Running head: CONDUCTING WEB-BASED SURVEYS

Strengths and Weaknesses of Conducting Web-based Surveys: A Review of the Literature

Jennifer A Weber[1] and Kelly D Bradley

University of Kentucky

[1] Use Jennifer Weber as author of contact, jweber@email.uky.edu

Abstract

Surveys have been and remain a popular method for data collection in the social, behavioral and consumer sciences. The induction of the Internet into survey research methods brought with it technological breakthroughs and difficulties. Being able to collect individual's thoughts, interests, opinions, behaviors and attitudes in this format has advantages and disadvantages. Advantages include quick response time, lower cost and a better sense of anonymity for those responding. One of the most pronounced limitations applies to sampling bias, specifically to the lack of representation of various groups. This review provides an overview of reliability and validity issues associated with conducting Web-based surveys.

Strengths and Weaknesses of Conducting Web-based Surveys: A Review of the Literature

Surveys have been and remain a popular method for data collection in the social, behavioral and consumer sciences. The advent of the Internet spun survey research onto the World Wide Web, an event which provided researchers with a new vehicle for data collection, a technological update for gathering participants' thoughts, interests, opinions, behaviors and attitudes (Rezabek, 2000). As Ladner, Wingenbach and Raven (2002) report, "a simple Internet query of 'Web-based surveys' produce[s] over 18,500 matches," most of which were advertising the for fee services of companies offering Web-based survey software and implementation support. Currently, an Internet search of the exact phrase 'Web-based surveys' produces 196,000 results using Yahoo! search engine, many of which appear to be websites for consumers purchasing or testing Web-based survey software. Skitka and Sargis (2006) report one-fifth of all 2003-2004 American Psychological Association Journals published at least one article which used Web-based data collection methods. Here, advantages and disadvantages of the Web-based approach are discussed as supported by the current body of literature.

The induction of Internet usage in the methods of survey research has provided several advantages compared to more traditional methods of data collection, such as pencil-and-paper questionnaires and phone interviews (Lyons, et. al., 2005; Wright, 2005;). Coupled with advancements are limitations of the more technological methods. Regardless of format, survey research techniques require reliability and validity of the instrument so that the measurement is credible and the subsequent data collected is of high quality. In general, surveys tend to have strong reliability and weak validity estimates (Babbie, 2001).

Purpose

This paper provides an overview of reliability and validity issues associated with conducting Web-based surveys, viewing these issues through the lenses of advantages and disadvantages. The research is guided by the underlying question: What are the strengths and limitations of using a Web-based format when conducting survey research? This paper outlines the current methodological strengths and limitations presented throughout the literature.

Method

This study reviews the literature in order to identify strengths and limitations with the integration of the internet, specifically Web-based techniques, to survey research methods. Questions regarding reliability and validity issues associated with the methodology are examined by searching the literature bases of PsycINFO, ERIC, and Professional Development collections hosted by WebSPIRS and EBSCOhost accessed via the university's library website. Keywords or phrases including survey research, Internet, Web-based, online, data collection methods, reliability and validity. Relevant research was also identified using a snowball approach when reviewing the reference sections of articles found through the computer-based literature search. Additional selection criteria included works published after 1997 in a peer-reviewed journal, national conference proceedings or dissertation.

Results

Strengths of Web-based Surveys

The methodological insurgence of Internet-based methods of collecting survey data has a number of advantages over the more traditional methods, like pencil-and-paper. Efficacy, specific to time and money, is the most frequently reported advantage (see Lyons, et. al., 2005; Wright, 2005; Yun & Trumbo, 2000; Edmunds, 1999; Tourangeau, 2004; Skitka & Sargis,

2006). Collecting data online, in general, takes comparatively less timely and is a less expensive avenue for tapping into basic human attitudes, opinions and behaviors. In an annual study on the influence of the Internet on Americans, the results indicated 78.6% of Americans went online in 2005 at an average of 13.3 hours per week, with 66.2% of those individuals using home access (USC, 2005). This capability of reaching thousands, even hundreds of thousands, of individuals in one click is a survey researcher's dream come true.

