Presented at the Association of Internet Researchers,
2nd International Conference, October 10-14, 2001
University of Minnesota

DRAFT

Web-based Surveys: Exploring Design and Implementation Issues

Dorine Andrews, D.C.D.
Georgetown University
dca4@georgetown.edu

Blair Nonnecke, Ph.D.
University of Guelph
rbn@acm.org

Jennifer Preece, Ph.D.
University of Maryland
Baltimore County
preece@umbc.edu

Abstract

Using the Internet to conduct research presents challenges not found in conventional research. Survey design, consent and subject privacy, sampling, subject solicitation, distribution media and technical operability are key issues that must be addressed in order to conduct sound online research. The authors' current research involves defining the nature of "non-public participation" (commonly referred to as lurking) in online discussion groups and gaining insights into related factors. It is research to follow-up earlier work that revealed significantly different volumes of non-public participation in different types of discussion groups [Nonnecke, 2000; Nonnecke and Preece, 2000]. This follow-up study involves an extensive online survey of discussion group participants. This paper examines the current knowledge about electronic surveys through the design and implementation decisions made for this research.

Introduction

It is well recognized in the behavioral sciences that surveys are not perfect vehicles for collecting data because surveys require subjects to recall past behavior [Schwarz, 1999]. Some social scientists contend that observation captures behavior more accurately [Bernard, et al., 1981; Bernard, et al., 1983] and there is ample data to support their position. For example, online consumers overestimate the amount they purchase online by 55% [Comscore, 2001]. Others suggest that the survey questions bias subject judgements and answers [Schwarz, 1999]. One alternative, many contend, is to collect behavioral data using multiple approaches [Sudweeks & Simoff in Jones, 1999; Rogers, 1987]. Observations, focus groups, individual interviews, email, Web-based, postal, and random digital dial telephone surveys can be used in combination to improve results quality [Smith, 1997] and sample representativeness [Yun & Trumbo, 2000; Swoboda, et al., 1997]. For example, Yun & Trumbo (2000) achieved a 72% return rate within a one-month period by combining postal, email and Web-based survey forms.

However, research costs, access to subjects, the scope of the research and the nature of behavior under study may make it impractical or financially unfeasible to use more than one data collection approach. Electronic surveys provide the ability to conduct large-scale data collection by others than organizations at the centers of power in society [Couper, 2000]. The technology provides an inexpensive mechanism for conducting surveys online instead of through the postal mail [Sheehan & Hoy, 1999; Weible & Wallace, 1998] and one in which costs per response decrease significantly as sample size increases [Watt, 1999]. Electronic surveys are becoming increasingly common [Lazar, J & Preece, J., 1999], and research comparing electronic vs. postal surveys is starting to confirm that electronic survey content results may be no different than

postal survey content results. In addition, electronic surveys provide the strong advantages of speedy distribution and response cycles [Yun & Trumbo, 2000; Swoboda, et al., 1997].

Some of our knowledge concerning the effective design and use of paper-based surveys does translate into electronic formats. However, electronic surveys have distinctive technological, demographic and response rate characteristics that affect how they should be designed, when they can be used and how they can be implemented [Sohn, 2001]. This paper focuses on those distinctive characteristics.

Two forms of electronic survey have emerged in the last fifteen years. The first, asynchronous email survey dates back to 1986 [Kiesler & Sproull, 1986]. The second, synchronous Web-based survey, started about 1994 [Kehoe & Pitkow, 1996]. There are several fundamental differences between email and Web-based surveys. First is database technology. Web-based surveys provide the ability to automatically verify and store survey responses using database technology and an HTML (hypertext markup language) user interface. Email survey responses, whether embedded directly within an email message or attached as a word processed document, must be manually entered into storage. Second, email is a "push" technology that allows researchers to directly communicate with prospective respondents. Web-based surveys do not provide this affordance of direct communication. This paper argues that Web-based surveys are superior to email surveys in many aspects, but that email combined, perhaps with offline media, is an excellent vehicle for inviting individuals to participate in Web-based surveys. The case study that follows informs development guidelines for Web-based survey design and use in five methodological components of online survey design and implementation. These are (1) survey design, (2) subject privacy and confidentiality, (3) sampling and subject selection, (4) distribution and response management, and, (5) survey piloting.

