

# Attrition Bias in a U.S. Internet Survey of Alcohol Use Among College Freshmen\*

THOMAS P. MCCOY, M.S.,<sup>†</sup> EDWARD H. IP, PH.D., JILL N. BLOCKER, M.S., HEATHER CHAMPION, PH.D.,<sup>†</sup> SCOTT D. RHODES, PH.D.,<sup>†</sup> KIMBERLY G. WAGONER, M.P.H.,<sup>†</sup> ANANDA MITRA, PH.D.,<sup>†</sup> AND MARK WOLFSON, PH.D.<sup>†</sup>

*Department of Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157*

**ABSTRACT. Objective:** Attrition bias is an important issue in survey research on alcohol, tobacco, and other drug use. The issue is even more salient for Internet studies, because these studies often have higher rates of attrition than face-to-face or telephone surveys, and there is limited research examining the issue in the field of drug usage, specifically for college underclassmen. This study assessed whether measures of high-risk drinking and alcohol-related consequences were related to attrition groups (“stayers” or “leavers”) in a cohort of college freshmen. **Method:** Data were collected in 2003 and 2004 from 2,144 first-year college students at 10 universities in the southeastern United States. Demographics, indicators of high-risk drinking, and alcohol-related consequences were

compared between cohort stayers and leavers in statistical analyses using two methods. **Results:** Analyses indicated that cohort leavers reported significantly higher levels of high-risk drinking (past-30-day heavy episodic drinking, weekly drunkenness) and past-30-day smoking but not significantly increased alcohol-related consequences. The directionality of bias was modestly consistent across outcomes and comparison methods. **Conclusions:** The current study’s findings suggest that intervention efforts to reduce smoking or high-risk drinking need to consider attrition bias during study follow-up or account for it in analyses. (*J. Stud. Alcohol Drugs* 70: 606-614, 2009)

THE SURVEY-BASED METHOD is one of the most commonly used approaches for assessing the prevalence of high-risk drinking and related behaviors and the consequences of alcohol consumption (McCabe et al., 2006). Although new media such as the Internet and cellular phones have created exciting opportunities for researchers to reach specific populations in cost-effective and efficient ways (Comley, 2000; Rhodes et al., 2003), old and persistent problems in cohort studies, such as lower follow-up rates and attrition bias, remain a concern. Attrition bias is an important issue in cohort studies of college alcohol use—and potentially even more so in Internet studies—because in this setting greater rates of attrition typically exist (although this does not necessarily imply there is an attrition bias).

This study focuses on attrition bias using data from a cohort of college freshmen concerning their alcohol use and related consequences. Attrition is defined as loss of participants (and their subsequent data) during follow-up (Bailey et al., 1992; Cox et al., 1998; Grekin et al., 2007; Jessor and Jessor, 1975; Locke and Newcomb, 2003; Paschall and Freisthlet, 2003; Prescott and Kendler, 2001). Definitions of attrition bias have differed slightly in previous literature:

Miller and Wright (1995) state that “attrition results in a potential threat of bias if those who drop out have unique characteristics such that remaining sample ceases to be representative of the original sample” (p. 921).

This definition echoes that of Cuddeback et al. (2004): “Selected individuals may agree to participate but then be ‘lost’ over time to transience, incarceration, death, or other reasons. The final sample might be biased if the individuals who are lost differ in some systematic way from the participants who remain. This is known as attrition bias” (p. 20). Miller and Hollist (2007) agreed with this theme, commenting that “attrition of the original sample represents a potential threat of bias if those who drop out of the study are systematically different from those who remain in the study. The result is that the remaining sample becomes different from the original sample, resulting in what is known as attrition bias” (p. 57). McGuigan et al. (1997) remark that “if attrition is systematically related to outcomes of interest and if nonresponse adjustments are made, bias may result” (p. 554).

Researchers should be cognizant of the difference between attrition and attrition bias. Attrition does not neces-

Received: July 30, 2008. Revision: December 11, 2008.

\*This research was supported by National Institute on Alcohol Abuse and Alcoholism (NIAAA) grant R01 AA14007 and by funds from the Division of Mental Health, Developmental Disabilities and Substance Abuse Service of the North Carolina Department of Health and Human Services; the U.S. Office of Juvenile Justice and Delinquency Prevention through the Enforcing Underage Drinking Laws program; and Wake Forest University Interim Funding. The content is solely the responsibility of the authors and does

not necessarily represent the official views of the NIAAA or the National Institutes of Health.

<sup>†</sup>Correspondence may be sent to Thomas P. McCoy at the above address or via email at: tmccoy@wfubmc.edu. Heather Champion, Scott D. Rhodes, Kimberly G. Wagoner, and Mark Wolfson are with the Department of Social Sciences and Health Policy, Wake Forest University School of Medicine, Winston-Salem, NC. Ananda Mitra is with the Department of Communication, Wake Forest University, Winston-Salem, NC.

sarily imply there is attrition bias, even if attrition is high (Outes-Leon and Dercon, 2008). From a statistical perspective, high attrition rate and hence smaller sample size leads to a decrease in the precision of estimates of population parameters (e.g., because sample size is in the denominator of the formula for standard error used in confidence intervals) and lower statistical power to detect associations.

This issue of decreased precision is distinct from the issue of attrition bias. Although a high attrition rate may suggest potential problems in the design of the survey, it does not necessarily imply the presence of attrition bias, which occurs when there is a systematic difference between “stayers” (those who remain in the cohort sample) and “leavers” (those who drop out of the cohort and subsequently have missing data) in terms of the outcome variables of interest, regardless of the magnitude of the follow-up rate.

