The Online Survey: Its Contributions and Potential Problems

Hisako Matsuo, Kevin P McIntyre, Terry Tomazic, and Barry Katz,
Saint Louis University

This paper examines the use of internet-based survey methodologies in the behavioral and social sciences. Over the past decade, the use of the internet has become more widespread, and many social scientists have started to turn to online designs to conduct survey research (Fraley, 2004). Although the use of internet is still more common among people in the middle and upper-middle class individuals than those in the lower socioeconomic class, evidence indicates that this gap is narrowing and internet samples are becoming more representative (Gosling, Vazire, Srivastava, & John, 2004; Kraut, Olson, Banaji, Brukman, Cohen, & Couper, 2004). As a result, online surveys may have the potential to reach populations that are more diverse than the typical college sample, and they may be as effective as mail surveys (Gosling, et al., 2004).

There are several reasons why social scientists may want to employ internet-based designs. First, certain populations may actually be more accessible via online survey methodologies than traditional methodologies. Some subsets of society use the internet to communicate on listserves and newsgroups, which are often accessible to researchers (Kraut et al., 2004; Reips, 2002). Also, cross-cultural samples may be more accessible via online surveys when internet access is available in the culture of interest (Reips, 2002). Second, internet surveys may achieve sample sizes that exceed mail and telephone surveys (Kraut et al., 2004), depending upon the characteristics of target populations. Gosling and colleagues (2004) report collecting survey data from 361,703 participants via an Internet questionnaire, while Nosek, Banaji, and Greenwald (2002) report obtaining responses from 2.5 million individuals for their online test of implicit attitudes. Such large sample sizes are possible because internet designs are often more cost effective than other techniques (Reips, 2002). Unlike mail or telephone surveys, the costs of email surveys are not dependent on the number of participants solicited. Further, when entire populations of interest have internet access, it becomes possible to recruit from the whole population, rather than a randomly selected sample. For example, rather than sending recruitment emails to a sample of the population, surveys conducted on campus have recruited from the entire population of a university by obtaining a complete email list of all students, faculty, staff, and administrators, in order to assess campus climate (e.g., Women's Commission at Saint Louis University, 2003; Affirmative Action and Diveristy Committee at Saint Louis University, 2003; Michigan State University Graduate Program Evaluation, 2000).

However, while online surveys offer many benefits to researchers, some researchers have begun to identify methodological issues that need to be addressed when employing internet-based designs. We categorize these issues into methodological and data analysis considerations.

METHODOLOGICAL CONSIDERATIONS
Sample Representativeness
Although the diversity of the population using the internet has increased, issues of representativeness must still be considered when designing web-based surveys. Couper (2000) suggests that even if internet access were universal, sample representativeness would remain an issue because it is difficult to generate a sampling frame from which to recruit participants. Unlike telephone methodologies using random-digit dialing, there is no equivalent way for researchers to access the population within the web environment. Thus, although one may have access to greater numbers of participants, there is no way to determine whether one's sample is representative (Couper, 2000). Gonzalez (2002) likens the problems of representativeness in current online surveys to the classic blunders of political posters in the 1930's presidential elections in which biased sampling led to incorrect predictions that Landon would beat Roosevelt and that Dewey defeated Truman.

Nevertheless, in studies comparing demographic characteristics among online surveys and traditional sampling techniques, results indicate that internet surveys are similar to other techniques on several dimensions. Gosling and colleagues (2004) report that their internet study, although perhaps not representative of the population in general, still compared favorably to other published findings with respect to gender, socioeconomic status, geographic location, age, and race. Thus, the question is whether or not internet surveys produce representativeness similar to mail and telephone surveys.

Response and Nonresponse
When interpreting the results of online surveys, it is necessary to consider both who responded to the survey and who did not. Among the responses, self-selection, multiple submissions, non-serious responses, and dropouts are especially problematic in web-based designs (Gosling et al., 2004; Porter & Whitcomb; Reips, 2002). Self-selection may be common in internet designs because participants who respond may be especially motivated or interested in the research topic, exacerbating the problem of sample representativeness. However, self-selection is not more problematic in internet surveys than mail and telephone surveys. More attention should be given to multiple responses and non-serious responses, which result from the inherent ease of completing online surveys and the lack of control over who responds. In the online environment, identity theft, variable IP addresses, and firewalls make it difficult, if not impossible, to determine who is responding. Similarly, it is possible for individuals to affect the quality of the results by deceptively or falsely answering questionnaire items (Nosek et al., 2002) or by simply submitting their response multiple times. Nosek and colleagues suggest several way for researchers to deal with these issues including requiring participants to provide a unique identifier at the beginning of the survey, to have a couple of questions with definite answers (e.g., birth date) repeated in the survey, and to provide questions specifically designed to catch malicious participants (i.e., questions that have intentionally odd choices). Gosling et al. (2004) also suggest searching submissions for strings of identical responses and examining the internal consistency of responses.

