Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Bennett McDonald Date

A Multilevel Model of Longitudinal Non-response: Implications for Studies of College Student Substance Use

By

Bennett McDonald
Master of Public Health

Department of Epidemiology

_____
Michael Goodman
Committee Chair

_____
Carla Berg
Committee Member

_____
Regine Haardoerfer
Committee Member

_____
Michael Windle
Committee Member

A Multilevel Model of Longitudinal Non-response: Implications for Studies of College Student Substance Use

By

Bennett McDonald

B.S. Psychology
University of Florida
2014

Faculty Thesis Advisor: Michael Goodman, M.D.

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2016

Abstract

A Multilevel Model of Longitudinal Non-response: Implications for Studies of College Student Substance Use

By Bennett McDonald

Background: Methodologically, unit non-response in longitudinal studies poses challenges in obtaining unbiased measures of prevalence and association. Differential non-response due to measured respondent characteristics should be explored to best correct for this bias in analysis of further waves.
Methods: We examined individual and institutional level predictors of unit non-response between waves 1 and 2 of a multi-wave longitudinal study of student tobacco and substance use using multilevel modeling. The sample constituted 3,418 college students aged 18-25 from 7 Georgia colleges/universities.
Results: 450 (12.9%) participants were lost to follow up at wave 2. Multilevel model results indicated the odds of non-response at wave 2 were higher for males, Blacks, those living off campus, those with lower academic motivation, and those who used tobacco within the past 30 days. At the institutional level, the odds of non-response were higher for those attending private colleges/universities and those attending schools with student populations fewer than 10,000.
Discussion: Future longitudinal studies should assess predictors of non-response between waves to allow for correction of selection bias due to differential non-response. Techniques, such as the use of propensity scores, can then be used to correct for selection bias.

A Multilevel Model of Longitudinal Non-response: Implications for Studies of College Student Substance Use

By

Bennett McDonald

B.S. Psychology
University of Florida
2014

Faculty Thesis Advisor: Michael Goodman, M.D.

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2016

Table of Contents

Acknowledgements

I would like to thank everyone from the Project DECOY team for their continued support throughout the thesis process and Dr. Berg for welcoming me to the team. None of this would have been possible without them. I'm also very grateful to my friends and family for keeping spirits high during the toughest parts of this process and for pushing me to do the best that I can in all facets of life.

Introduction

Rapid development of new information technology tools has revolutionized survey implementation. In parallel with this innovation, however, researchers have noticed a universal and consistent decline in survey participation (1-4). Evidence suggests that the rate of non-response in survey research has increased within the past decade when compared to the latter half of the 20th century (5). This lack of participation poses potential threats to the validity of population based studies that rely on surveys for data collection (6-9). Investigation of the potential threats to validity (a.k.a. biases) due to non-participation has not been sufficiently explored in studies of student tobacco and substance use.

There are two recognized types of survey non-response: item non-response and unit non-response. Item non-response refers to the situation in which a participant refuses or provides no answer for a particular survey question and is common when items ask about sensitive topics, such as income, race, or illicit substance use. Item non-response is often treated in the analysis as a missing data problem. Unit non-response refers to failure to enroll or loss to follow up and will be the focus of this study (10).

Although avoiding non-response is desirable in public health research, a lack of participation does not always result in biased estimates (1, 11). Survey dropout or refusal can occur for a variety of situational or individual reasons, but problems arise when the nature of the non-response of the study population becomes systematic. Both unit and item non-response data can be characterized as missing completely at random, missing at random, and missing not at random. Missing completely at random occurs when no factors are related to survey non-response, suggesting missing data is due to chance

occurrence and minimal to no bias exists. Missing at random occurs when there is some association between an outcome of interest and other variables, but the missingness can be explained by measured factors available to the researcher; methods, such as multiple imputation, can then be used to make reasonable estimates of outcome variables (12). Missing not at random is the most problematic. This occurs when systematic differences between responders and non-responders are present even after accounting for available variables, leaving outcomes of interest prone to bias (13, 14).

When data are not missing completely at random, selection bias becomes a concern. Often data are incorrectly assumed to be missing completely at random and only analyses of individuals with complete data are reported, potentially resulting in biased associations or skewed prevalence estimates.

Non-response in university students

University students and other young adults, a commonly studied population, have shown declining participation rates in research during the past decade (15, 16). Literature investigating non-response patterns of university students has found similarities to patterns seen in general population health surveys, with males, minorities, and those less academically motivated having lower odds of participation (17-20). Mixed results have been reported concerning the effect of socioeconomic status on survey response, with one study reporting lower participation among those coming from more affluent households (21) and another reporting lower participation among those using financial aid (20). In terms of residence, those living on campus are more likely to respond than those living off campus (22). Although many studies have found lower levels of non-response as

academic achievement increased, one study found that first year university students failed to participate due to school related responsibilities or a lack of time (17). Furthermore, many studies have found that tobacco use predicted cross-sectional or longitudinal non-response in surveys of the general population or younger adults in general (23-26), yet this association has not been sufficiently explored in the college student population. For example, one study found that past 30-day alcohol and past 30-day tobacco use predicted longitudinal study attrition in a student population (27), but further research is needed to help elucidate this relationship, particularly in studies of student tobacco, alcohol, and substance use.