Attrition bias threatens results and conclusions from alcohol studies on college students for many reasons. As pointed out by Miller and Hollist (2007), attrition bias can affect the external and internal validity of the study. They consider the external validity of the study in danger if the composition of the sample changes to the point (from leavers dropping out) that results are no longer generalizable to the original study population. The internal validity of the study can be affected if associations among study variables are substantially affected or if differential dropout in a treatment group design is observed.

Several approaches for assessing attrition bias have been used in prior studies. These strategies can be grouped into two categories: ones for detecting attrition bias and ones that attempt to correct or account for attrition bias in analyses. For detecting attrition bias, previous studies have used the following: (1) bivariate comparisons of baseline data for stayers versus leavers, (2) multivariable comparisons of baseline data for stayers versus leavers, (3) comparing individual correlation coefficients from variables using the entire overall sample at baseline with those using the longitudinal sample of only stayers using Fisher's  $z$  tests, and (4) comparing the correlation matrices of the entire overall sample at baseline with the longitudinal sample of only stayers using a test of invariance (Miller and Wright, 1995).

Tests of invariance for entire correlation matrices seem like an improvement over individual tests for correlation coefficients in terms of multiplicity and Type I error considerations (i.e., one  $p$  value for invariance test vs many individual  $p$  values from different correlations). Cuddeback et al. (2004) remark that bivariate comparisons of stayers and leavers cannot be used to estimate the independent effects of variables used to determine differences between stayers and leavers on outcomes but, instead, that multivariable comparisons better accomplish this.

Bivariate comparisons of stayers versus leavers on baseline data appear to be one of the most common strategies for detecting attrition group differences (Bailey et al., 1992;

Boys et al., 2003; Cox et al., 1998; Miller and Wright, 1995; Prescott and Kendler, 2001; Smith et al., 1995; White et al., 2006). The drawback of this approach is that the same comparisons on the missing follow-up data could result in different conclusions, but unfortunately this is not checkable. One could also investigate how consistent (correlated) stayers' data are from baseline to follow-up in an attempt to validate baseline comparisons, but this may or may not generalize to leavers and could be confounded with intervention group in an intervention study design setting.

For explicitly accounting for attrition bias, previous studies have explored the following: (1) using logistic regression where attrition group (stayer, leaver) is the outcome and compute weights from predicted probabilities to use in subsequent weighted outcome analyses (McGuigan et al., 1997), (2) selection modeling where a probit model for attrition and a regression model for the outcome are used, with focus on the test of correlated error terms between the two models for a significance test of attrition bias (McGuigan et al., 1997), (3) comparing analyses based on all available data using maximum likelihood methods with those with complete-case data (stayers only), (4) using multiple imputation to first impute missing follow-up data from leavers and then analyze the imputed data (Schafer and Graham, 2002).

Martino et al. (2005) used weighting created from modeling attrition groups and commented that weighting in analysis of outcomes had removed “more than 90% of the bias caused by attrition” (p. 141). In a study by McGuigan et al. (1997), they concluded that weighting performed better than selection modeling but only because the underlying assumptions of selection modeling were not satisfied. However, Schafer and Graham (2002) propose that using all available data from stayers and leavers in a maximum likelihood methods-based analysis or using multiple imputation is the practical state of the art for missing data. They conclude that either of these two is better than weighting but do not consider selection modeling.

One assumption in selection modeling is that the model for attrition is not misspecified (no omitted predictors, interactions, or higher order terms). Selection modeling also assumes selection on observables, which is equivalent to the “missing at random” (MAR) assumption (Fitzgerald et al., 1998). The MAR assumption is the same assumption that underlies methods based on maximum likelihood with all available data and using multiple imputation as well. The literature on these methods is rich, especially in economics for selection models and in statistics for missing data analyses assuming MAR.

Data that are missing at random assume that missing data (the value of the outcome at follow-up that is missing) can depend on observed data (covariates, previous follow-up values of outcome, baseline value of outcome) but not on the missing outcome itself. For example, the MAR assumption would hold if a student's eventual attrition group status

at follow-up was dependent on his or her outcome value at baseline or measured demographics but not dependent on the actual outcome value at follow-up itself (which is missing).

If this dependency on missing outcome does occur, this is known as *informative dropout*. The difficulty in attempting to check the MAR assumption is that one cannot check if attrition grouping is dependent on the outcome value at follow-up because the outcome value at follow-up was not observed. However, Collins et al. (2001) showed that non-MAR data that assume MAR may have a smaller impact on results. Schafer and Graham (2002) reiterate that in some practical applications, departures from MAR may not be large enough to invalidate results from analysis that assumes MAR.

Another strategy outside of statistical analysis of attrition is to conduct follow-up studies on leavers, often using alternative modes of contact. For example, leavers could be contacted by phone, residence visit, or mail rather than email as in the initial phase of the study. The limitations of this method are the following: (1) follow-up quite often generates only a low yield, (2) differences in responses could be attributable to variation in survey mode, (3) the approach does not offer direct information about individuals who still do not ultimately respond, even after extensive efforts (e.g., after expending increased incentives), and (4) often there is an increased cost of tracking down leavers, which could be more costly than offering an increased incentive up front (Boys et al., 2003).

Although there have been several studies that investigated attrition bias among young adults (Martino et al., 2005; White and Widom, 2008) and adolescents (Smith et al., 1995), fewer studies have examined attrition bias in college students (White et al., 2006). In the present study, we employed two methods to study attrition bias in an Internet survey of U.S. college students' alcohol consumption. We focused on bias that resulted from attrition in a cohort study specifically of college freshmen. Thus, this study did not consider nonresponse bias as a result of baseline nonresponders.