Research Background

The current research explores and defines the nature of participation in online discussion groups, especially those aspects of "non-public participation" commonly referred to as "lurking." When lurking is defined as "not posting to a discussion group," the mean level of "non-public participation" for all discussion groups is lower than the previously reported 90% and the volume of lurking can vary dramatically among different online discussion groups [Nonnecke & Preece, 2001]. For example, health-support discussion groups have, on average, significantly fewer lurkers (46%) than software support discussion groups (82%) [Nonnecke & Preece, 2000].

As recommended in Sudweeks & Simoff [1999], this research sought to balance the strengths of quantitative and qualitative methods to provide multiple perspectives that can be refined and integrated into a single model. Preceding the quantitative logging study [Nonnecke & Preece, 2001] a qualitative study was conducted to collect preliminary data as to why peripheral members of online communities "lurk." It consisted of ten face-to-face semi-structured interviews. The results guided discussion group selection for the logging study as well as the current study to examine causal factors. The interview results indicated that non-public participation is a strategic activity. Researchers were able to propose a preliminary model, called the gratification model, to categorize the reasons for lurking. It suggests that support and

information gathering needs can be met through non-public participation [Nonnecke & Preece, 2001]. Seventy-nine stated reasons for non-public participation were grouped into four categories: member characteristics, group characteristics, membership life cycle stage, and external constraints [Nonnecke & Preece, 2001]. This study, a quantitative online survey, will be used refine and validate the proposed model of the qualitative study. The research is seeking answers to the following questions:

1. Do age, gender, education level, experience with Internet, experience with online community technologies, work status, and/or work environment affect the amount of lurking one does and the reasons for doing it?

2. How does membership in multiple online communities, frequency of access, the type of community or the technology used by the community influence the amount of lurking one does or the reasons for it?
3. What, if any, are the relationships between lurking and stage of membership within an online community (e.g., joining and leaving an online community)?

4. What is the relationship of "feeling like a member of a community" and lurking?

The researchers were faced with many decisions regarding design and distribution of the survey. As researchers made decisions regarding survey design, subject privacy and confidentiality, sampling and subject selection, distribution and response rate management, and survey piloting, they used the knowledge from the literature (a complete reference list is included). In addition, these assumptions about the nature of the behavior under study and the scope of their research were made:

- All potential survey subjects are Internet users
- Getting people who do not post in discussion groups to respond to a survey may be problematic
- Subject identities are not required for anything other than follow-up interviews with a small subset of respondents to clarify response patterns, if needed
- Topic salience is going to be a major problem for almost all respondents (i.e. reasons for joining or leaving a discussion group are, most likely, of no interest to members of a sports discussion group)
- Using the population of all online discussion groups as the universe from which a representative random sample is drawn is not possible
- Researchers' university affiliations and reputations must be leveraged to establish survey credibility
- Funding to incent survey participation is not available
- The survey could easily become lengthy based upon the previous qualitative study results
- Results from the qualitative survey will guide the language used in the online survey

After reviewing the literature on electronic surveys, their first decision was to use a web-based survey with email invitations and reminders posted to public online discussion groups. The

research team had a working knowledge of survey development, web and database technologies as well as access to additional technical resources.

Survey Design Decisions

Cold Fusion, Microsoft Access and HTML were selected for the technical design. This allows researchers to (1) support multiple platforms and browsers [Yun & Trumbo, 2000], (2) prevent multiple submissions [Kehoe & Pitkow, 1996], (3) provide multiple opportunities for saving respondent answers [Smith. 1997], (4) collect both coded and open-ended responses [Yun & Trumbo, 2000], and, (5) provide immediate "thank-you" feedback upon survey completion [Smith, 1997]. This technical approach also provides the ability to track respondent identity for follow-up interviews if they "opt in" to follow-up interviews and, at the same time, protect individual privacy without the use of cookies [Cho & LaRose, 1999]. Also, if the survey, in piloting, takes too long to complete, it allows researchers the option of providing re-entry access using non-cookie passwords. The technology also allows a respondent who belongs to more than one community in the study to complete the survey for each community while at the same time preventing "ballot stuffing." The processing logic produces a gently worded error message: "It appears you have already completed a survey for this online community. Please contact the survey administrator at the link below to investigate the problem", if ballot stuffing occurs.