Just as researchers need to control the quality of participant responses, they also need to consider the problem of nonresponse. Online surveys often do not have a defined

sampling frame (Couper, 2000); therefore, it is impossible to calculate the response rate for such studies. Research comparing response rates among online, mail, and telephone surveys suggest that response rates are generally lower for online surveys than for mail or telephone surveys (Kraut et al., 2004). Further, techniques commonly used to increase response rates for mail and telephone surveys (e.g., personalizing the survey, increasing the authority of the researcher) do not seem to increase rates for online surveys (Porter & Whitcomb, 2003). Additionally, even when one can attract potential participants to a survey, it seems difficult to get them to complete the survey. In a research study by Porter and Whitcomb (2003), of the total number of individuals who clicked the hyperlink to their survey, relatively few attempted to complete the survey. Of those that began to complete the survey, there was a high dropout rate and only a small number actually finished the survey.

Controlling Testing Conditions
In many survey designs, properties of the testing conditions are controlled in order to minimize measurement error. For example, in telephone surveys, the researcher can determine the order in which questions are asked, thereby establishing ordered relationships among independent and dependent variables. In the case of online surveys, the researcher has little control over the testing environment and is susceptible to increased variability in testing conditions. For example, it is difficult for researchers to control the order in which participants complete online surveys (Nosek et al., 2002). This is especially the case when questionnaire items are presented on the same page, thereby allowing participants to scroll down the screen and answer questions in the order they desire. Some survey software programs (e.g., TestPilot, SurveyMonkey) can prevent participants from advancing to latter items without first answering previous questions by separating questions onto different pages; however, this may make the perceived length of the survey longer leading to greater dropout. Additionally, some web-based questionnaires employ multimedia features, such as images, sounds, and other graphics. Unfortunately, it is impossible for the researcher to be sure that these features will be displayed properly across different computer systems, if they are displayed at all. As such, researchers should especially consider the degree to which their surveys require strict control over testing conditions before employing web-based designs.

DATA ANALYSIS CONSIDERATIONS
Large Sample Sizes
Most research conducted in the behavior and social sciences use inferential statistics for data analysis. By convention, researchers typically determine statistical relationships between variables by employing various tests of significance (Soric, 1989). These tests (e.g., t-test, analysis of variance) determine the likelihood of obtaining the results, assuming no relationship between two or more variables (Soric, 1989). However, null hypothesis significance tests are dependent on the sample size to determine the degrees of freedom, and the subsequent critical values. As such, large sample sizes may yield significant p-values for relationships that may not be meaningful in a practical sense (Borenstein, 1997; Soric, 1989). For this reason, researchers and journals are emphasizing estimates of effect size, rather than only on significance tests. In studies employing internet survey designs, it is possible to obtain very large sample sizes, thereby increasing the statistical power to detect an effect (Borenstein, 1997). Thus, although the results of a statistical test may be significant, the effect may be too small to be meaningful. In this way, researchers employing internet designs should consider calculating estimates of effect sizes, as well as determining what constitutes a meaningful effect prior to analysis.

Use of Inferential Statistics in Nonprobability Sampling
Inferential statistics should be used when the data are obtained through probability sampling. However, researchers often use nonprobability sampling, such as quota sampling, purposive sampling, and convenience sampling, because probability sampling is either costly or impossible to conduct due to unavailability of a sampling frame. Nonprobability sampling is frequently used, especially in market research and public opinion surveys (Levy and Lemeshow , 1999). Although inferential statistics is also used to analyze data collected through nonprobability sampling, it is fairly difficult to evaluate validity and reliability of the data. Similarl, researchers using nonprobability sampling for online data collection need to evaluate statistical estimates with more caution, considering representativeness of the sample and validity of the resulting estimates.

Analysis of Written Comments
Unlike traditional mail surveys, online surveys tend to yield a large amount of qualitative data obtained from open-ended questions. For example, about 50% of the sample (n=1300) of a needs assessment, conducted at one private university in the Midwest, responded to three open-ended questions, yielding rich qualitative information, which would not have been obtained through a traditional paper survey (Women's Commission at Saint Louis University). This example suggests that there might be some psychological impact of dealing with cyberspace and the ease of typing up written comments instead of writing with a pen.

A major advantage of administering open-ended questions through an online survey is that the data are easily converted to Word files, which enables researchers to analyze data both manually and via computer programs. Although qualitative data obtained through surveys cannot provide in-depth information, unlike data obtained through qualitative face-to-face interviews, an online survey surpasses a paper survey in terms of collecting rich data from open-ended questions.