Although large-scale surveys of alcohol use in the United States often collect repeated cross-sectional data, the data set for this study comprised two years of cohort survey data (a baseline year in 2003 and a follow-up year in 2004). We used cohort data to provide two comparisons in exploring potential differences between cohort stayers and leavers (Menard, 2002). The first comparison retrospectively examined differences between stayers and leavers on baseline data from 2003 (Figure 1: Method I). The second comparison used all available data with a maximum-likelihood approach that contrasts baseline data from leavers to follow-up data from stayers, after adjusting for stayers' baseline data (Figure 1: Method II).

Our goal was to assess attrition bias using previously collected response data from cohort leavers using two comparisons. The comparisons were based on these explicit assumptions: In the first comparison, we assumed that baseline data for stayers and leavers would provide valid comparisons, although missing data at follow-up for leavers was of primary interest. The second comparison, which used follow-up data for stayers, assumed that outcomes were stable over the period of study (i.e., no serial effects). If there were serial changes in outcomes over the year of follow-up, then differences between stayers and leavers might be mistakenly attributed to effects of attrition rather than serial effects (Manning et al., 1993). However, by including covariates for survey period, we could account for serial effects among stayers in this comparison by allowing variation in the outcome to be decomposed into that related to attrition group and into that related to survey time point, independently of each other.

## Method

### *Participants and procedures*

*Study design.* A Web-based survey was conducted as part of a randomized community trial of an intervention to prevent high-risk alcohol consumption and its consequences on college campuses and surrounding communities. The meth-

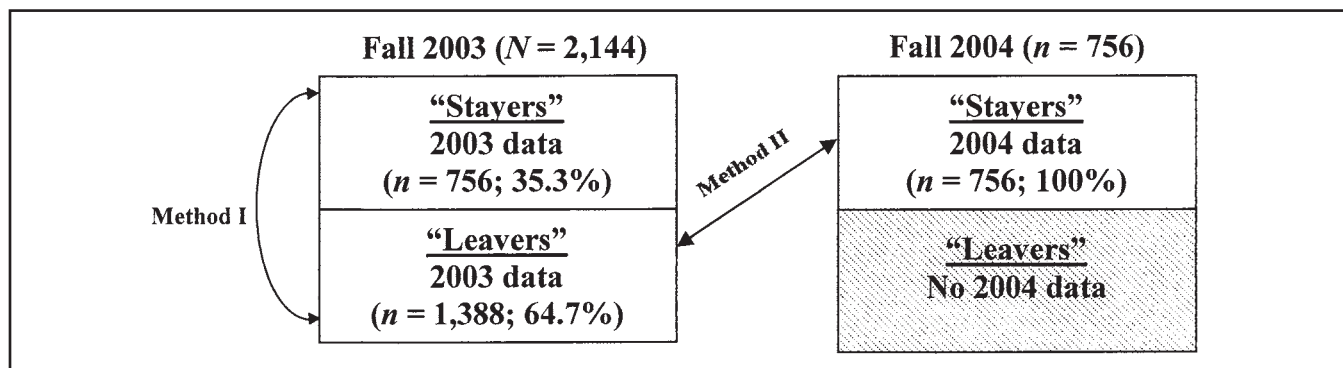


FIGURE 1. Diagram of freshmen cohort follow-up from fall 2003 to fall 2004. Method I compares leavers' 2003 data versus stayers' 2003 data. Method II compared leavers' 2003 data versus stayers' 2004 data, adjusting for stayers' 2003 data.

ods used for data collection in this study have been described previously (Mitra et al., 2008) and are summarized below. Participation in the survey was voluntary. The protocol was approved by the institutional review board at Wake Forest University School of Medicine.

*Study setting and population.* In the fall of 2003 (baseline year) and 2004 (follow-up year), class-year stratified random samples of undergraduate college students attending the 10 participating universities (8 public and 2 private) were invited to complete a Web-based survey. A random sample of students was selected for each school after undergraduate enrollment lists were provided to the study team by each university's registrar.

Initially, posters were placed in common areas of the residence halls and other places on campus (student union, cafeteria, etc.) asking students to check their emails for an offer to participate in the study. Prenotification postcards were mailed to randomly selected students selected by their campus email address. The postcards encouraged students to check their campus email addresses for an invitation to participate in the survey. It also informed them of a \$10 completion incentive and study institutional review board approval.

The class-stratified random sample of students received an email invitation in mid-October to complete the survey, which was electronically accessible on a secure Web site. The Web-based survey was hosted by an encrypted server. Students selected for the survey each received one invitation email message; nonresponders received up to four reminder email messages. Students were each sent an email message with a unique identifying number so that only they could access the survey. After students logged in to complete the survey, they had the option of leaving the survey incomplete and returning to it later. However, after the student "submitted" their responses, they could not return to the survey.

For the freshmen cohort, the Web site was shut down shortly after the target number of 215 freshmen (or close to target number) from each of the 10 schools was achieved. The target sample size was determined based on the statistical power for the initial study design. Students who completed the survey were sent email messages awarding them a \$10 check paid through an electronic payment service and entry into a \$100 cash prize lottery (one per school).

*Survey content and administration.* The Web-based survey was developed after reviewing the Harvard College Alcohol Survey (Wechsler et al., 1994), the Core questionnaire (Presley et al., 1994), the Youth Survey questionnaire used in the National Evaluation of the Enforcing Underage Drinking Laws program (Preisser et al., 2003; Wolfson et al., 2004), DeJong's College Drinking Survey (DeJong et al., 2006), and the Youth Risk Behavior Survey (Kolbe, 1990).

The number of questions increased slightly between years and had between 250 and 300 items (with multiple skip patterns based on different responses). The survey took be-

tween 17 and 24 minutes to complete. The survey measured demographic variables, smoking and drinking behaviors and consequences experienced from one's own drinking and from others' drinking, and other risk behaviors. The survey was pretested by college students at campuses not participating in the study for clarity and length and was subsequently revised by the study team.

