aligned, right-aligned) affected the respondents' reaction to the survey, with some who received the right-aligned survey stating that they were confused and rating the design unfavorably, although the placement did not affect response rates. Notably, one study found that although sophisticated formatting can make the survey more attractive and interesting, surveys with advanced features and sophisticated designs (HTML tables, multiple colors, motion, sound) had a 5% lower response rate than did simple surveys (black letters on a white screen; Dillman et al., 1998). These authors noted that the sophisticated design was slower to load, particularly on older browsers, and some older browsers were more likely to crash when attempting to load the survey. Respondents to the sophisticated questionnaire took more than twice as long to complete it as did those who used the plain questionnaire. In another study, researchers developed a complex Web-based survey. They found that although 523 potential respondents said they had access to the Internet, only 73 of those had the capability to respond to this technologically sophisticated survey (Nichols & Sedivi, 1998). Although it is clear that more research must be done and that the capacity of browsers to handle complex designs has undoubtedly improved since these studies were completed, there still seems to be a point at which adding more features to the surveys becomes self-defeating in relation to increasing response rates.

Dillman et al. (1999) developed a list of principles for respondent-friendly, Web-based survey designs that includes the following: a motivational welcome screen; formats that are similar to paper-and-pencil formats; limited line length to reduce the need for left-to-right scrolling; and computer-operation instructions for different question formats at the location in the survey where the instructions will be implemented, rather than at the beginning of the instrument. Piloting the survey with a representative sample of the target population and using a wide variety of computing formats should help to reduce technical difficulties (Wyatt, 2000).

STEPS IN ONLINE DATA COLLECTION

Researchers wishing to engage in online data collection can use the advantages and limitations previously outlined to decide if their research needs can be adequately met through the Internet. Despite the advantages of online data collection, the limitations must be addressed before a decision

can be made about whether this is the appropriate methodology for a particular study. For example, we decided to use a Web survey for a large-scale research project because, in addition to the advantages just outlined, many of the significant aforementioned limitations could be controlled. The potential population for the study was discrete and knowable (workers at a large public agency), all had Internet access, and an e-mail mailing list was available for all potential respondents. Thus, we could control for problems with representativeness (everyone knew about and had access to the survey) and we could ultimately know the response rates. Therefore, we decided that a Web-based survey would be a sound methodology (Wheaton & Granello, 2001). Other researchers will need to determine if online data collection can meet their research needs. On the basis of our own experience with conducting an online study and a comprehensive review of the literature, we recommend the following steps for those considering this type of data collection.

1. *Determine the population to be measured.* Do all members of the population have equal access to the Internet? Some examples would be students or faculty at a university or members of a work site where everyone has access to computers and the Internet. In these cases, is there a standard method that the target population uses to access the Internet (e.g., a standard configuration for work computers for a work site study) or will individuals be accessing the survey using a variety of formats? Is there a readily available online method (e.g., electronic mailing lists, Web sites) to solicit participation? If the population does not have equal access to the Internet, then clearly an online-only data collection methodology is not appropriate for the research design. In this case, it may be most appropriate to send out a traditional paper-and-pencil survey with a Web address that allows participants their choice of response formats. The operative question is "Will the sample accurately reflect the population?"

2. *Determine whether an e-mail or Web-based survey will be used.* Consider the limitations and advantages of both, as previously described. In addition, consider whether the information to be collected is too sensitive to be collected via e-mail (e.g., evaluations of employers, sensitive mental health data). If the survey is to be Web-based, is there adequate server space available? Researchers will need access to a www server that has the ability to execute CGI programs. Prior to writing the survey, researchers should contact appropriate Web administrators to discuss the survey. Some researchers have found that buying "neutral" server space from a commercial vendor reduces some of the bias associated with particular servers. For example, a researcher wanting to evaluate the program quality of a business may find that participants resist completing a survey on the company's server. On the other hand, academic researchers may deliberately choose to place a survey on a university server to gain credibility for the research.

3. *Develop the layout of the survey and the type of format for the questions.* With e-mail surveys, layout means using plain text format without the special features available in HTML.

It is important to use plain text because bullets, font changes, and tabs may appear differently on different systems or fail to appear at all. In some cases, the HTML code itself may appear in the e-mail message, making the message difficult, if not impossible, to read. Moreover, radio buttons or check boxes (commonly used in HTML-based forms) are not available in an e-mail message. An alternative to these formats is to create a space where an *X* can be typed, which serves the same purpose as the buttons.

