Comparing web and mail responses in a mixed mode survey in college alcohol use research

Sean Esteban McCabe[a,*], **Alison Diez**[b], **Carol J. Boyd**[c], **Toben F. Nelson**[d], and **Elissa R. Weitzman**[d]

[a] Substance Abuse Research Center, The University of Michigan, 2025 Traverwood, Suite C, Ann Arbor, Michigan 48105-2194, United States

[b] Government, Foundation, and Academic Research Division, MSI, Livonia, MI, United States

[c] Substance Abuse Research Center, Institute for Research on Women and Gender, The University of Michigan, Ann Arbor, MI, United States

[d] Society, Human Development and Health, Harvard School of Public Health, Boston, MA, United States

Abstract

Objective—This exploratory study examined potential mode effects (web versus U.S. mail) in a mixed mode design survey of alcohol use at eight U.S. colleges.

Methods—Randomly selected students from eight U.S. colleges were invited to participate in a self-administered survey on their alcohol use in the spring of 2002. Data were collected initially by web survey (*n* =2619) and non-responders to this mode were mailed a hardcopy survey (*n* =628).

Results—College students who were male, living on-campus and under 21 years of age were significantly more likely to complete the initial web survey. Multivariate analyses revealed few substantive differences between survey modality and alcohol use measures.

Conclusions—The findings from this study provide preliminary evidence that web and mail surveys produce comparable estimates of alcohol use in a non-randomized mixed mode design. The results suggest that mixed mode survey designs could be effective at reaching certain college sub-populations and improving overall response rate while maintaining valid measurement of alcohol use. Web surveys are gaining popularity in survey research and more work is needed to examine whether these results can extend to web surveys generally or are specific to mixed mode designs.

Keywords

Survey mode; Web; Mail; College students; Alcohol use

1. Introduction

Web surveys are a viable mode of data collection on many college campuses because of the near universal use of the Internet among college students (Couper, 2000; Jones, 2003; Rainie, 2001). With the growing popularity of web surveys, it has become more important to compare this innovative mode of data collection to other, more traditional, survey methods. Therefore, the impact of collecting data among undergraduate students using web surveys

*Corresponding author. Tel.: +1 734 998 6510; fax: +1 734 998 6508. plius@umich.edu (S.E. McCabe).

versus traditional approaches such as mailed paper surveys has become a topic of college-based research (e.g. Carini, Hayek, Kuh, Kennedy, & Ouimet, 2003; Pealer, Weiler, Pigg, Miller, & Dorman, 2001; Wygant & Lindorf, 1999).

Past research suggests that data collection modality can lead to substantially different answers to questions regarding alcohol and other drug use (e.g. Link & Mokdad, 2005; Turner et al., 1998; Wright, Aquilino, & Supple, 1998). However, at least three randomized college-based studies suggest minimal differences in the reporting of alcohol use between web surveys and more traditional paper-based survey approaches (e.g., Bason, 2000; McCabe, Boyd, Couper, Crawford, & d'Arcy, 2002; Miller et al., 2002). For instance, Miller et al. (2002) randomly assigned 255 college students ages 18–29 to complete a survey in one of three conditions, including one paper-based and two web-based conditions. The survey contained several well-known alcohol-related measures such as the Alcohol Use Disorders Identification Test (AUDIT: Saunders, Aasland, Babor, de la Fuente, & Grant, 1993), Rutgers Alcohol Problem Index (RAPI: White & Labouvie, 1989), University of Rhode Island Change Assessment (URICA: Prochaska & DiClemente, 1986), Alcohol Dependence Scale (ADS: Skinner & Allen, 1982), and a quantity-frequency-peak index used to assess drinking rates. Miller et al. (2002) found there were no significant mean differences among the assessment conditions on any measures of alcohol use. The authors concluded that the web-based survey mode of data collection represented a suitable alternative to a paper-based approach because data integrity was not compromised. In another study, Bason (2000) randomly assigned 3000 students to complete the CORE survey (Presley, Meilman, & Cashin, 1996) between four modes of data collection (telephone, mail, web-based and interactive voice recognition) and examined substantive differences in prevalence estimates of alcohol and other drug use. Although the web mode produced the lowest response rate of the four survey modes, there were minimal substantive differences in prevalence rates of substance use between web respondents and other modes. Finally, McCabe et al. (2002) randomly assigned 7000 students to complete a survey via two distinct survey modes (web and mail). The web mode produced a sample that was more representative of the overall student population with respect to gender and there were no significant differences in alcohol use measures. In addition, there were no differences in the distribution of race, class year, academic credit hours, or age between the final samples obtained from the two survey modes.