### Measures

*Attrition group.* Attrition groups were created based on two categories: stayers or leavers. Leavers were operationalized as those students who were in the cohort in 2003 but dropped out in 2004 (missing data for all survey items). An indicator variable for attrition group was created using 1 = leaver and 0 = stayer.

*Smoking outcomes.* One item was used to measure the number of days of cigarette smoking in the past 30 days. This outcome was measured on an ordinal scale with seven response options (0 days, 1-2 days, 3-5 days, 6-9 days, 10-19 days, 20-29 days, and all 30 days). Midpoints of response intervals were taken and treated continuously.

*Drinking outcomes.* For the purposes of this study, two items were used to measure high-risk drinking: estimated number of days of heavy episodic drinking in the past 30 days and getting drunk in a typical week (O'Brien et al., 2006). Heavy episodic drinking in the past 30 days was estimated using the question, "In the past 30 days, how many days have you had FOUR/FIVE or more drinks in a row?" Women were asked about having four or more drinks in a row, and men were asked about having five or more drinks in a row. Response options to this item were 0 days, 1-2 days, 3-5 days, 6-9 days, 10-14 days, 15-19 days, 20-29 days, or all 30 days. Past-30-day heavy episodic drinking was defined as drinking four or more drinks in a row for women and five or more drinks in a row for men. Getting drunk was measured with the question, "In a typical week, how many days do you get drunk?"

Four drinking outcomes were used in the analyses reported below: (1) the estimated number of days the subject consumed alcohol in past 30 days calculated by taking the midpoints of the response intervals (same as heavy episodic drinking), (2) the estimated number of days of heavy episodic drinking (five/four drinks in a row; gender specific) in the past 30 days calculated by taking the midpoints of the response intervals, (3) the number of days the subject got drunk in a typical week, and (4) the maximum number of alcoholic drinks consumed in the past 30 days on any one drinking occasion.

*Alcohol-related consequences.* Five alcohol-related consequences were considered in this study. These were measured for a subset of students who indicated they had drunk alcohol in the past 30 days. Students were asked the number of times that each consequence was experienced "while you

were drinking alcohol or due to your alcohol use.” These included the following: taken advantage of sexually, drove a car or rode in a car with someone under the influence, was in an automobile accident, was in a physical fight, and was hurt or injured or required medical treatment. Indicator variables were created for each consequence, coded as 1 = 1 or more times and 0 = 0 times experienced.

*Covariates.* Student demographics were summarized for sample description. An indicator variable for gender was created (1 = female, 0 = male). Four indicator variables for race/ethnicity were created with white as the referent group (black, Hispanic, Asian/Pacific Islander, and other). Student age was considered as a continuous variable. An indicator variable for fraternity/sorority pledge/member status was created (1 = pledge/member, 0 = no), as well as an indicator variable for residence location (1 = on campus, 0 = off campus).

### Statistical analyses

*Sample descriptives.* Student demographics were summarized by attrition groups using frequency tables with percentages for categorical variables or means and standard deviations for continuous variables. Bivariate mixed-effects logistic regression was used to test if demographics were associated with attrition group (dependent variable). Within-school clustering that resulted from randomization at the campus level (Murray, 1998) was accounted for in the analyses using a random effect for school.

*Correlation of measures for cohort stayers.* To investigate how consistent data from stayers were, Pearson correlations were calculated for stayers using their drinking outcomes data from the 2003 and 2004 surveys.

*Method I.* Because we had two waves of data for the stayers (baseline and follow-up) and only one wave for

leavers (baseline), we analyzed attrition patterns using two methods (denoted Methods I and II). In Method I, we used both groups’ baseline data and compared the two in bivariate analyses (Figure 1: Method I). Linear mixed-effects modeling was performed for smoking and drinking outcomes, and logistic mixed-effects modeling was performed for alcohol-related consequences, unadjusted for other covariates. The within-school clustering that resulted from randomization at the campus level (Murray, 1998) was accounted for in the analyses using a random effect for school (range of school intraclass correlation was <.001 to .031).

*Method II.* In Method II, we compared stayers’ data at follow-up with leavers’ data at baseline using multivariable mixed-effects modeling, after adjustment for baseline measures (Figure 1: Method II). Thus, this method used all available data from stayers and leavers with maximum-likelihood methods as an approach. Linear mixed-effects modeling was performed for smoking and drinking outcomes, and logistic mixed-effects modeling was performed for alcohol-related consequences. These models adjusted for survey year (2004 vs 2003), condition (control or intervention school), and Condition  $\times$  Year interaction. The focal predictor variable in these analyses was the binary indicator variable for attrition group (leaver vs stayer). Within-school clustering as a result of randomization at the campus level (Murray, 1998) and repeated measures as a result of multiple survey responses for the stayers were accounted for in the analyses using random effects for each (range of school intraclass correlation was <.001 to .032). Specific contrasts between attrition groups were performed to test if there were significant differences between stayers and leavers using these models.