The survey is designed to have a professional, simple layout using a straightforward navigation strategy, keeping graphics and color to a minimum in an effort to add credibility to the survey as well as keep downloading time as short as possible [Couper et al., 2001; Dillman et al., 1998; Preece et al., 2002]. The survey, in its current state of design, has 28 primary questions, about 20 sub-set questions, and 12 demographic items. Researchers, following an introduction page, divided these questions into three sections, each having a "submit" function: 1) demographic questions, 2) questions related to the discussion group where invitation to participate was posted, and, 3) questions related to a discussion group that was permanently left. If the respondent abandons the survey, the data from the completed sections is not lost. This was a compromise between having the whole survey on a single page vs. displaying each question on its own page. The download and submit processing time required for over 50 single pages or one very large page was considered too burdensome. When a survey is completed, the respondent is thanked immediately and notified that his/her survey was successfully submitted.

The survey begins with a single introduction page. Like the email invitation that brought the respondent to the survey website, its purpose is to establish a trusting relationship with the prospective respondent that encourages him/her to proceed into the survey. To do this, there is text to (1) establish the authority and credibility of the researchers, (2) explain the survey purpose, (3) explain benefits of the results to online communities to address the salience issues of the survey, (4) establish respondent confidentiality and privacy, (5) provide open access to researchers through email address links to answer questions before starting the survey, (6) explain the sampling methodology, and, (7) provide a third party guarantee of the survey's authenticity and credibility using the Institutional Review Board approval with supporting links [Cho & LaRose, 1999].

Before demographic information is collected, a page for "opt-in" informed consent is presented along with links to small pop-up windows to display term definitions. For example, the terms *active*, *occasional*, *join*, *participate*, *leave*, *member* and *visitor* have links throughout the survey.

A small incentive to participate in the survey is provided, as the survey results will be available on line. By providing an email address, researchers will be able to notify the respondent of results availability. Respondents are also asked to "opt-in" to a follow-up telephone interview. Researchers explain that they intend to randomly sample a few respondents to explore results patterns more deeply, if needed.

The researchers are comfortable with asking both coded and open-ended questions. However, they are limiting the open-ended questions to optional opportunities to add information at the end of a coded question set and are using text-input boxes with wrapping and scrolling, not single line entry [Preece et al., 200l; Stanton, 1998; Andrews et al., 2001]. Skipping these questions does not affect the coded survey results. Coded questions use nominal scales, Likert scales, semantic differential scales, single and multiple choice selection options [Oppenheim, 1992].

Following Dillman's [2000] four stage development process, survey design required four rounds of prototype development before researchers felt they could proceed To Stage 2 – Cognitive Pre-testing. Unlike paper surveys where questionnaire presentation is stable, web-based survey question presentation requires the extensive use of HTML tables to control layout, wording and selection option alignment with testing on numerous browsers and preferences within browsers [Preece et al., 2002]. This was particularly important for any scales where a shift in alignment can cause misinterpretation of the question or make it unanswerable.

Question language proved more challenging than first anticipated by the researchers. In addition to maintaining question objectivity to control for bias, shorter sentences are better for reading on the screen. As Nielsen [2000] and others have demonstrated, people do not read web pages, they scan them, looking for key word and phrases. Therefore, survey questions and instructions became briefer as researchers reviewed the prototyped screens. For example, the original statement "*This second set of questions is similar to the first set, but focus on an online group which you have permanently left and no longer consider yourself to be a member of.*" became "*The questions below pertain to an online group you have permanently left."* There is a constant struggle to maintain the balance between brevity and a friendly tone.