Despite evidence that web surveys can produce comparable sample demographics and estimates of substance use as mailed paper surveys when college students are randomly assigned to survey modes (e.g. McCabe et al., 2002; Miller et al., 2002), there remains a need for comparative research that examines the impact of using a non-randomized mixed mode design on sample demographics and substance use estimates. Non-randomized mixed mode studies, including web and mailed paper surveys have been conducted at multiple universities, and very few substantive differences between survey mode have been found (e.g. Carini et al., 2003). However, such studies have not examined highly sensitive behaviors such as substance use.

There are several different types of web surveys and the current investigation focuses on a probability-based survey approach within a mixed mode design (Couper, 2000). The main objective of this study was to examine whether demographic characteristics and alcohol use data collected initially using a web survey as the first mode of data collection differed from a subsequent mailed paper survey as the second mode of data collection.

2. Methods

2.1. Data collection

The present study investigated potential mode effects in self-reported measures of alcohol consumption using survey data collected as part of the "A Matter of Degree" (AMOD) national prevention demonstration program evaluation (Weitzman, Nelson, Lee, & Wechsler, 2004). AMOD is a comprehensive community change program that aims to alter patterns of heavy and harmful drinking among college youth through environmental change strategies. Survey data from eight AMOD colleges were used for the present mode effects study. AMOD schools were selected because of their high rates of binge drinking (i.e., >50%) in 1993, prior to the program's inception, and their commitment to ongoing participation in an environmental prevention program for their respective campus communities. Details about the AMOD program, site selection, survey methods and findings are published elsewhere (Weitzman et al., 2004).

The study was conducted during a one-month period during the spring of 2002. Potential participants were sent information making it clear that participation was voluntary, explaining the relevance of the study, and that responses were kept confidential. The wording, response categories and skip patterns for survey questions in each mode were identical. Potential participants were informed that a research firm was contracted to set-up the web survey as well as store and maintain data from both modes of data collection. Finally, the web survey was maintained on a hosted secure web site running under the secure sockets layer (SSL) protocol to insure respondent data was safely transmitted between the respondent's browser and the server. The Harvard School of Public Health's Human Subjects Committee approved the study's protocol.

Administrators at each of the schools provided a random sample of 750 undergraduate students from the total population of students enrolled full-time and one school provided a sample of 1250 students. The original sample included nine schools but one school was excluded from this investigation because the response rate was considerably lower than the other eight schools. Two of the universities were located in the Northeast, three in the South, two in the North Central and one in the West; seven were public schools and one was a private school. Five of the schools were large (over 10000 undergraduate students), two were medium (5001 to 10000 undergraduate students), and one was small (under 5000 undergraduate students). Six schools were located in an urban or suburban setting and two in small town/rural settings.

The present study was not designed to be a randomized mode experiment with subjects randomly assigned to two different modes of data collection. A web survey was used as the first mode of data collection followed by a mailed paper survey mode. In the web mode, e-mail invitations were sent to 6500 full-time undergraduate students attending eight different colleges and universities inviting them to participate in the study by clicking a URL link and self-administering the web survey. Following the initial e-mail invitation, students were sent a first and/or second e-mail reminder if they had not completed the survey. All of the 3880 students who did not respond to the web survey were mailed hardcopies of survey questionnaires along with a cover letter inviting them to respond. Reminder post cards were also sent to all potential respondents. Additional cover letters and hardcopies of the survey questionnaires were sent to students who did not respond following the postcard reminder.