Data collection time has also been well documented. In a study using both Web-based and pencil-and-paper surveys, Cobanoglu, Warde and Moreo (2001) found a mean response speed of 5.97 days for the Web-based surveys compared to 16.46 days for mailed surveys. The turn around time for Web-based surveys has been reported as two to three days by Yun and Trumbo (2000), with 80% of responses collected in the first three days, most of which are submitted within the first 24 hours. The speed of return is due to such factors as the absence of printing and mailing (Lyons, Cude, Lawrence, & Gutter, 2005). Another contributor is the increased frequency with which Americans check email. A report from the USC Annenberg School (2005) found 90% of Americans who go online use email and check electronic messages several times a day on average. Survey researchers are encouraged by the quicker turn-around time, especially in light of declining response rates over the last four decades for traditionally disseminated national surveys (Krosnick, 1999). The quickness of response may also increase researchers' ability to implement test-retest and parallel form reliability strategies.

Cobanoglu and colleagues (2001) also looked compared costs associated with conducting Web-based and hard copy survey research. Cost comparisons revealed an estimated total cost for mailed surveys to be 2.4 times greater than costs for Web-based surveys. Ladner, Wingenbach and Raven (2002) compared data collection methods of the Internet and pencil-and-paper in a

study on agricultural education and found the costs of conducting pencil-and-paper survey research to be 11 times more expensive than the Web-based counterpart, even when factoring in the purchase of a Web-based software package at $550. Likewise, Watt's (1999) evaluation concluded a typical cost of $0.65 per 10,000 respondents via the Web and $1.64 for emailed surveys, both much less expensive than mailed surveys.

In more traditional survey settings, such as face-to-face interviews, telephone interviews, or pencil-and-paper surveys completed in a lab, researchers must be vigilant about response bias. Participants in these settings have been found to provide more socially desirable information and observable behaviors due to the simple presence of the researcher, or assistant (http://www.utexas.edu/learn/surveys/advantages.html). Although Yun and Trumbo (2000) offer a point of dissention, a vast number of studies reveal a less inhibited, more open and honest response pattern with online participants, which may be due to the physical distance inherent in Web-based survey research (Lyons, et. al., 2005). Edmunds (1999) and Rezabek (2000) even point out that bias is reduced when using Web-based measures compared to mail and telephone surveys. Studying school-aged children, Vereecken (2001) found participants were more likely to endorse involvement in risky behaviors using a Web-based survey format than when using a pencil-and-paper format. In general, Web-based surveys potentially reduce response bias and improve the accuracy of the data collected (Wright, 2005).

Lyons and colleagues (2005) argue the quality of responses gathered using Internet-based methods is at least equal and is some cases better, especially with regards to sensitive topics of inquiry, to the quality of traditional methods. In a study on gay and lesbian family dynamics, Moorefield and Proulx (2004) found an increase in the breadth and depth of responses given by families using Web-based data collection methods when compared to data obtained from face-to-

face interviews. Buchmann, Elfrink, and Vazzana (2000) reported increased response rates and more thorough answers with greater self-disclosure to open-ended questions when utilizing a Web-based format, as compared to telephone and traditional mail surveys. While confidentiality is difficult to guarantee in any setting, Web-based survey methods seem to offer individuals a better sense of anonymity, leading to a decreased likelihood of response bias and increased response rate (Skitka & Sargis, 2006: Daley, McDermott, McCormack-Brown, & Kittleson, 2003).

Several studies have provided researchers with evidence suggesting Web-based surveys are more practical and desirable than traditional methods of survey research with "hard-to-reach" (Lyons, et. a., 2000, pg. 345; Andrews, Nonnecke, & Preece, 2003). Because of the Internet's world-wide nature, Web-based surveys have the potential to reach participants around the globe very quickly, creating an international sample that allows for a more worldly view of the construct being studied (Edmunds, 1999). The collection of data from vast numbers and geographical regions of people can aid in establishing better convergent and divergent construct validity. Access to international samples provide researchers with a convenient grouping variable for comparing responses on the basis of differences between languages, religion and other cultural factors (Litvin & Kar, 2001). Swoboda, Muehlberger, Weitkunat and Schneeweiss (1997) were able to receive responses to a Web-based survey from all parts of the world, with 90% of participants responding within four days. Their study is an indication that e-surveys, even in the English-language, may overcome international barriers. Internet access to populations from all over the world may result in larger sample sizes, which increases the power and reliability.