Researchers also had to work to eliminate redundant questions and refine the ones kept. High non-response rates are anticipated to be due to the continual growth of email and use of electronic surveys [Couper et al., 1999; Schafer & Dillman, 1998]. They also anticipate a high attrition (drop out) rate if the survey was too long or irrelevant to the respondent. The online community shared interests (e.g., stock market, dieting, soap operas, health) are not those addressed in the survey. The survey is collecting data about "the different ways people use and participate in online groups." Therefore, each review round resulted in eliminating questions and revising the introduction and invitation.

Another challenge in question development faced by the researchers was how to present the 79 reasons and four categories for being attracted to, participating and/or leaving an online group generated from the initial interview-based study. The original idea was to use the results of the qualitative study directly, but researchers were concerned that the categories could be limiting and prototype testing proved that 79 items are overwhelming to review and select from on a single screen display without constant scrolling up and down. In the end, the researchers removed all items that could be considered duplications, put similar items together into groups of 3 and 4, and removed all category headings. An open-ended question, "Please add any other reasons you might have…" was added.

When all of this was completed, the survey was considered ready for Stage 2 -- Cognitive Pre-testing.

Subject Privacy and Confidentiality Decisions

To protect privacy and reduce intrusion, researchers decided to post the survey invitation only to public online discussion groups rather than email individuals. Direct email to individuals would have provided a mechanism for tracking individual responses, but was considered unnecessarily intrusive and, for some participants, would have been considered spam [Cho & LaRose, 1999]. In addition, obtaining email addresses for each group member is becoming increasingly difficult. As a result, respondent identity will be obtained only if he/she opts to provide it. The original intention was to post the invitation only after obtaining permission from the discussion group owner, but this proved problematic in the pilot work. The invitation will be posted without owner permission, unless discussion group policies directly require owner permission to post a message that is not directly "on topic."

The invitation text (1) explains the nature of the posting, (2) builds researcher credibility and authority, (3) demonstrates third party guarantee of trustworthiness by mentioning IRB approval, (4) explains discussion group selection methodology, and, (5) explains how taking the survey may benefit the potential respondent [Cho & LaRose, 1999]. After reading the invitation, discussion group readers can ignore the post or self-select to take the survey when they click on the survey URL in the invitation. The IRB process ensures that respondents understand what they are participating in, are told of any known risks, and requires documented respondent acceptance to participate in the research before research is conducted.

An "opt-in" informed consent approach was selected for this survey, as mentioned earlier. If a respondent clicks on "I do not accept" then proceeds with the study, the data will be permanently removed from the database. Under age respondents are asked to provide an email address of an adult who can give consent. Researchers will follow-up to obtain consent. If consent is not given, the responses of the underage respondent will be removed from the database before analysis begins.

To provide additional protection for respondent privacy, participation in the follow-up sample is completely optional through "opt-in" selection. In addition, an email address is not required and cookies are not used. As a result, survey tracking is at the discussion group, not the individual

level. If an email address is provided, this identity data is stripped from the main database table before analysis begins and is accessible only by the research team in a separate database table.

Sampling and Self-selection Decisions

Acknowledging the continuing decrease in response rates as online surveys proliferate, researchers anticipate low response rates from the target population. For this reason an easily replicated population (sampling frame) definition and sampling approach was developed [Couper, 2000; Coomber, 1997; Yun & Trumbo, 2000]. Researchers decided to use a probabilistic sampling method that begins with knowledge of the target population to permit the measurement of non-response at the discussion group level. However, the target population is not the total population of discussion groups, which is impossible to identify. Therefore, the results from this research will be considered indicative and no attempt will be made to infer to the general population of discussion group readers. Researchers hypothesize, based upon the high degree of variation in lurking among types of discussion groups, that there may be great variation in responses among readers in different discussion groups [Nonnecke & Preece, 2000]. Given the decision that results will only be indicative, not predictive, researchers created a sampling process that can be replicated across numerous discussion group populations. This will provide multiple, comparable indicative results which may be all Internet researchers can hope to attain given the ubiquitous and changing nature of the Internet. The following process documents the sampling process.