Due to the variation in academic calendars between the eight schools in the present study, four slightly different data collection schedules were used during April and May of 2002 that ranged from 19 days to 26 days. By participating in the study, students from each school became eligible for a sweepstakes drawing of $500 at each school. The response rates for

the web survey mode of data collection ranged as follows: (School A=46.1%, School B=42.6%, School C=39.2%, School D=37.2%, School E=33.1%, School F=30.9%, School G=30.7%, School H=14.8%). The response rates for the mail survey mode of data collection ranged as follows: (School A=10.8%, School B=16.2%, School C=41.0%, School D=23.6%, School E= 26.1%, School F= 31.2%, School G= 25.3%, School H= 42.2%). A total of 3247 students returned questionnaires from the eight schools resulting in total response rates from 46.3% to 61.9%.

2.2. Measures

The present study used data from the Harvard School of Public Health College Alcohol Study (CAS) survey. The survey instrument contained questions regarding demographic characteristics such as gender, race, age, and fraternity/sorority membership as well as several alcohol use measures.

Living arrangement was determined by whether students lived on-campus or off-campus during the current school year. For purposes of these analyses, on-campus residence consisted of living in a residence hall, other university housing, or a fraternity or sorority house. Off-campus residences consisted of living in a house, apartment, or other off-campus locations during the school year.

Abstinence was defined as not consuming alcohol in the past year.

Alcohol use in the past 30 days was assessed with a single item asking "On how many occasions have you had a drink of alcohol in the past 30 days?" The response scale was 1) 1 to 2 occasions, 2) 3 to 5 occasions, 3) 6 to 9 occasions, 4) 10 to 19 occasions, 5) 20 to 39 occasions, and 6) 40 or more occasions. For purposes of these analyses, responses were dichotomously coded as 1) less than ten occasions in the past 30 days and 2) ten or more occasions in the past 30 days.

Monthly drunk occasions were assessed by asking "In the past 30 days, how often did you drink enough to get drunk?" The response scale was 1) not at all, 2) 1 to 2 occasions, 3) 3 to 5 occasions, 4) 6 to 9 occasions, 5) 10 to 19 occasions, 6) 20 to 39 occasions, and 7) 40 or more occasions. For purposes of these analyses, responses were dichotomously coded as 1) less than three occasions and 2) three or more occasions.

Binge drinking was defined as having five or more drinks in a row for men (and four or more drinks in a row for women) in the past two weeks. A drink was defined as a 4-ounce glass of wine, a 12-ounce bottle or can of beer or wine cooler, or a shot of liquor straight or in a mixed drink. The response scale was 1) none, 2) once, 3) twice, 4) 3 to 5 times, 5) 6 to 9 times, and 6) 10 or more times. For purposes of these analyses, responses were dichotomously coded as 1) none and 2) at least once.

Frequent binge drinking was defined as having three or more binge drinking episodes in the past two weeks.

Usual heavy drinking behavior was also assessed by asking respondents who had consumed alcohol in the past 30 days the following question: "In the past 30 days, on those occasions when you drank alcohol, how many drinks did you usually have?" The response scale ranged from 1) 1 drink to 9) 9 or more drinks. For purposes of these analyses, responses were dichotomously coded as 1) usual non-heavy drinking (less than 4 drinks for women and less than 5 drinks for men) and 2) usual heavy drinking (4 or more drinks for women and 5 or more drinks for men).

Drink to get drunk was assessed with one item that asked students how important drinking "to get drunk" was as a reason to drink alcohol. The response scale was 1) very important, 2) important, 3) somewhat important, and 4) not at all important. For purposes of these analyses, responses were dichotomously coded as 1) not important at all and 2) very important, important, and somewhat important.

2.3. Data analysis

A sample weight variable was created to account for colleges' varying sampling fractions and data were weighted throughout all analyses to increase the representativeness of our results. The weight variable was centered (normalized) within each study site to insure that sample sizes for all sites remained the same after weighting.

Chi-square tests of homogeneity were conducted to determine whether the respondents' distributions for demographic variables differed significantly between survey modality. Next, multivariate logistic regression analyses were conducted to estimate survey mode effects on several dichotomously coded variables for alcohol use measures. The multivariate logistic regression analyses tested the null hypothesis that there were no significant differences in substantive responses to each alcohol use measure between survey modes while adjusting for possible confounding variables including gender, age, and living arrangement. Separate logistic regression models were conducted for all eight schools in aggregate and for each of the eight schools separately.