*General considerations.* The Kenward-Roger degrees of freedom method was used in mixed-effects modeling (Kenward and Roger, 1997; Littell et al., 2006). A two-sided  $p$  value < .05 was considered to be statistically significant. All

TABLE 1. Demographics by attrition group using 2003 baseline data ( $N = 2,144$ )

Characteristic	“Stayers” ( $n = 756$ ; 35%) % or mean (SD)	“Leavers” ( $n = 1,388$ ; 65%) % or mean (SD)	Test statistic	$p^a$
Gender			$z = -2.17$	.03
Male	40.9%	35.1%		
Female	59.2%	64.9%		
Race			$\chi^2 = 3.45$ , 4 df	.49
All other	4.9%	4.1%		
Asian/Pacific Islander	5.6%	3.8%		
Black	6.4%	7.5%		
Hispanic	3.3%	4.5%		
White	79.9%	80.0%		
Age	18.4 (1.5)	18.3 (1.1)	$z = 1.47$	.14
Fraternity/sorority member or pledge, yes vs no	9.3%	9.2%	$z = 0.72$	.47
Residence location			$z = -0.83$	.41
On campus	90.7%	87.7%		
Off campus	9.3%	12.3%		

<sup>a</sup> $p$  value from mixed-effects bivariate logistic regression for attrition group, adjusting for school clustering.

TABLE 2. Analysis using Method I of cohort leavers versus stayers (N = 2,144)

Outcome	“Stayers” (n = 756; 35%) % or mean (SE)	“Leavers” (n = 1,388; 65%) % or mean (SE)	b or OR <sup>a</sup>	Test statistic (df)	p <sup>a</sup>
<b>Smoking outcomes</b>					
Estimated no. days smoking cigarettes in past 30	3.05 (0.57)	4.27 (0.51)	1.23	t = 3.00, 2,084 df	<.01
<b>Drinking outcomes</b>					
Estimated no. days heavy episodic (5/4) in past 30 days	2.28 (0.27)	2.76 (0.23)	0.47	t = 2.28, 1,930 df	.02
No. days get drunk in typical week	0.57 (0.07)	0.67 (0.06)	0.10	t = 2.04, 2,061 df	.04
Estimated no. days drinking in past 30 days	3.43 (0.31)	4.08 (0.28)	0.64	t = 2.67, 2,034 df	<.01
Most no. drinks in past 30 days	3.87 (0.32)	4.58 (0.29)	0.72	t = 3.27, 2,074 df	<.01
<b>Alcohol-related consequences<sup>b</sup></b>					
Was taken advantage of sexually	5.4%	7.1%	1.33	z = 1.17	.24
Drove a car under the influence/ rode with a driver under the influence	32.4%	34.3%	1.09	z = 0.65	.52
Automobile accident	1.3%	0.7%	0.50	z = -1.16	.25
Physical fight	3.9%	6.8%	1.81	z = 2.16	.03
Was hurt or injured/ required medical treatment	10.5%	13.5%	1.33	z = 1.58	.11

<sup>a</sup>p value, b (unstandardized coefficient) from mixed-effects linear regression for smoking and drinking outcomes or odds ratio (OR) from mixed-effects logistic regression for consequences comparing leavers vs stayers; <sup>b</sup>asked past-30-day drinkers only; % reporting one or more times.

analyses were performed using SAS v9.1.3 (SAS Institute, Cary, NC) and Stata v10 (StataCorp, College Station, TX).

**Results**

*Sample descriptives*

Table 1 provides the sample demographics by attrition group. Of the 2,144 initial freshmen cohort students in fall 2003, 1,388 (65%) were lost to attrition in fall 2004 (Figure 1). Overall, in fall 2003, 26% of the students were past-30-day smokers. A little less than two thirds (63%) of the students were past-30-day drinkers and thus were asked questions about alcohol-related consequences in the past 30 days. Leavers were significantly more likely to be female (65% vs 59% of stayers). No other demographic characteristics were associated with attrition group (p > .10).

*Correlation of drinking measures for cohort stayers*

For stayers, who responded to two surveys one academic year apart, all drinking-outcome measures were moderately highly correlated (range of r = .54-.72, p < .0001).

*Bias under Method I*

*Bias in smoking outcome.* The average number of days of past-30-day smoking was significantly different between stayers and leavers (Table 2). Leavers reported number of

smoking days to be 1.2 days higher on average than that of stayers.

*Bias in drinking outcomes.* All four drinking outcomes (i.e., estimated number of days of any drinking and heavy episodic drinking in the past 30 days, getting drunk in a typical week, and most number of drinks in past 30 days) were significantly different for stayers and leavers (Table 2). The average frequency or quantity for these outcomes was significantly higher among the cohort leavers. The estimated number of drinking days reported by leavers was 0.7 days higher than that of stayers, as was the maximum number of drinks on a drinking occasion in the past 30 days. Leavers also reported 0.5 more days on average of estimated heavy episodic drinking and 0.1 days more of being drunk in a typical week.

*Bias in alcohol-related consequences.* Only one of the five past-30-day alcohol-related consequences was significantly associated with attrition (Table 2). A higher proportion of cohort leavers reported that they “got into a physical fight” at least once (stayers = 3.9% vs leavers = 6.8%, p = .03).

*Bias under Method II*

*Bias in smoking outcome.* Similar to the results using Method I, leavers reported significantly more days of past-30-day smoking than stayers (Table 3). Here, leavers reported 0.6 more days of smoking cigarettes on average than stayers.