Web pages created in HTML, on the other hand, can have specially formatted form fields, such as radio buttons, pull-down selection menus, and open-ended text boxes. The Web also offers easy movement within the survey so that subsections specific to one group can be created. For example, in our study, we had all participants complete the first portion of the instrument, then "click" on the appropriate link to go to specialized questions based on their job classification. Using this method, participants went directly to the pages that applied only to them. The drawback is that such forms must either be created directly in HTML or with a Web page development tool (e.g., Amaya, FrontPage, Dreamweaver, GoLive).

4. *Write the questions.* The development of the questions for an e-mail or Web-based survey is similar to the process used for a paper-and-pencil survey, with the same concerns about response-set bias, variation in formats, clear wording, and so forth (Dillman et al., 1999).

5. *Keep the layout simple, with easy-to-read fonts and a consistent layout throughout.* Good Web-page design dictates that pages be easy to read and navigate (World Wide Web Consortium [W3C], 1999). All documents should be accessible to persons with disabilities (see the Web Accessibility Initiative [WAI], 2002, for detailed instructions). Forms present special problems for persons with visual impairments, especially those using screen readers. Ensure that tab sequences follow the logical order of the questions (WAI, 2002, Guideline 9.4). A link at the beginning of the page that goes to a "Help For Persons Using Screen Readers" page can be very helpful. This type of help page should include an overview of the structure of the document, noting in particular that the document is a form. Screen readers such as JAWS (Job Access with Speech), one of the most common screen readers, have a special "form mode" that helps the user navigate forms. Such a page should also include instructions for alternative formats, such as a method for completing the form over the telephone, with assurances of confidentiality or anonymity.

6. *Be sure to address informed consent issues, including the name and contact information of the researcher.* To ensure informed consent over the Internet, Schmidt (1997) recommended that a separate consent screen appear before the respondent can access the survey. The consent screen would require participants to take some action (e.g., click on an "I agree" button) before they can move forward into the questionnaire.

7. *Determine how data will be entered into the computer.* This issue is best decided by planning how the data will be analyzed. Data from Web-based surveys can be saved di-

rectly to a database or saved as text for importing into another data analysis program. Obviously, e-mailed surveys are returned as e-mail messages, and responses must be manually keyed into the data analysis program. When using the Web, data can be saved in text format for import into another program. In such cases, we recommend saving the output as tab-delimited text. The reasons for this are twofold. First, if spaces or commas are used as the delimiter, respondents can enter (even inadvertently) these characters in places where the researcher does not expect them. If this happens, the data will become misaligned in the output, and all the data after the spurious entry will be in error. Second, tab delimiters have particular advantages when data are to be analyzed in a spreadsheet (such as Microsoft's Excel) or by a statistical package such as SPSS. When used with a spreadsheet, the data can be copied from the text file and then "pasted" directly into the spreadsheet with no formatting required. Tab-delimited data can also be imported directly into an SPSS program. If planning to use a statistical package, however, be sure that the variables' names in the text file conform to the naming requirements of the statistical package (SPSS, for example, requires that all variable names be eight characters or less and begin with a letter).

8. *Practice putting in data.* If using the Web, submit multiple entries of the data. Fill in every field on the survey. Enter sets of fake data to make sure that when the information is downloaded, it appears in the correct format. We found that entering sets of identical data (e.g., a set of all "5" responses, then a set of all "4" responses) could help us determine whether all data were downloaded and whether it appeared in the correct columns. The goal is to find errors by purposely making the system fail.

9. *Include "error detection" variables in anonymous Web-based surveys.* It is axiomatic to say that errors will occur. These errors can occur for a wide variety of reasons, many of which are beyond the researcher's control. When surveys are completed anonymously, there is no way to verify the accuracy of the data by contacting the respondent. Thus, a method of error detection must be used. We gathered the date, time, and Internet protocol (IP) address as error detection variables. This became an important method to solve the problem of several people submitting the form several times. This situation can occur when persons use computers with slow connections because with slow connections submissions can take several seconds, which caused some respondents to believe they had not submitted the form correctly and they submitted it again. We could check for such multiple submissions by noting that the date and IP address were exactly the same and the times of submission were very close together (typically within only a few seconds of each other), and, of course, the data were exactly the same for all submissions. (*Note.* The IP address does not identify the person who submitted the form but the computer from which it was submitted. In environments where a single IP address is assigned to one person, it would be possible to surmise that the owner of that computer submitted the form—although someone else might have been using the computer. Nevertheless, it would be unethical to tell participants that the survey was anonymous and then identify them surreptitiously. Outside of such rare closed environments, identification of the person from the IP address is virtually impossible.)