Step 1: Select a population to be sampled

This study is limited to asynchronous discussion groups because the previous studies were based asynchronous discussion groups. Discussion groups are aggregated by many different portal and non-portal resources such as MSN, Yahoo!, Catalist, Talkcity, Google, Altavista, http://tile.net/lists and http://webcom.com/impulse/list.html, as well as many websites that aggregate discussion groups of a particular character or interest. Because previous research clearly demonstrated that different types of discussion groups have different lurker characteristics, an aggregation of heterogeneous discussion groups was determined to be ideal to continue the study of this diversity. The "MSN web communities" have such diversity and was selected for that reason. There are 16 discussion group categories at the highest level of the MSN's community hierarchy. Twenty-five percent (25%) of these categories were selected using a random number generator to narrow the sampling frame. The categories selected were (1) health and wellness, (2) government, (3) sports & recreation, and (4) organizations. This population was further narrowed to ensure that discussion groups had sufficient critical mass (at least 50 members), were open to public participation, and were not just mailing lists, but active discussion groups (4-5 people posted within the past 90 days). A total of 1304 discussion groups were identified as members of this target population.

Step 2: Select a stratified random sample

Given this population, it is now possible to select a random sample from the frame. A stratified random sampling approach was used to ensure that each category was proportionally sampled (Table 1). If a category had additional hierarchical clustering of discussion groups, these were ignored for sampling purposes. All groups meeting the criteria were counted within that group as if there was no categorization below the highest level.

Table 1: Stratified Sample within the Sampling Frame

Category	Groups Meeting Criteria	Pop. %	Sample Taken
Health & Wellness	435	33%	122
Government	139	11%	41
Sports & Recreation	531	41%	152
Organizations	199	15%	56
Total	1304	100%	371

If this survey research is replicated with a non-categorized, but a heterogeneous discussion group population, then an inspection method should be added to categorize the discussion groups to ensure proportional representation in the sample. To attain a 95% confidence level that the sample results are inferable to the sampling frame, 359 discussion groups need to be surveyed. However, knowing the response rate may be less than 50%, over sampling was estimated to compensate for this lack of coverage. For this reason, the sample was adjusted to 371 discussion groups.

To measure non-response rates, the survey captures the name of the discussion group as entered by the respondent and/or the source URL of the discussion group on which the invitation was placed. This will also allow results to be analyzed by category. If one category appears to be proportionally underrepresented in the response rates, additional samples can be pulled in for the under-represented category. The process can be repeated using various discussion group aggregations with this sampling method until researchers decide that a sufficient number of survey responses have been collected for analysis.

Distribution Methods and Response Rate Management Decisions

Based on the way the survey is designed and the technology is applied, researchers will be able to estimate the non-response rate at the discussion group level and will be able to calculate the attrition rates at the respondent level by survey section.

The invitation to participate in the research is separate from the survey itself [Witmer et al., 1999; Mehta & Sivadas, 1995; Sheedan, 2001; Cho & LaRose, 1999]. It has been designed to build a trusting relationship from the beginning of the survey experience. The invitation will be posted to each group in the sample. Knowing that the majority of responses to electronic invitations occur very shortly after invitation posting, reminders will be posted each week for three weeks following the initial invitation [Yun & Trumb, 2000; Claycomb et al., 2000]. By that time the invitation should have been read by all readers who visit the discussion group at least once a month. The database will be examined before each reminder to determine if a reminder is warranted. If, after three reminders, few responses from a particular discussion group have been recorded, a decision may be made to select another discussion group for the sample.

Many design features are used to reduce attrition. The survey introduction will use full disclosure, direct access to researchers, and third party guarantor (IRB) to built trust and credibility in the researchers. Demographic data is gathered at the beginning of the survey [Frick et al., 1999]. A realistic estimate of the time required to complete the survey, a description of the survey structure and indicators of survey progress (using static statements) are provided.