*Bias in drinking outcomes.* Again, all four drinking

TABLE 3. Analysis using Method II of cohort leavers versus stayers ( $N = 2,144$ )

Outcome	"Stayers" ( $n = 756$ ; 35%) % or mean (SE)	"Leavers" ( $n = 1,388$ ; 65%) % or mean (SE)	$b$ or OR <sup>a</sup>	Test statistic (df)	$p^a$
Smoking outcomes					
Estimated no. days smoking cigarettes in past 30	3.63 (0.56)	4.27 (0.51)	1.23	$t = 3.00$ , 2,231 df	<.01
Drinking outcomes					
Estimated no. days heavy episodic (5/4) drinking in past 30	2.34 (0.26)	2.75 (0.23)	0.50	$t = 2.44$ , 2,321 df	.01
No. days get drunk in typical week	0.57 (0.07)	0.67 (0.06)	0.10	$t = 2.04$ , 2,408 df	.04
Estimated no. days drinking in past 30 days	3.76 (0.30)	4.06 (0.27)	0.64	$t = 2.68$ , 2,404 df	<.01
Most no. drinks in past 30 days	4.33 (0.32)	4.57 (0.29)	0.74	$t = 3.39$ 2,312 df	<.01
Alcohol-related consequences <sup>b</sup>					
Was taken advantage of sexually	5.9%	7.1%	1.35	$z = 1.17$	.24
Drove a car under the influence/ rode with a driver under the influence	35.0%	34.6%	1.25	$z = 1.14$	.25
Automobile accident	1.9%	0.7%	0.50	$z = -1.20$	.23
Physical fight	6.1%	6.9%	2.01	$z = 2.17$	.03
Was hurt or injured/ required medical treatment	12.0%	13.0%	1.48	$z = 1.71$	.09

<sup>a</sup> $p$  value,  $b$  (unstandardized coefficient) from mixed-effects linear regression or odds ratio (OR) from logistic regression adjusting for survey year, condition, Condition  $\times$  Year interaction, and school clustering comparing leavers versus stayers; <sup>b</sup>asked past-30-day drinkers only; % reporting one or more times.

outcomes were significantly different for stayers and leavers after adjusting for intervention group, survey year, and Condition  $\times$  Year interaction (Table 3). Similar to the results obtained using Method I, the average frequency or quantity for these outcomes was significantly higher among the cohort leavers. Leavers' reported estimated heavy episodic drinking was 0.4 days higher and weekly drunkenness was 0.1 days higher than those of stayers. The estimated number of days drinking in the past 30 days was 0.3 days higher for leavers than that of stayers, and the maximum number of drinks in a single occasion was 0.2 drinks higher for leavers.

*Bias in alcohol-related consequences.* Results for consequences using Method II were similar to those under Method I (Table 3). Only one of the five alcohol-related consequences ("got into a physical fight") was significantly associated with attrition group (stayers = 6.1% vs leavers = 6.9%,  $p = .03$ ). There was a marginally significant difference between the two groups in "Was hurt or injured/required medical treatment" ( $p = .09$ ). Point estimates of prevalence were higher for three of five consequences for cohort leavers.

## Discussion

Accurate and unbiased estimation of alcohol-related outcome variables is an integral component of efforts to understand and prevent high-risk drinking and consequences related to alcohol consumption. Our findings suggest that the gender of leavers is different from stayers; leavers generally have higher levels of smoking, alcohol-related behavior, and alcohol-related consequences; the differences between

leavers and stayers is moderate and varies across outcomes; and, for some variables, the difference could translate into substantial bias.

In this study, in addition to comparing demographic information of stayers and leavers, we used two methods to directly examine attrition bias on nine alcohol-related outcomes (four related to drinking behavior and five related to consequences) using cohort data from college freshmen. Direct assessment of outcome bias using Methods I and II show that attrition bias is statistically significant in the same 6 of 10 outcomes: smoking, estimated heavy episodic drinking, weekly drunkenness, estimated number of days drinking in past 30 days, most number of drinks in past 30 days, and reporting being in at least one physical fight in past 30 days.

Although statistical testing for identifying variables that reflect attrition bias identifies 6 significant differences of 10 variables in both Method I and II, there was also a strikingly consistent pattern in the direction of the effects using either method. In 9 of 10 outcomes reported in Table 2 using Method I, leavers have higher levels of smoking, alcohol consumption, and alcohol-related incidences than stayers. For Method II (Table 3), the proportion is 8 of 10. The only variable for which the direction is reversed under both methods is "automobile accident," which may be an artifact of the small number of automobile accidents that were reported in both groups (only 12 accidents at baseline and 10 at follow-up).

Even including "automobile accident," it is likely that the stayers and leavers differ in alcohol-related behaviors and consequences. Indeed, using a two-sided binomial test, the  $p$

value for observing nine events in one direction of 10 trials, assuming that there is no attrition bias, is .025. Therefore, the binomial test of direction supports the hypothesis that leavers tend to have a higher level of alcohol consumption and consequences, although the level of difference may be moderate and may also vary across different outcome measures.

For some of the variables, the differences between stayers and leavers may reach levels that are of clinical interest. As an example, the difference between stayers and leavers in the maximum number of drinks reported is 0.7 (Method I), which represents approximately an 18% difference. The range of bias for the drinking outcomes was -0.46 to -0.07 across the two methods (Groves, 2006), and for alcohol related consequences, the range of bias was -3.7% to 0.8%.

Although we found gender differences between stayers and leavers, White and Widom (2008) did not find gender differences in attrition groups in a cohort study of abused and nonabused children followed into adulthood. Our finding that leavers were proportionally more likely to be women seems to be opposite of what Grekin et al. (2007) found in a cohort of college students examining their Spring Break behaviors. Our results on leavers' alcohol use seem to agree with those of Bailey et al. (1992), who examined attrition groups in a cohort of middle school adolescents, but are different from those of White et al. (2006), who found no differences in a cohort of college students required to attend a drug assistance program. Regarding smoking, our findings agree with those of Cunradi et al. (2005), in that tobacco use was associated with increased attrition.

Although the findings in the current study provide some support for the position suggesting that attrition bias is generally small to moderate, our findings also suggest that attrition biases are prevalent and should not be ignored across variables. It is common in survey studies, especially those with a large cohort, to assess more than a single outcome or conduct substudies within the larger study that focus on a few selected outcome variables. If some of the attrition biases in a survey that assesses multiple outcomes are large enough to reach a level of clinical significance, then across-the-board neglect of attrition bias might lead to biased results for substudies that happen to focus on these variables.