CASE STUDY
One of the authors was a principal investigator of a campus climate survey conducted by the Affirmative Action and Diversity Committee of the Faculty Senate at her university. The survey consisted of 90 Likert-scale questions and three open-ended questions, which explored respondents' perceptions about campus climate, experiences of discrimination, and perceptions of the university's commitment to diversity. The target population of the survey included students, staff, and faculty of the university. The research team chose not to conduct a stratified random sampling because of the complexity of the issues, which tapped not only race and sex but also religion, nationality, and sexual orientation. Furthermore, there is no established response rate for an online survey in the

existing literature (Vehovar, Batagelj, Manfreda, and Zaletel, 2002). The survey was sent to all students (N=8500), and part-time and full-time staff and faculty (N=5,000), as a bulk e-mail with a Web site for the survey included in the message. In order to reach the segment of population which might not use e-mails frequently as a means of communication, such as staff in the Facilities Department, 500 paper questionnaires were also mailed.

The online survey demonstrated three major advantages in comparison with the mail survey: speed of response, cost for distribution of questionnaires, and speed and precision of data compilations. The project yielded a sample size of 1870 (984 students, 537 staff, and 349 faculty). Among these1870, about 35% responded within 24 hours after launching the bulk mailing, about 25% within 48 hours, about 20% within 72 hours, about 15% within a week, and the remaining 5% after the first follow-up message. Although follow-up messages were sent twice, at three-week intervals, only a few subjects responded to the second follow-up message. This project did not involve any costs for printing or postage for distribution of the survey. Both numeric data and written comments were downloaded immediately, and converted to SPSS and Word files respectively, enabling the research team to conduct data analysis on day after the data collection was completed. The online survey was very cost effective for this project because of electronic distribution and speed of data collection. Further, about 50% of the sample responded to either all of the three open-ended questions or at least one of them, yielding 126 pages of written comments, typed in 10 point.

Although the sample size was large, the project team had to consider the validity of the data because of the low response rate (about 12% for students, and about 25% for the full-time faculty and staff). However, it is not appropriate to calculate a response rate by dividing the sample size by the number of questionnaires mailed out. Online surveys go through several layers of potential nonresponses: E-mails might be lost in cyberspace; one might not click the e-mail; one might not go the Web site; and one might not be able to access the Web site (Vehovar, Batagelj, Manfreda, and Zaletel, 2002). Therefore, the response rate of this project should be the proportion of response to the entire population. In order to justify representativeness of the sample, racial and sex compositions of each sample of the students, staff and faculty were carefully compared to those of the population. Also composition by college of the student sample was compared to that of the population. Although this was done in a post-hoc manner, the sample represents some of the demographic characteristics of the population. The research team agreed that the sample is relatively representative of the population in terms of demographic characteristics.

CONCLUSIONS

Online surveys are cost effective in general. They are probably the most cost effective means of data collection when the target population is students and employees on a college campus. Some of the advantages of online surveys over mail surveys are: no need for printing or postage; speed of data collection; and precision of data compilation as demonstrated in the case study. However, there are some disadvantages in online surveys.

These are low response rates, nonresponses, nonrepresentativeness of sample, and lack of validity of the data. Furthermore, a large sample size is likely to lead to a small significance level, thus requiring additional statistical tests to justify the results. When nonprobablity sampling is used, researchers should use inferential statistics with caution. Although there is some literature on online surveys, there is much need for additional contributions to the body of knowledge in this area.

REFERENCES

Borenstein, M. (1997). Hypothesis testing and effect size estimation in clinical trials. *Annals of Allergy, Asthma, & Immunology, 78,* 5-16.

Couper, M.P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64,* 464-494.

Fraley, R.C. (2004). *How to conduct behavioral research over the internet.* New York: Guilford Press.

Gonsalez, J.E. (2002). Present day use of the internet for survey-based research. *Journal of Technology in Human Services, 19,* 19-31.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist, 59,* 93-104.

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: Report of board of scientific affairs' advisory group on the conduct of research on the internet. *American Psychologist, 59,* 105-117.

Levy, P.S., Lemeshow, S. (1999). *Sampling of Populations.* New York: John Wiley & Sons, Inc.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). E-research: Ethics, security, design, and control in psychological research on the Internet. *Journal of Social Issues, 58,* 161-176.

Porter, S.R., & Whitcomb, M.E. (2003). The impact of contact type on web survey response rates. *Public Opinion Quarterly, 67,* 579-588.

Reips, U. D. (2002). Standards for Internet-based experimenting. *Experimental Psychology, 49,* 243-256.

Soric, B. (1989). Statistical "discoveries" and effect-size estimation. *Journal of the American Statistical Association, 84,* 608-610.

Vehovar, V., Batagelj, Z., Manfreda, K. L., and Zaletel, M. (2002). Nonreponse in Web Surveys. In Groves, R., Dillman, D., Eltinge, J, and Little, R. J.A., eds. *Survey Nonreponse.* New York: John Wiley & Sons, Inc.