Survey Piloting Decisions

Following Dillman's [2000] four stage piloting process, Stage 1 – Initial Survey Development is complete, as is Stage 2 – Cognitive Pre-testing although a slightly different approach was taken than is discussed by Dillman [2000] and Preece et al [2002]. The researchers constructed draft survey questions using a word processor. One researcher developed the online prototype, which went through three rounds of review with colleagues to ensure question completeness, efficiency, relevancy and format completeness. Stage 2 consisted of several subjects, not involved in the research, who completed the survey under the observation of a researcher using "think out loud" protocols with retrospective interviews. These cognitive pretests resulted in language simplification on the invitation and survey questions, changes in sequencing, and feedback on the look and feel of the survey. After the prototype was updated once more, an invitation to review the survey was placed on the AoIR listserve. Almost 40 people completed the survey and 15 people provided email feedback to varying degrees of detail. This pretesting produced an array of technical testing changes to privacy and confidentiality language and requirements, numerous recommendations for question wording, inconsistencies among questions and elimination of several questions.

The survey and the invitation will soon undergo final changes so that a Stage 3 - Pilot Test of the sampling technique with "live" discussion groups can be conducted. The language must encourage non-public participants in discussion groups to participate in the survey without alienating more active discussion group members. The invitation will be posted on approximately 30 discussion groups not included in the sample but included within the MSN online communities' portal. This will allow researchers to monitor for negative reactions to the posting, estimate non-response and attrition rates, and test analysis procedures before going to Stage 4 – Last Formal Review before the full study is conducted.

Conclusions

All the researchers involved in the development of this web-based survey are seasoned academics and professionals in their representative fields. Yet, despite this, they are continually challenged by the trade-off decision making required at every step of the survey process. Among their chief concerns are the sampling process and building a trusting relationship with prospective respondents. Every review uncovers new apparent weaknesses that require still more adjustments in either the survey or the distribution method. It is hoped that from this process, the researchers can not only address the issues of the subject of their research (i.e., further knowledge of lurkers and lurking), but also add to the knowledge regarding the design and using web-based surveys.

References:

_____. Papers and WWW user surveys methodology since 1994: http://www.cc.gatech.edu/gvu/user_surveys/

Andrews, D., Preece, J., & Turoff, M. (January 2001). "A Conceptual Framework for Demographic Groups Resistant to Online Community Interaction." Presented at the HICSS 2001 Conference. Honolulu, HI.

comscore. (February, 2001). "Online consumers proving erratic." NUA. Available at: http://www.nua.com/surveys/index.cgi?f=VS&art_id=905356472&rel=true

Bachmann, D., & Elfrink, J. (1996). Tracking the progress of email versus snail-mail. *Marketing Research* 8(2), 31-35.

Birnbaum, M.H., (May, 2000) "SurveyWiz and FactorWiz: JavaScript Web pages that make HTML forms for research on the Internet." *Behavior research methods, instruments and computers*, 32(2), 339-346.

Burgoon, J.K., Parrott, R., LePoire, B.A., Kelly, D.L., Walther, J.B., & Perry, D. (1989). "Maintaining and restoring privacy through communication in different types of relationships. *Journal of Social and Personal Relationships* 6, 131-158.

Cho, H. & LaRose, R. (Winter, 1999). "Privacy issues in Internet surveys." *Social Science Computer Review*, 17(4), 421-434.

Coomber, R. (June 30, 1997). "Using the Internet for survey research." *Sociological Research Online*, 2(2), 14-23.

Couper, M.P. (2000). "Web-based surveys: A Review of Issues and Approaches." *Public Opinion Quarterly* 64, 464-494.

Couper, M.P., Traugott, M. W., & Lamias, M.J. (2001). "Web survey design and administration." *Public Opinion Quarterly* 65, 230-253.

Couper, M.P., Blair, J. & Triplett, T. (1998). "A comparison of mail and e-mail for a survey of employees in federal statistical agencies. *Journal of Official Statistics* 15(1), 39-56.