10. *Pilot the study using a subset of the target population.* As with a paper-and-pencil survey, the goal is to check for clarity of wording, participant acceptance of the questions, and so forth. An additional goal of the piloting process specific to online surveys is to have participants submit the survey from a variety of computers and Internet connections, using different browsers (e.g., Netscape, Internet Explorer), including all possible versions, on different platforms (e.g., MacIntosh and Windows). Also, employ persons with a variety of technological expertise, especially a novice group. Make sure the directions are clear, particularly for a novice user. Instructing the pilot group to "try to mess it up" will give researchers valuable insight into potential problems that can be remedied before the survey is administered. If possible, the researcher should be present with the members of the pilot group as they complete the survey so they can listen to comments and detect and make note of any misconceptions (Wyatt, 2000).

11. *Determine the schedule for initial mailing, including e-mail posting and reminders.* Remember, research shows that most online surveys are completed within the first few days. Thus, the traditional calendar for follow-up reminders that is used for paper-and-pencil surveys (for example, a follow-up postcard, 2 weeks after initial mailing) may have to be condensed.

12. *Download the data frequently.* It is advisable to download the data perhaps as frequently as every night for two reasons: to minimize the possibility that the data will be inappropriately retrieved by others (Harris & Dersch, 1997) and to have a backup of the data.

CONCLUSION

Online data collection provides researchers with numerous possibilities—and significant challenges. The benefits in reduced cost, ease of data entry, format flexibility, and ability to access different populations make this type of data collection extremely appealing. Nevertheless, as with any survey method, measurement errors, low response rates, and possible nonrepresentativeness of the sample must all be addressed before meaningful data can be obtained. The benefits of accuracy, low cost, speed, and data entry become meaningless if the limitations are not adequately addressed. In the case of online data collection, special attention must be paid to representativeness of the sample, and it is not an appropriate methodology in all cases. The suggestions outlined in this article are meant to assist researchers in determining whether online data collection is a viable methodology for their research and, if so, to aid in the planning and implementation in ways that emphasize the advantages and manage the limitations.

REFERENCES

Arnau, R. C., Thompson, R. L., & Cook, C. (2001). Do different response formats change the latent structure of responses? An empirical investigation using taxometric analysis. *Educational and Psychological Measurement, 61,* 23–44

Bachmann, D., Elfrink, J., & Vazzana, G. (1996). Tracking the progress of e-mail versus snail mail. *Marketing Research*, *8*, 31–35.

Bauman, S., Airey, J., & Atak, H. (1998, May). *Effective use of Web-based technology: Using the Internet for data collection and communication applications*. Paper presented at the 53rd Conference of the American Association for Public Opinion Research, St. Louis, MO. Retrieved February 11, 2002, from http://www.wirthlin.com/strategic/casefiles/wes.phtml

Bosnjak, M., & Tuten, T. L. (2001). Classifying response behaviors in Web-based surveys. *Journal of Computer-Mediated Communication*, *6*(3). Retrieved January 10, 2002, from http://www.ascusc.org/jcmc/vol6/issue3/boznjak.html

Bowker, D., & Dillman, D. A. (2000). *An experimental evaluation of left and right oriented screens for Web questionnaires*. Retrieved February 11, 2002, from http://survey.sesrc.wsu.edu/dillman/papers/AAPORpaper00.pdf

Cartwright, D. W., Thompson, R. J., Poole, M. C., & Kester, D. D. (1999). *Assessing distance learning using a Website survey*. (ERIC Document Reproduction Service No. 433781)

ClickZ Network. (2002). *U.S. Internet population continues to grow*. Retrieved February 11, 2002, from http://www.clickz.com/stats/big_picture/geographics/article.php/5911_969541

Conboy, L., Domar, A., & O'Connell, E. (2001). Women at mid-life: Symptoms, attitudes, and choices, an Internet-based survey. *Maturitas*, *38*, 129–136.

Couper, M. P., Blair, J., & Triplett, T. (1997, May). *A comparison of mail and e-mail for a survey of employees in federal statistical agencies*. Paper presented at the annual conference of the American Association for Public Opinion Research, Norfolk, VA.

Crawford, S. D., Couper, M. P., & Lamias, M. J. (2001). Web surveys: Perception of burden. *Social Science Computer Review*, *19*, 146–162.

Dillman, D. A., Phelps, G., Tortora, R. D., Swift, K., Kohrell, J., & Berck, J. (2001). *Response rate and measurement differences in mixed mode surveys using mail, telephone, interactive voice response, and the Internet*. Retrieved February 13, 2002, from Washington State University, Social & Economic Sciences Research Center Web site: http://survey.sesrc.wsu.edu/dillman/papers/Mixed%20Mode%20ppr%20_with%20Gallup_%20POQ.pdf

Dillman, D. A., Tortora, R. D., & Bowker, D. (1999). *Principles for constructing Web surveys*. Retrieved February 11, 2002, from Washington State University, Social & Economic Sciences Research Center Web site: http://survey.sesrc.wsu.edu/dillman/papers/websurveyppr.pdf

Dillman, D. A., Tortora, R. D., Conradt, J., & Bowker, D. (1998). *Influence of plain vs. fancy design on response rates for Web surveys*. Retrieved February 16, 2002, from Washington State University, Social & Economic Sciences Research Center Web site: http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf

Duffy, M. (2000). Web-based research: An innovative method for nursing research. *Canadian Oncology Nursing Journal*, *10*(2), 45–49.