Yet another advantage of Web-based surveys is the programming services and software involved, which helps assure reliability and validity of surveys. Web-based survey instruments can be created and deployed through various Web-based programs and software packages. Most programs have the ability to require participants to respond to certain, if not all questions, on the survey. The response requirements of Web-based instruments decrease the likelihood of missing data and response errors, making the response set more reliable and valid (Skitka & Sargis, 2006; Yun & Trumbo, 2000). Many of the Web-based formats of surveys offer instant response collection over the Internet, which eliminates the time needed for entering data into analytical software packages. Web-based data processing is beneficial because it means less threat of unreliable researcher observations and less handling of the data by research personnel, which results in lower risk of data coding and entry errors (Lyons, et. al., 2005; Wright, 2005). While the advantages and new opportunities provided by Web-based survey research are far-reaching, the limitations imposed by the methodologies also need to be noted.

Limitations of Web-based Surveys

It is true that there are clear advantages to implementing surveys in a Web-based format; however, there are limitations associated with this method as well. Nardi (2003) clearly outlines concerns associated with sampling. Even though traditional survey methods share similar issues with sampling frames, and specifically representation (Krosnick, 1999), Web-based surveys are confronted with limited access to certain demographic groups, which restricts generalizability (Eastin & LaRose, 2000; Tourangeau, 2004; Skitka & Sargis, 2006). While internet access continues to grow, certain populations are being excluded from technological advancements (Birnbaum, 2004). Eastin and LaRose (2000) suggest the digital divide “is undoubtedly one of

the most important social equity issues facing the information society and is international in scope" (Introduction section, ¶ 2).

Race, socioeconomic status and age are variables by which Internet use differentiation is present. The Pew International and American Life Project (2003) characterized users and nonusers in 2003. Of the 58% of Americans estimated to go online in the Pew study, only 8% were African American and 9% were Hispanic. The study's estimates found only 18% of users had incomes less than $35,000 and less than half were from rural areas. The report described nonusers as persons from a minority group with a modest or less than modest income and education (25% of nonusers were without a high school diploma). Fifty percent of nonusers were older than 50 years, 30% were retired, 16% were unemployed, and 26% were disabled (compared to only 12% of users). Simply stated, Web-based instruments are limited by a threat to external validity. Skitka and Sargis (2006) further claim Internet users within a subculture are different than nonusers within the same culture. Therefore, the internal validity may also be threatened if the variables distinguishing users from nonusers interfere with the constructs intended to be measured by the instrument.

Another disadvantage threatening the reliability and validity of Web-based surveys is nonresponse. There is disagreement over how great the difference is in response rates between Web-based and pencil-and-paper surveys (Tourangeau, 2004; Yun & Trumbo, 2000). Nonresponse errors refer to solicited participants' choice not to take part in a study and can include non-received emails and solicitations deleted by potential participants (Skitka & Sargis, 2006). Initial success with Web-based surveys was argued to be an effect of the novelty of the Internet (Dillman, 2000). As internet abuse, privacy concerns (Tourangeau, 2004), commercial advertisements and 'junk' emails have increased, it has become extremely challenging, if not

impossible, to compute actual response rates (Birnbaum, 2004). Lyons and colleagues (2005) suggest a number of methods to increase response rates and mitigate nonresponse biases. They suggest 1) advance notice, such as a personalized cover letter, to potential participants, 2) email and mailed reminders to participate, and 3) incentives to participate, such as a chance to win money or gift certificates, may increase response rate. Dillman (2000) found a significant increase in responses with reminders but recommend limiting follow-up to two emails so as not to dissuade people from participating.