We recommend that investigators carry out baseline comparisons for stayers versus leavers using logistic regression, as previously described by Miller and Hollist (2007). We also recommend that investigators have an open dialogue with their study team about the likelihood of the believability of the MAR assumption for missing data as a result of attrition. If the MAR assumption is tenable, then analyses such as selection modeling, analyses using all available data with maximum-likelihood methods, or use of multiple imputation could be performed. Results and conclusions from these analyses can be compared with those from complete case data (i.e., stayers-only data) as one way to examine sensitivity. If conclusions are sensitive to the analysis, the sensitiv-

ity should be noted and described (Schafer and Graham, 2002).

The implications from this manuscript are that when reporting similar outcomes from a similar cohort study of college freshmen, researchers should be cognizant of the magnitude and direction of possible attrition bias. Suggestions for future research include examining cohorts of other college students who are nonfreshmen among different classes or examining longitudinal studies with extended follow-up (more than two time points) to see if the current findings are replicated.

There are several limitations to this study. One is that assumptions need to be made about how leavers would behave at follow-up and that being a leaver is not related to outcome. For Method II, statistical adjustments were made that included time lapse between the two survey periods and the effect of intervention. Thus, the validity of our findings would depend on the quality of the assumed modeling adjustments.

An additional limitation is that the study does not investigate nonresponders to the baseline 2003 survey. Thus, this study investigates only sample attrition effects. The sample in 2003 is treated as the baseline sample in the cohort data, but whether this baseline sample is representative of the population is debatable. This study considers only leavers who are included in the baseline sample, which limits the generalizability.

Also, this study did not look at academic success indicators as a possible explanatory variable. Although we did collect current cumulative grade-point average in our survey, 70% of the freshmen cohort reported not yet having a current grade-point average at the time of survey implementation, and thus grade-point average could not be incorporated into the analyses. Other studies have found that students with lower grades had higher dropout rates (Bailey et al., 1992; Paschall and Freisthler, 2003).

Finally, the study uses data collected in a college drinking study using an Internet-based survey in 10 colleges located in the southeastern United States. Although the data were gathered from a reasonable level of geographically, demographically, and culturally diverse populations, our findings on attrition bias may not be generalizable to larger populations or to issues other than alcohol consumption. The purpose of these analyses was to explore the issue of attrition bias using a large dataset with complementary methods. The two methods showed congruence of empirical results. A practical implication of these findings is that for surveys similar to the current study, considerations should be made for limiting, assessing, and correcting for attrition bias.

## References

- BAILEY, S.L., FLEWELLING, R.L., AND RACHAL, J.V. The characterization of inconsistencies in self-reports of alcohol and marijuana use in a longitudinal study of adolescents. *J. Stud. Alcohol* **53**: 636-647, 1992.