2.4. Sample

A total of 3247 students responded, 628 by U.S. mail and 2619 by web. As illustrated in Table 1, the overall sample consisted of 52.8% male, 85.2% White, and 54.2% of students were under 21 years of age. In addition, 19.4% of students were members of fraternities or sororities and 52.0% lived on-campus which included residence halls, other university housing, and fraternity and sorority houses.

3. Results

As illustrated in Table 1, there were several demographic characteristics that differed significantly by survey modality. For instance, web respondents were more likely to be male (55.5% vs. 41.3% of mail respondents, $p < 0.001$), live on-campus (56.5% vs. 33.1% of mail respondents, $p < 0.001$) and tended to be less than 21 years of age (58.4% vs. 37.1% of mail respondents, $p < 0.001$).

Multivariate logistic regression analyses in the aggregate across all eight schools were conducted to examine the impact of survey modality on substantive responses to alcohol use after adjusting for sex, age and living location. As presented in Table 2, logistic regression models examining alcohol use measures revealed no substantive differences by survey mode across several measures including 12-month abstinence, drinking occasions in the past 30 days, three or more occasions of being drunk in the past 30 days, usual heavy drinking in the past 30 days, importance of getting drunk, binge drinking in the past two weeks and frequent binge drinking in the past two weeks.

The regression models were also conducted separately at each of the eight schools (results not shown). Although there were some instances where the sample size from individual schools was too small to permit reliable results using multivariate analysis, the results for individual schools indicated the same trend as the aggregate findings. With very few exceptions, there were no individual school differences between survey modes for reporting 12-month abstinence, drinking occasions in the past 30 days, three or more occasions of being drunk in the past 30 days, usual heavy drinking in the past 30 days, importance of

getting drunk, binge drinking in the past two weeks and frequent binge drinking in the past two weeks.

4. Discussion

While relying exclusively on the web for data collection within the general population might not yet be possible, web surveys have shown promise for restricted populations with near universal web access such as college students (e.g. Carini et al., 2003; Kypri & Gallagher, 2003; McCabe et al., 2002). The results of the present study provide preliminary evidence from multiple institutions that web surveys produce few substantive differences in estimates of alcohol use when college students complete web surveys as the first mode of data collection followed by a mailed hardcopy survey as the second mode of data collection. These findings complement the growing evidence suggesting web and paper surveys produce comparable results regarding alcohol use based on randomized college-based studies (e.g. Bason, 2000; McCabe et al., 2002; Miller et al., 2002). Web surveys and mailed paper survey modes share several important characteristics so it is not surprising that several randomized college-based studies have found that responses to web surveys closely resemble the responses of hardcopy mailed paper surveys. For example, both survey modes are self-administered and they each rely on visual information through which the respondent controls the speed of answering questions.

The present study found that undergraduate men were more likely than undergraduate women to respond to the web survey and less likely to respond to the subsequent mail survey. These gender differences in response patterns between web and subsequent mail surveys are consistent with a non-randomized study (Carini et al., 2003) and a randomized mode study (McCabe et al., 2002) within undergraduate student populations. The present study showed that students living on-campus were more likely to respond to a web survey and less likely to respond to the following mail survey relative to students living off-campus. This could be a function of differences in computer access and use between students living on-campus versus those living off-campus.

The results of the present study are important for researchers considering whether to use mixed-mode data collection strategies within college student populations. Past research has shown that switching to a second mode of data collection increases survey response rates (e.g. Shettle & Mooney, 1999). Because of the cost savings when web surveys are the first mode of data collection, web surveys have a great deal of promise for conducting large-scale studies using mixed mode data collection strategies.

4.1. Limitations

Despite the strengths of the current study, there were some noteworthy limitations that need to be taken into account when considering the implications of the findings. First, the non-randomized design of the present study may have led to some selection effects among responders. For example, the mail survey responders could contain college students who are late or reluctant responders and not necessarily representative of mail survey responders in general. Although we attempted to minimize these selection effects by controlling for demographic characteristics that differed between the two survey modes, there could be other selective factors operating in a group that does not respond to one mode of the survey. Second, non-response may have introduced some bias in the present study. There was variation in the response rates between individual schools suggesting web-based approaches might be more feasible at some institutions. Non-response remains one of the main concerns for web surveys and more work is needed to increase web survey response rates (Couper, 2000). Finally, the eight schools used in the present study are not representative of American colleges and universities as the schools were selected because of their high rates of binge

drinking in 1993 and their ongoing participation in an intervention to reduce alcohol abuse on their respective campuses (Weitzman et al., 2004). Therefore, more comparative research is needed at colleges and universities with diverse binge drinking profiles.