Farmer, T. (1998, December). *Understanding this thing we call Internet research*. Retrieved March 23, 2002, from http://infotekonline.com/Info_Online_New/White_paper_on_internet.htm

Franceschini, L. A., III. (2000). *Navigating electronic survey methods: Three pilot studies*. (ERIC Document Reproduction Service No. 448183)

Graphics, Visualization, and Usability Center. (1999). *GVU's tenth WWW user survey*. Retrieved February 11, 2002, from Georgia Institute of Technology, GVU Center, College of Computing Web site: http://www.gvu.gatech.edu/user_surveys

Harris, S. M., & Dersch, C. A. (1997). Conducting research on the Internet: Potential, concerns, and reflections. *Kappa Omicron Nu Forum: Technology*, *11*(1). Retrieved February 16, 2002, from http://www.kon.org/archives/forum/11-1/harris.html

Joinson, A. N. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavior Research Methods, Instruments and Computers*, *31*, 433–438.

Joinson, A. N. (2001). Knowing me, knowing you: Reciprocal self-disclosure in Internet-based surveys. *CyberPsychology & Behavior*, *4*, 587–591.

Kittleson, M. (1997). Determining effective follow-up of e-mail surveys. *American Journal of Health Behavior*, *21*, 193–196.

Lazar, J., & Preece, J. (1999). Designing and implementing Web-based surveys. *Journal of Computer Information Systems*, *39*(4), 63–67.

Medin, C., Roy, S., & Ann, T. (1999). *World Wide Web versus mail surveys: A comparison and report*. Retrieved February 16, 2002, from http://www.anzmac99.unsw.edu.au/anzmacfiles/papers.htm

Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of Consumer Research*, *27*, 323–339.

Nichols, E., & Sedivi, B. (1998). Economic data collection via the Web: A census bureau case study. Proceedings of Survey Methods Section, 1998 American Statistical Association Meetings, Dallas, Texas.

Niles, S. G., Akos, P., & Cutler, H. (2001). Counselor educators' strategies for success. *Counselor Education and Supervision*, *40*, 276–291.

Parker, L. (1992, July). Collecting data the e-mail way. *Training and Development*, *46*, 52–54.

Schaefer, D. R., & Dillman, D. A. (1998). *Development of a standard e-mail methodology: Results of an experiment*. Retrieved February 11, 2002, from Washington State University, Social & Economic Sciences Research Center Web site: http://survey.sesrc.wsu.edu/dillman/papers/E-Mailppr.pdf

Schleyer, T. K. L., & Forrest, J. L. (2000). Methods for the design and administration of Web-based surveys. *Journal of the American Medical Informatics Association*, *7*, 416–425

Schmidt, W. (1997). World Wide Web survey research: Benefits, potential problems, and solutions. *Behavior Research Methods, Instruments, & Computers*, *29*, 274–279.

Solomon, D. J. (2001). Conducting Web-based surveys. *Practical Assessment, Research, and Evaluation*, *7*(19). Retrieved February 16, 2002, from http://pareonline.net/getvn.asp?v=7&n=19

Tse, A. C. B., Tse, K. C., Yin, C. H., Ting, C. B., Yi, K. W., Yee, K. P., et al. (1995). Comparing two methods of sending out questionnaires: E-mail versus mail. *Journal of the Market Research Society*, *37*, 441–446.

Web Accessibility Initiative. (2002, March 5). *Web Accessibility Initiative (WAI)*. Retrieved April 22, 2002, from http://www.w3.org/WAI/

Wheaton, J. E., & Granello, D. H. (2001). *The training needs of employees at the Ohio Rehabilitation Services Commission: Results of a multiphase needs assessment*. Unpublished manuscript.

Winzelberg, A. (1997). The analysis of an electronic support group for individuals with eating disorders. *Computers in Human Behavior*, *13*, 393–407.

World Wide Web Consortium. (1999). *Web content accessibility guidelines*. Retrieved April 22, 2002, from http://www.w3.org/TR/1999/WD-WAI-PAGEAUTH-19990217/#transform-gracefully

Wyatt, J. C. (2000). When to use Web-based surveys. *Journal of the American Medical Informatics Association*, *7*, 426–430.