- BOYS, A., MARSDEN, J., STILLWELL, G., HATCHINGS, K., GRIFFITHS, P., AND FARRELL, M. Minimizing respondent attrition in longitudinal research: Practical implications from a cohort study of adolescent drinking. *J. Adolesc.* **26**: 363-373, 2003.
- COLLINS, L.M., SCHAFER, J.L., AND KAM, C.-M. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychol. Meth.* **6**: 330-351, 2001.
- COMLEY, P. *Pop-Up Surveys: What Works, What Doesn't Work and What Will Work in the Future*, Hook, Hampshire, United Kingdom: Virtual Surveys Limited, 2000 (available at: [www.virtualsurveys.com/papers](http://www.virtualsurveys.com/papers)).
- COX, G.B., WALKER, R.D., FRENG, S.A., SHORT, B.A., MEIJER, L., AND GILCHRIST, L. Outcome of a controlled trial of the effectiveness of intensive case management for chronic public inebriates. *J. Stud. Alcohol* **59**: 523-532, 1998.
- CUDDEBACK, G., WILSON, E., ORME, J.G., AND COMBS-ORME, T. Detecting and statistically correcting sample selection bias. *J. Social Serv. Res.* **30**: 19-33, 2004.
- CUNRADI, C.B., MOORE, R., KILLORAN, M., AND AMES, G. Survey nonresponse bias among young adults: The role of alcohol, tobacco, and drugs. *Subst. Use Misuse* **40**: 171-185, 2005.
- DEJONG, W., SCHNEIDER, S.K., TOWVIM, L.G., MURPHY, M.J., DOERR, E.E., SIMONSEN, N.R., MASON, K.E., AND SCRIBNER, R.A. A multisite randomized trial of social norms marketing campaigns to reduce college student drinking. *J. Stud. Alcohol* **67**: 868-879, 2006.
- FITZGERALD, J., GOTTSCHALK, P., AND MOFFITT, R. An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics. *J. Human Resour.* **33**: 251-299, 1998.
- GREKIN, E.R., SHER, K.J., AND KRULL, J.L. College spring break and alcohol use: Effects of spring break activity. *J. Stud. Alcohol Drugs* **68**: 681-688, 2007.
- GROVES, R.M. Nonresponse rates and nonresponse bias in household surveys. *Publ. Opin. Q.* **70** (5 Spec. Issue): 646-675, 2006.
- JESSOR, R. AND JESSOR, S.L. Adolescent development and the onset of drinking. *J. Stud. Alcohol* **36**: 27-51, 1975.
- KENWARD, M.G. AND ROGER, J.H. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* **53**: 983-997, 1997.
- KOLBE, L.J. An epidemiological surveillance system to monitor the prevalence of youth behaviors that most affect health. *Hlth Educ.* **21**: 44-48, 1990.
- LITTELL, R.C., MILLIKEN, G.A., STROUP, W.W., WOLFINGER, R.D., AND SCHAUBENBERGER, O. *SAS for Mixed Models*, 2nd Edition, Cary, NC: SAS Institute, 2006.
- LOCKE, T.F. AND NEWCOMB, M.D. Psychosocial outcomes of alcohol involvement and dysphoria in women: A 16-year prospective community study. *J. Stud. Alcohol* **64**: 531-546, 2003.
- MCCABE, S.E., DIEZ, A., BOYD, C.J., NELSON, T.F., AND WEITZMAN, E.R. Comparing web and mail responses in a mixed mode survey in college alcohol use research. *Addict. Behav.* **31**: 1619-1627, 2006.
- MCGUIGAN, K.A., ELLICKSON, P.L., HAYS, R.D., AND BELL, R.M. Adjusting for attrition in school-based samples: Bias, precision, and cost trade-offs of three methods. *Eval. Rev.* **21**: 554-567, 1997.
- MANNING, W.G., JR., DUAN, N., AND KEELER, E.B. *Attrition Bias in a Randomized Trial of Health Insurance*, Minneapolis, MN: School of Public Health, University of Minnesota, 1993 (available at: [www.hcp.med.harvard.edu/files/Attrite\\_2.pdf](http://www.hcp.med.harvard.edu/files/Attrite_2.pdf)).
- MARTINO, S.C., COLLINS, R.L., AND ELLICKSON, P.L. Cross-lagged relationships between substance use and intimate partner violence among a sample of young adult women. *J. Stud. Alcohol* **66**: 139-148, 2005.
- MENARD, S. *Longitudinal Research*, 2nd Edition, Thousand Oaks, CA: Sage, 2002.
- MILLER, R.B. AND HOLLIST, C.S. Attrition bias. In: SALKIND, N.J. (Ed.) *Encyclopedia of Measurement and Statistics*, Vol. 1, Thousand Oaks, CA: Sage, 2007, pp. 57-60.
- MILLER, R.B. AND WRIGHT, D.W. Detecting and correcting attrition bias in longitudinal family research. *J. Marr. Fam.* **57**: 921-929, 1995.
- MITRA, A., JAIN-SHUKLA, P., ROBBINS, A., CHAMPION, H., AND DURANT, R. Differences in rate of response to web-based surveys among college students. *Int. J. E-Learn.* **7**: 265-281, 2008.
- MURRAY, D.M. *Design and Analysis of Group-Randomized Trials*, New York: Oxford Univ. Press, 1998.
- O'BRIEN, M.C., MCCOY, T.P., CHAMPION, H., MITRA, A., ROBBINS, A., TEUSCHLER, H., WOLFSON, M., AND DURANT, R.H. Single question about drunkenness to detect college students at risk for injury. *Acad. Emer. Med.* **13**: 629-636, 2006.
- OUTES-LEON, I. AND DERCON, S. *Survey Attrition and Attrition Bias in Young Lives*, Young Lives, Technical Note No. 5, Oxford, United Kingdom: Department of International Development, University of Oxford, March 2008.
- PASCHALL, M.J. AND FREISTHLER, B. Does heavy drinking affect academic performance in college? Findings from a prospective study of high achievers. *J. Stud. Alcohol* **64**: 515-519, 2003.
- PREISSER, J.S., YOUNG, M.L., ZACCARO, D.J., AND WOLFSON, M. An integrated population-averaged approach to the design, analysis and sample size determination of cluster-unit trials. *Stat. Med.* **22**: 1235-1254, 2003.
- PRESCOTT, C.A. AND KENDLER, K.S. Associations between marital status and alcohol consumption in a longitudinal study of female twins. *J. Stud. Alcohol* **62**: 589-604, 2001.
- PRESLEY, C.A., MEILMAN, P.W., AND LYERLA, R. Development of the Core Alcohol and Drug Survey: Initial findings and future directions. *J. Amer. Coll. Hlth* **42**: 248-255, 1994.
- RHODES, S.D., BOWIE, D.A., AND HERGENRATHER, K.C. Collecting behavioural data using the World-Wide-Web: Considerations for researchers. *J. Epidemiol. Commun. Hlth* **57**: 68-73, 2003.
- SCHAFFER, J.L. AND GRAHAM, J.W. Missing data: Our view of the state of the art. *Psychol. Meth.* **7**: 147-177, 2002.
- SMITH, G.T., MCCARTHY, D.M., AND GOLDMAN, M.S. Self-reported drinking and alcohol-related problems among early adolescents: Dimensionality and validity over 24 months. *J. Stud. Alcohol* **56**: 383-394, 1995.
- WECHSLER, H., DAVENPORT, A., DOWDALL, G., MOEYKENS, B., AND CASTILLO, S. Health and behavioral consequences of binge drinking in college: A national survey of students at 140 campuses. *JAMA* **272**: 1672-1677, 1994.
- WHITE, H.R., MORGAN, T.J., PUGH, L.A., CELINSKA, K., LABOUVIE, E.W., AND PANDINA, R.J. Evaluating two brief substance-use interventions for mandated college students. *J. Stud. Alcohol* **67**: 309-317, 2006.
- WHITE, H.R. AND WIDOM, C.S. Three potential mediators of the effects of child abuse and neglect on adulthood substance use among women. *J. Stud. Alcohol Drugs* **69**: 337-347, 2008.
- WOLFSON, M., ALTMAN, D., DURANT, R., SHRESTHA, A.H., ZACCARO, D., FOLEY, K.L. ET AL. *National Evaluation of the Enforcing Underage Drinking Laws Program: Year 4 Report*, Winston Salem, NC: Wake Forest University School of Medicine, 2004.