4.2. Future research

Future research would benefit from evaluating how various contact strategies, within different mode designs, can produce higher response rates. Recent research indicates pre-notification, pre-paid incentives, intensive follow-up, and mixed modes of data collection featuring web, phone, and mail can achieve high response rates among college students (e.g., Kypri & Gallagher, 2003). However, the use of web surveys needs to be carefully considered at individual schools based on the mixed results of the present study and several past studies (e.g. Bason, 2000; Kwak & Radler, 2002; McCabe et al., 2002; Wygant & Lindorf, 1999). Nearly all (95%) of the students in the eight schools used in the present study were 18–24 years of age rendering the web an especially practical mode of data collection as over three out of every four individuals within this age group report online computer use (Rainie, 2001). In addition, previous research has shown approximately 98% of traditional-age (18–24 years old) undergraduate students used electronic mail (Couper, Traugott, & Lamias, 2001). However, it remains to be shown if web surveys can be effective at schools with higher proportions of non-traditional age students. There is also relatively little known about the institutional characteristics that impact the efficacy of web surveys at individual universities and this represents an important area of future research. Future research could greatly benefit from conducting a multi-campus mode experiment with randomized conditions to test what types of characteristics improve response rates. Given the variation in institutional characteristics across different colleges and universities, it is plausible that various schools might require different mixed mode combinations depending on certain conditions.

Acknowledgments

References

Bason, JJ. Comparison of telephone, mail, web, and IVR surveys of drug and alcohol use among university of Georgia students. Paper presented at the American Association of Public Opinion Research; Portland, Oregon. 2000.

Carini RM, Hayek JH, Kuh GD, Kennedy JM, Ouimet JA. College student responses to web and paper surveys: Does mode matter? Research in Higher Education. 2003; 44:1–19.

Couper MP. Web surveys: A review of issues and approaches. Public Opinion Quarterly. 2000; 64:464–494. [PubMed: 11171027]

Couper MP, Traugott M, Lamias M. Web survey design and administration. Public Opinion Quarterly. 2001; 65:230–253. [PubMed: 11420757]

Jones, S. The Internet Goes to College. Pew Internet and American Life Project report [Online]. 2003 [accessed Feb 2003]. Available: http://www.pewinternet.org/reports/pdfs/PIP_College_Report.pdf

Kwak N, Radler B. A comparison between mail and web surveys: Response pattern, respondent profile, and data quality. Journal of Official Statistics. 2002; 18:257–273.

Kypri K, Gallagher SJ. Incentives to increase participation in an internet survey of alcohol use: A controlled experiment. Alcohol and Alcoholism. 2003; 38:437–441. [PubMed: 12915520]

Link MW, Mokdad AH. Effects of survey mode on self-reports of adult alcohol consumption: Comparison of web, mail and telephone approaches. Journal of Studies on Alcohol. 2005; 66:239–245. [PubMed: 15957675]

McCabe SE, Boyd CJ, Couper MP, Crawford S, d'Arcy H. Mode effects for collecting alcohol and other drug use data: Web and U.S. mail. Journal of Studies on Alcohol. 2002; 63:755–761. [PubMed: 12529076]

Miller ET, Neal DJ, Roberts LJ, Baer JS, Cressler SO, Metrik J, et al. Test–retest reliability of alcohol measures: Is there a difference between internet-based assessment and traditional methods? Psychology of Addictive Behaviors. 2002; 16:56–63. [PubMed: 11934087]

Pealer L, Weiler RM, Pigg RM, Miller D, Dorman SM. The feasibility of a web-based surveillance system to collect health risk behavior data from college students. Health Education Behavior. 2001; 28:547–559. [PubMed: 11575685]

Presley, CA.; Meilman, PW.; Cashin, JR. Alcohol and drugs on American college campuses: Use, consequences, and perceptions of the campus environment, volume IV: 1992–94. Carbondale, IL: Core Institute, Southern Illinois University; 1996.

