

Should Educational Researchers Use Web-Based Surveys to Collect Sensitive Data?

Jay Tarby
 Robert A. Dubick
 Isadore Newman
 The University of Akron

The purpose of this study was to determine if a World Wide Web-based survey given to college students would elicit responses to sensitive questions that were different from those obtained via an equivalent mailed paper-and-pencil survey. Such a study was warranted on the grounds that, despite the growing popularity of web-based surveys in educational settings due primarily to cost savings and flexibility, very little has been done to verify the validity of the data obtained from such an approach. Moreover, studies by Erdman, Klein, and Greist (1983); Kiesler and Sproull (1986); Martin and Nagao (1989); and Lautenschlager and Flaherty (1990) and others have indicated that other forms of electronic surveys, especially when used to collect personal or sensitive information in educational settings, may produce responses that differ both quantitatively and qualitatively from those obtained from mailed paper-and-pencil surveys.

Methods

The population ($N = 1,967$) consisted of undergraduate students who lived in on-campus co-educational residence halls. From this population, a sample ($n = 1,115$) was derived consisting only of those students who had university-provided internet/e-mail accounts. Using a true experimental design adopted from Campbell & Stanley (See Figure 1), we randomly (R) divided this sample into two groups (X_1, X_2). One of these groups was sent a letter asking them to complete an attached paper version of Core Institute's Alcohol and Drug Survey and the other group was asked, via an e-mail message, to complete a web-based version of the same survey. After three weeks of data collection, which included one follow-up mailing for each group, chi-square tests were used to determine if there were significant differences (.05) between the two groups when looking at student responses (O_1) to sensitive questions.

Results

Of the 1,115 surveys sent out, 145 (13%) were returned. Of those returned, 63 (43%) came from the WWW group and 82 (57%) came from the mailed survey group. Although the raw response rates of the two survey methods were not significantly different at the $p < .05$ level ($p = 0.062$), the content of the responses given by the subjects to questions regarding their alcohol consumption and illegal drug use was. In looking at Table 1, for example, we can see that subjects who responded via the World Wide Web reported significantly lower levels of alcohol consumption at the $p < .05$ level than those responding via the paper survey on all items except the first. Likewise, in Table 2, we can see that those responding via the World Wide Web reported significantly lower levels of illegal drug use than those responding via the paper survey on all of the items except the last two.

Table 1

Comparative summary of students' self-reported data on the use of alcohol when collected using a web-based survey ($n = 63$) versus a paper survey ($n = 82$)

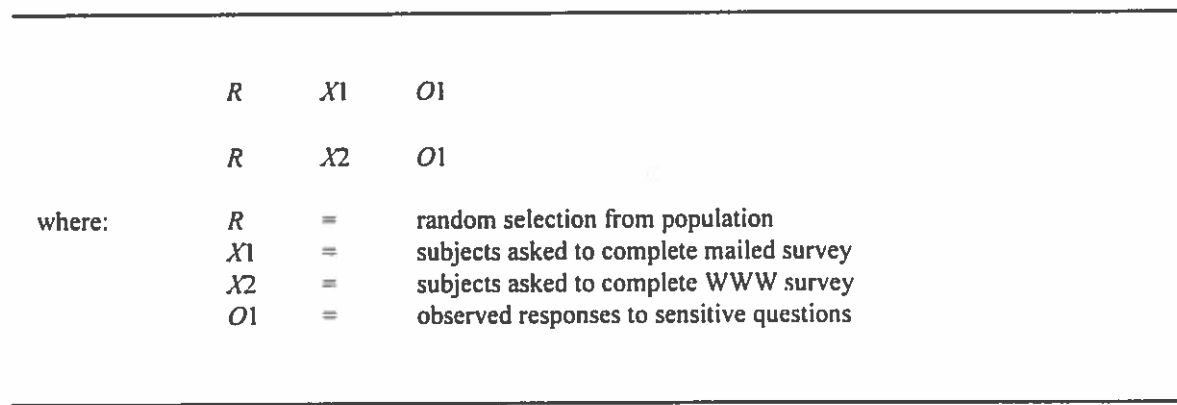
Item Summary	Method of survey delivery		
	Web	Paper	p
% reporting alcohol consumption in past year	59%	73%	.076
% reporting alcohol consumption in past 30 days	43%	63%	.017
% reporting alcohol consumption in past 30 days (underage)	36%	61%	.003
% reporting binge drinking in last 2 weeks	18%	38%	.009

Note: Binge drinking is defined as five or more drinks at a sitting.