Dillman, D.A. (2000.) *Mail and web-based survey: the tailored design method.* New York: John Wiley & Sons.

Frick, A., Bachtinger, M.T. & Reips, U-D. (1999). "Financial Incentives, personal information and drop-out rates in online studies." In U-D. Reips et al. (eds.). *Current Internet Science. Trends, techniques, results.* Available: http://www.dgof.de/tband99/pdfs/i_p/knapdf.

Kehoe, C.M. & Pitkow, J.E. (1996). "Surveying the territory: GVU's five www user survey's." *The Worldwide Web Journal 1(3).* 77-84. [Also on-line]. Available: http://www.cc.gatech.edu/gvu/user_surveys/papers/w3j.html.

Dillman, D., Totora, R.D., Conradt, J. & Bowker, D. 91998). "Influence of plain versus fancy design on response rates for web-based surveys." Paper presented at annual meeting of the American Statistical Association, Dallas, TX.

Kehoe, C.M., Pitkow, J.E. & Morton, K. (1997). "Eighth WWW user survey [online-line]. Available: http://www.gvu.gatech.edu/user_surveys/survey-1997-10.

Kiesler, S. & Sproull, L.S. (1986). "Response effects in the electronic survey. *Public Opinion Quarterly* 50, 402-413.

Knapp, F., & Heidingsfelder, M. (1999). "Drop out analysis: The effect of research design." In U-D. Reips et al. (eds.). *Current Internet Science. Trends, techniques, results.* Available: http://www.dgof.de/tband99/pdfs/i_p/knapdf.

Krosnick, J. A. (1999). "Survey Research." *Annual Reviews: Psychology* 50(1), 537-567.

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

Loke, S.D., & Gilbert, B.O. (1995) Method of psychological assessment, self disclosure, and experiential differences: A study of computer, questionnaire and interview assessment formats. *Journal of Social Behavior and Personality* 10, pp 255-263.

McCoy, S., & Marks, P.V., Jr. (August, 2001). "Using electronic surveys to collect data: experiences from the field." Paper presented at the AMCIS conference. Boston, MA.

Mehta, R. & Sivada, E. (1995). Comparing response rates and response content in mail vs. electronic mail surveys. *Journal of the Market Research Society,* 17(4), 429-440.

Nielsen, J. (2000). *Designing Web usability.* Indianapolis: New Riders.

Nonnecke, B. (2000). *Lurkers in email-based discussion lists.* South Bank University, London.

Nonnecke, B. & Preece, J. (2000). "Lurker demographics: counting the silent." *Human Factors in Computing Systems*. Paper presented at CHI'00. The Hague, Holland.

Nonnecke, B. & Preece, J. (August, 2001). "Why lurkers lurk." Paper presented at the AMCIS conference. Boston, MA.

Norman, K., Friedman, Z., Norman, K. & Stevenson, R. (2001). "Navigational issues in the design of online self-administered questionnaires." *Behaviour & Information Technology*, 20 (1), 37-45.

NUA. (2001). "Women now outnumber men on the Internet. http://www.nua.ie/surveys/index.cgi?f=VS&art_id=905356873&rel=true.

NUA. (2000). Income & Age Are Largest Gap in Digital Divide. http://www.nua.ie/surveys/index.cgi?f=VS&art_id=905355848&rel=true.

Oppenheim, A.N., (1992). *Questionnaire Design, Interviewing and Attitude Measurement*, New York: Pinter Publishers.

Paolo, A.M., Bonamino, G.A., Gibson, D., Patridge, T. & Kallail, K. (2000). "Response rate comparisons of email and mail distributed student evaluations." *Teaching and Learning in Medicine*, 12(2), 81-84.

Pereira, J., Bruera, E., & Quan, H. (Spring, 2001). "Palliative care on the net: An online survey of health care professionals." *Journal of Palliative Care* 17(1), 41-45.