Prochaska, JO.; DiClemente, CC. Toward a comprehensive model of change. In: Miller, WR.; Heather, N., editors. Treating addictive behaviors: Processes of change. Applied clinical psychology. New York: Plenum; 1986. p. 3-27.

Rainie, H. The changing online population: It's more and more like the general population. [Online]. Pew Internet and American Life Project report. Available: http://www.pewinternet.org/reports/index.asp[2001, October, 9]

Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption: II. Addiction. 1993; 88:791–804. [PubMed: 8329970]

Shettle C, Mooney G. Monetary incentives in U.S. government surveys. Journal of Official Statistics. 1999; 15:231–250.

Skinner HA, Allen BA. Alcohol dependence syndrome: measurement and validation. Journal of Abnormal Psychology. 1982; 91:199–209. [PubMed: 7096790]

Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. Science. 1998; 280:867–873. [PubMed: 9572724]

Weitzman ER, Nelson TF, Lee H, Wechsler H. Reducing drinking and related harms in college: Evaluation of the "A Matter of Degree" program. American Journal of Preventive Medicine. 2004; 27:187–196. [PubMed: 15450630]

White HR, Labouvie EW. Toward the assessment of adolescent problem drinking. Journal of Studies on Alcohol. 1989; 50:30–37. [PubMed: 2927120]

Wright DL, Aquilino WS, Supple AI. A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol, and other drug use. Public Opinion Quarterly. 1998; 62:331–353.

Wygant, S.; Lindorf, R. Surveying collegiate net surfers–Web methodology or mythology [Online]. Quirk's marketing research review. 1999. Available: http://www.quirks.com/articles/article.asp?arg_ArticleId=515 [2002, May, 24]

Table 1

Sample characteristics: overall population and respondents to each survey modality

Sample characteristic	Total (*n* =3247) *n*	%	Web (*n* =2619) %	Mail (*n* =628) %	χ^2 (*df*), *p*-value
Sex					
Male	1709	52.8	55.5	41.3	40.5 (1)
Female	1530	47.2	44.5	58.7	<0.001
Race					
White	2220	85.2	85.1	85.3	0.2 (1)
Non-white	386	14.8	14.9	14.7	0.899
Living arrangement					
On-campus	1678	52.0	56.5	33.1	110.3 (1)
Off-campus	1548	48.0	43.5	66.9	<0.001
Fraternity/sorority membership					
Non-member	2599	80.6	80.2	82.2	1.3 (1)
Member	626	19.4	19.8	17.8	0.251
Age (in years)					
Under 21	1759	54.2	58.4	37.1	96.1 (1)
21–23	1314	40.6	37.3	54.3	<0.001
24 or older	167	5.2	4.3	8.7	

Chi-square tests indicate whether each sample characteristic differs significantly by survey modality.

Table 2

Logistic regression results of survey mode effects for alcohol use measures

	Web %	Mail %	Adjusted OR[a,b]	95% Confidence interval
Past 12 months				
Abstained	12.5	10.0	1.05	0.77 – 1.41
Past 30 days				
Was drunk 3 or more times	32.1	31.1	0.99	0.81 – 1.21
Used alcohol 10 or more times	24.1	26.9	0.87	0.70 – 1.07
Drink to get drunk[c]	63.1	57.9	1.13	0.91 – 1.40
Usual heavy drinking	41.2	42.6	0.92	0.75 – 1.13
Past 2 weeks				
Binge drinking	60.3	64.2	0.87	0.72 – 1.06
Frequent binge drinking	35.8	38.5	0.90	0.74 – 1.09

There were no statistically significant differences ($p < 0.05$) for any of the alcohol use measures.

[a] The mail survey was the reference group so the adjusted odds ratios reflect the relative odds that a respondent to the Web survey would report each alcohol measure.

[b] Regression model was adjusted for sex, living arrangement, and age.

[c] Regression model for "drink to get drunk" includes only those who drank within the past 30 days ($n = 2173$).