Lower response rates do not necessarily equate to an increase in nonresponse errors. Assuring valid responses and thereby minimizing the creation of error in statistical measurement can be a difficult task for survey researchers regardless of the approach taken for data collection. Even so, strategies for the assessment of online responses can be formulated to improve the accuracy of response patterns. Many Web-based survey software programs provide controls for response patterns and item positions that may cause fatigue and lead to nonresponse bias or response errors. In terms of the general construction, careful wording, format and content selection through empirical and theoretical review of the constructs and methods can significantly reduce the unreliability of participant responses and increase the face, content and construct validity of an instrument (Nardi, 2003).

Because the quality of the sample is only as good as the sampling frame, rather than selecting particular closed-set lists (LISTSERVs) for the sampling frame, a broader dissemination to free-access, commercial email directories (e.g., Yahoo) may target a broader sample of the Internet-user population. Keeping the Pew (2003) study's characterization of users versus nonusers in mind, over sampling underrepresented populations may be necessary to boast the generalizability of a survey's results. Therefore, efforts to not only include but specifically

target larger numbers of minority groups may yield a more representative sample of the general population. However, researchers must also consider other characteristic differences between Internet users and nonusers, such as SES, age, and ability.

Other limitations related to Web-based survey methodologies include the occurrences of multiple responses from a single participant and the receipt of unsolicited responses. Participants may either intentionally submit their responses multiple times, possibly to increase their chances at winning incentives, or unintentionally hit the submit button more than once. Along the same lines, unsolicited responses may occur if the solicitation for participation is passed from the intended party to an outside participant that was not originally included in the sampling frame nor detected in the final data set (Lyons, et. a., 2005). Researchers may be able to counter this by using e-surveying services which can provide assistance in validating the origin and uniqueness of responses by tracking email and IP addresses.

Other technical issues exist for Web-based survey users. Yun and Trumbo (2000) suggest hardware and software problems exist including a difference in browsers, which may present the same Java Script differently, platforms and processors, and monitors, which display graphical images differently or not at all. Differences in appearance of the survey for participants can cause differential responses, leading to poor reliability. Another technical problem is computer freezes and crashes (Skitka & Sargis, 2006). Although the computer science technologies have advanced exponentially, damaging viruses and similar problems continue to plague Internet users.

Conclusion

The advantages of Web-based surveys have been given much attention in the literature, with particular interest into the benefits to time and cost effectiveness. Other advatanges, such as the more precise data entry and coding with decreased chance of human error, increase the

potential reliability for data collection. The disadvantages of using the Internet as a research tool have also been reviewed and threaten the consistency and accuracy of Web-based measurements. One of the most important limitations applies to sampling, related to a lack of representation for certain subgroups or populations. The number of advantages, however, has been shown to out-weigh the disadvantages of moving away from pencil-and-paper surveys as the sole method of collecting survey data. In general, surveys tend to have strong reliability and weak validity estimates. Triangulation through the utilization of several e-surveying methods to include non-Web users is a way for researchers to overcome the limitations of Web-based survey methods, as well as help illuminate and eliminate the differences between traditional and Web-based survey methodologies. As the Internet becomes more of a societal norm, there will be an increased access to all groups leading to an improved avenue for data collection via Web-based surveys.

This paper addresses fundamental strengths and limitations associated with Web-based survey research. Given the prevalence of survey research in social and behavioral settings, especially in education, this review provides valuable information for those teaching survey methods or conducting said research. By reviewing the literature, a foundation is provided for sound methodological practices while attention is still drawn to the concerns which need to be addressed by researchers employing Web-based survey methods of data collection.

References

Andrews, D., Nonnecke, B., & Preece, J. (2003). Electronic survey methodology: A case study in reaching hard-to-involve Internet users. *International Journal of Human - Computer Interactions,* 16(2), 185-210.

Babbie, E. (2001). *The Practice of Social Research* (9th ed.) Belmont, CA: Wadsworth/Thompson Learning.

Birnbaum, M. H. (2004). Human research and data collection via the Internet. *Annual Review of Psychology, 55,* 803-832.