Preece, J., Rogers, Y. & Sharp, S. (2002). *Interaction Design: Beyond Human-Computer Interaction*. New York: John Wiley & Sons.

Preece, J. (1994). *Human Computer Interaction.* Reading, MA: Addison-Wesley.

Robson, C. (1993). *Real World Research*. Oxford, UK: Blackwell.

Rogers, E. (1987). "Progress, problems and prospects for network research: Investigating relationships in the age of electronic communication technologies." Clearwater Beach, FL: Sunbelt Social Networks Conference.

Schaefer , D.R. & Dillman, D.A. (1998). "Development of standard email methodology: Results of an experiment." *Public Opinion Quarterly* 62(3), 378-397.

Schillewaert, N., Langerak, F. & Duhamel, T. (1998). Non probability sampling for www surveys: A comparison of methods. *Journal of the market Research Society* 4(40), 307-313.

Schwarz, N. (1999). "Self-reports: How the questions shape the answers." *American Psychologist* 54(2), 93-105.

Schwarz, N. & Sudman, S. (1996). *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Francisco: Jossey-Bass.

Sheehan, K. B. & Hoy, M. B.. (March, 1999). "Using e-mail to survey internet users in the united states: ethodology and assessment." *Journal of Computer Mediated Communication.*

Sheehan, K. B. & McMillan, S.J. (1999). "Response variation in email surveys; An exploration. *Journal of Advertising Research.* 39(4), 45-54.

Sheehan, K B. "E-mail survey response rates: a review." (January, 2001). *Journal of Computer Mediated Communication*.

Smith, C. B. (June, 1997). "Casting the net: surveying an Internet population." *Journal of Computer Mediated Communication.*

Stanton, J.M. (1998). Am empirical assessment of data collection using the Internet. *Personnel Psychology* 51(3), 709-726.

Sudweeks, F. & Simoff, S.J.(1999). "Complementary explorative data analysis: the reconciliation of quantitative and qualitative principles." In *Doing Internet Research: Critical Issues and Methods for Examining the Net*. S. Jones (editor). Thousand Oaks, CA: Sage.

Swoboda, S. J., Muehlberger, N., Weitkunat, R. & Schneeweiss, S. (1997). Web-based surveys by direct mailing: An innovative way of collecting data. *Social Science Computer Review* 15(3).

Taylor, H. (Winter, 2000). "Does Internet research work? Comparing online survey results with telephone survey." *International Journal of Market Research*, 42(1), 51-63.

Tuten, T.L., Bosnnjak, M., & Brandilla, W. (2000). "Banner – advertised web-based surveys." *Marketing Research*, 11(4), 17-21.

Walsh, J.P., Kiesler, S., Sproull, L.S., & Hess, B.W. (1992) Self-selected and randomly selected respondents in a computer network survey. *Public Opinion Quarterly*, 546, 241-244.

Watt, J.H. (1999). "Internet systems for evaluation research." In G. Gay & Bennington (eds), *Information technologies in evaluation: social, moral epistemological and practical implications*, 23-44. San Francisco: Josey-Bass, no. 84.

Weible, R. & Wallace, J. (1998). "The impact of the Internet on data collection." *Marketing Research*, 10(3), 19-23.

Witmer, D.F., Colman, R.W. & Katzman, S.L. (1999). "From paper-and-pencil to screen-and-keyboard." In *Doing Internet Research: Critical Issues and Methods for Examining the Net*. S. Jones (editor). Thousand Oaks, CA: Sage.

Witte, J.C., Amoroso, L.M., Howard, P.E.N. (Summer, 2000). "Research methodology—Method and representation in Internet-based survey tools—Mobility, community, and cultural identify in Survey 2000." *Social Science Computer Review*, 18(2), 179-195.

Yun, Gi Woong and Trumbo, Craig W. (Sept. 2000). "Comparative response to a survey executed by post, e-mail, & web form." *Journal of Computer Mediated Communication*.

Zhang, Y. (January, 2000). "Using the Internet for survey research: A case study." *Journal of the American Society for Information Science*, 51(1), 57-68.