Table 2
Comparative summary of students' self-reported data on the use of illegal drugs when collected using a Web-based survey (n = 63) versus a paper survey (n = 82)

Item Summary	Method of survey delivery		p
	Web	Paper	
% reporting marijuana usage in past year	3.3%	28%	.000
% reporting current marijuana usage	1.7%	18%	.002
% reporting usage of drug other than marijuana in past year	4.9%	11%	.187
% reporting current usage of drugs other than marijuana	0.0%	5.0%	.072

Figure 1. Diagram of research design adopted from Campbell & Stanley (1963).



Conclusions

The data reported here indicate that World Wide Web surveys may indeed produce results that are substantively different than those produced by paper surveys. One reason for this discrepancy may be that subjects taking the web-based survey introduced deliberate response biases (Dutka and Frankel, 1993). In other words, subjects may fear that their responses will be entered into personal records or will come to the attention of peers or supervisors and, as a result, may stretch the truth or engage in what Paulhus (1984) has referred to as impression management. Since impression management can lead to inaccurate and potentially misleading results, it should be of central concern to anyone conducting organizational survey research electronically. This is especially true of individuals interested in using Web-based surveys which are new and remain largely untested.

Recommendations

Although the results of this study suggest that World Wide Web surveys may elicit biased responses to sensitive questions in comparison to mailed surveys, researchers should keep in mind that there are a number of techniques that can be used with electronic surveys to minimize such biases. One such option would be to employ randomized response techniques (e.g., Boruch, 1972; Greenberg, 1969; Greenberg, Kuebler, Abernathy, and Horvitz, 1971; Warner, 1965) which, as described by Hosseini and Armacost, "...are probability-based methods that guarantee that an individual respondent cannot be identified with the response to a particular question" (1993, p. 445). Another option for those wanting to use Web-based surveys would be to employ a hybrid solution involving multiple methods of data collection. For example, one third of a population could be ap-

proached face-to-face, one third could be sent a paper questionnaire, and one third could be asked to complete a Web-based survey. The results of these three surveys could then be compared to determine if any response biases were introduced.

References

Boruch, R. F. (1972). Relations among statistical methods for assuring confidentiality of social research data. *Social Science Research, 1*, 403-414.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.

Dutka, S., & Frankel, L. R. (1993). Measurement errors in organizational surveys. *American Behavioral Scientist, 36*(4), 472-484.

Erdman, H., Klein, M. H., & Greist, J. H. (1983). The reliability of a computer interview for drug use/abuse information. *Behavior Research Methods and Instrumentation, 15*(1), 66-68.

Greenberg, B. G., Abul-Ela, A. A., Simmons, W. R., & Horvitz, D. G. (1969). The unrelated question randomized response model: Theoretical framework. *Journal of the American Statistical Association, 64*, 520-539.

Greenberg, B. G., Kruebler, R. R., Jr., Abernathy, J. R., & Horvitz, D. G. (1971). Application of the randomized response technique in obtaining quantitative data. *Journal of the American Statistical Association, 66*, 243-250.

Hosseini, J. C., & Armacost, R. L. (1993). Gathering sensitive data in organizations. *American Behavioral Scientist, 36*(4), 443-471.

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

Lautenschlager, G. J., & Flaherty, V. L. (1990). Computer administration of questions: More desirable or more social desirability? *Journal of Applied Psychology, 75*, 310-314.

Martin, C. L., & Nagao, D. H. (1989). Some effects of computerized interviewing on job applicant responses. *Journal of Applied Psychology, 74*, 72-80.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 598-609.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association, 60*, 63-69.

Jay Tarby is a doctoral student in the College of Education at The University of Akron, Akron, Ohio.

Robert A. Dubick is an Associate Professor in the College of Education at the University of Akron, Akron, Ohio.

Isadore Newman is a licensed psychologist and Distinguished Professor in the College of Education at the University of Akron, Akron, Ohio.