Buchmann, D. P., Elfrink, J., & Vazzana G. (2000). E-mail and snail mail face off in rematch. *Marketing Research,* 11(4), 10-15.

Cobanoglu, C., Warde, B., & Moreo, P.J. (2001). A comparison of mail, fax, and Web-based survey methods. *International Journal of Market Research, 43*(4), 441-452.

Daley, E., McDermott, R., McCormack-Brown, K., & Kittleson, M. (2003). Conducting Web-based survey research: A lesson in Internet designs. *American Journal of Health Behavior,* 27(2), 116-124.

Dillman, D. A. (2000). *Mail and Internet Surveys: The Tailored Design Method*. (2nd ed.) New York: John Wiley & Sons.

Eastin, M. S., & LaRose, R. (2000, September). Internet self-efficacy and the psychology of the digital divide. *Journal of Computer-Mediated Communication, 6*(1). Retrieved February, 1, 2006, from http://jcmc.indiana.edu/vol6/issue1/eastin.html

Edmunds, H. (1999). *The Focus Group Research Handbook.* Lincolnwood, IL: NTC. Business Books/Contemporary Publishing.

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50,* 537-567.

Ladner, M. D., Wingenbach, G. J., & Raven, M. R. (2002). Internet and paper based data collection methods in agricultural education research. *Journal of Southern Agricultural Education Research, 52*(1), 40-51.

Pew International and American Life Project. (2003, April 16). *The ever-shifting Internet population: A new look at Internet access and the digital divide*. Retrieved February 3, 2006, from http://www.pewinternet.org/report_display.asp?r=88

Litvin, S. W., & Kar, G. H. (2001). E-surveying for tourism research: Legitimate tool or researcher's fantasy? *Journal of Travel Research, 39*(3), 308.

Lyons, A. C., Cude, B., Lawrence, F. C., & Gutter, M. (2005). Conducting research online: Challenges facing researchers in family and consumer sciences. *Family and Consumer Sciences Research Journal, 33*(4), 341-356.

Nardi, P. M. (2003). *Doing survey research: A guide to quantitative methods*. New York: Pearson Education, Inc.

Rezabek, R. J. (2000, January). Online focus groups: Electronic discussions for research. *Forum: Qualitative Social Research, 1*(1). Retrieved October, 26, 2005, from http://qualitative-research.net/fqs

Skitka, L. J., & Sargis, E. G. (2006). The Internet as psychological laboratory. *Annual Review of Psychology, 57,* 529-555.

Swoboda, S. J., Muehlberger, N., Weitkunat, R., & Schneeweiss, S. (1997). Internet surveys by direct mailing: An innovative way of collecting data. *Social Science Computer Review, 15*(3), 197-205.

The University of Texas at Austin, Information Technology Services. (2005). *Online Surveys.* Retrieved January, 10, 2006, from http://www.utexas.edu/learn/surveys/index.html

Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology, 55,* 775-801.

University of Southern California, Annenberg School Center for the Digital Future. (2005). *Fifth study of the Internet by the digital future project finds major new trends in online use for political campaigns*. Retrieved February 5, 2006, from http://www.digitalcenter.org/pdf/Center-for- the-Digital-Future-2005-Highlights.pdf

Vereecken, C. (2001). Paper pencil versus PC administered querying of a study on health behaviour in school-aged children. *Archives of Public Health, 59*(1), 43-61.

Watt, J. H. (1999). Internet systems for evaluation research. In G. Gay and T.L. Bennington (Eds.), *Information Technologies in Evaluation: Social, Moral, Epistemological, and Practical Implications* (pp. 23-44). San Francisco: Jossey-Bass.

Wright, K. B. (2005) Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and Web survey services. *Journal of Computer-Mediated Communication, 10*(3). Retrieved February 5, 2006, from http://jcmc.indiana.edu/vol10/issue3/wright.html.

Yun, G. W., & Trumbo, C. W. (2000, September). Comparative response to a survey executed by post, e-mail, and Web form. *Journal of Computer-Mediated Communication,* 6(1). Retrieved February 6, 2006, from http://jcmc.indiana.edu/vol6/issue1/yun.html