Very few studies have examined associations of non-response at the institutional level. Porter et al found that public schools or those located in urban areas showed higher odds of non-response (28). Geographically, other studies found differences were important, with rates of response reported to be higher in the Midwest compared to the West (29) and in the West compared to the East (21). Lastly, institutions with larger student population size have exhibited lower levels of survey participation (16, 28).

Non-response in longitudinal tobacco use studies

Previous longitudinal studies of tobacco use in student populations have encountered varying levels of non-response. Few of those studies addressed non-response bias by comparing characteristics of participants and non-participants (30, 31) or early vs. late respondents in terms of recruitment period (32), or by reporting conservative baseline estimates (33). In order to obtain accurate estimates of prevalence or measures of association concerning tobacco use, correlates of non-response should be explored. In

addition, the association of baseline tobacco use with subsequent wave attrition needs to be estimated in order to ensure changes in tobacco use are unbiased.

In this study, we aim to estimate the association between 30-day tobacco use at baseline and wave 2 attrition controlling for relevant factors. In addition, we will also explore individual and institutional factors associated with attrition, such as sociodemographic and academic variables, urban/rural location, school size, and private school status, using a mixed effects modeling approach.

Methods

Sample design and data collection

Project DECOY, **D**ocumenting **E**xperiences with **C**igarettes and **O**ther **T**obacco in **Y**outh, is a quantitative, longitudinal assessment of tobacco use predictors in Georgia college students. The study was initiated in 2014. Data are being collected from 7 Georgia colleges, including two public schools, two private schools, two community/technical colleges, and one historically black university. The colleges were selected with the goal of obtaining a diverse sample with regards to tobacco policies and institutional factors. Surveys were administered every 4 months across 6 waves of data collection during spring, summer, and fall. Eligible participants were between 18 and 25 years of age who were able to read English.

A list of students was obtained from each institution's office of the registrar. Using these lists as the sampling frame, 3,000 randomly selected students from each of one private and two public schools were invited to participate. The remaining institutions contained fewer than 3,000 students, and invitations were sent to all eligible participants.

Recruitment was closed after a recruitment goal was reached. Response rates ranged from 12% to 59%, with an overall response rate of 23% (N=3574/15,607). A week after completion of the baseline survey, participants were asked to confirm their participation in the study via an emailed link and were provided their first gift card. The final response rate after confirmation was 22% (N=3418/3574).

Several techniques were used to increase retention of participants during follow up. First, social media contacts (in addition to primary and alternate e-mail addresses) were obtained in order to increase probability of reaching participants. An option to provide contact information of relatives and others likely to know the whereabouts of the participant was given. Second, small gifts were provided via social media to enhance engagement with participants. Updates and reminders regarding survey procedures were provided through the DECOY Twitter account and Facebook page. Third, text messages were sent to participants one month before survey administration via Trumpia (DoCircle, Inc., Anaheim, California) to remind participants of the survey and to allow them to update their email address. Lastly, survey incentives were provided on a gradual schedule, increasing from $30 at baseline to $50 for the final wave. If participants completed all 6 surveys, they received an additional $100. Further details can be found elsewhere (34).

Measures

The baseline DECOY survey was administered via the internet and took between 30 and 45 minutes to complete. Measures included broader multilevel variables, such as community, organization, and interpersonal factors; intrapersonal characteristics, such as

sociodemographics, general health information, attitudes towards tobacco, and market segmentation factors; and health outcomes, such as prevalence and frequency of tobacco use.

For analysis of non-response, three relevant institutional factors were examined. Rural/urban status of the institution was assessed with the hypothesis that non-response would be higher for those in urban settings following general patterns reported in the literature (28). We also predicted that non-response would be higher at institutions with greater student population, and a binary variable of student population >10,000 was created. Type of school was represented by a binary variable characterizing private school status (yes/no). We hypothesized that private schools would shower lower non-response following previous results (28).

Several individual level characteristics were also hypothesized to have an effect on non-response. Following the relatively robust findings concerning demographic characteristics on non-response in the general and student population, we hypothesized that non-response would be higher for males, non-Whites, and those less academically motivated. Three categories of academic motivation, measured as the highest degree the student wished to achieve, were created and included having an undergraduate certificate or some college, a bachelor's degree, or an advanced degree. Parental education was used as a proxy for socioeconomic status and measured as the highest degree which either of a student's parent received, classified as some college or less, an associates or bachelor's degree, or an advanced degree. We hypothesized individuals whose parents had higher degrees would be more likely to retain participation in the survey following previous findings linking financial aid status to non-response (20). Lastly, as noted above,

previous research found that those living on campus were more likely to respond to surveys, and for this reason we included campus residence as another variable of interest.

Our primary research question was whether past 30-day tobacco use (yes/no) was associated with non-response at wave 2, after controlling for relevant covariates. Alcohol use within the past 30 days was also included, and we predict those who used alcohol within the past 30 days will have higher non-response than those who did not. Given that the overarching study is primarily concerned with tobacco use, the literature suggests that those dropping out might do so due to a lack of topic salience or relevancy (35). On the other hand, it is also possible that those who use tobacco might be dropping out at a higher rate because of perceived stigma (36).

Analysis

The aims of the statistical analysis were to 1) investigate the association between any past 30-day tobacco use and attrition at wave 2 of the DECOY survey, 2) assess which institutional characteristics are associated with DECOY survey attrition, and 3) examine the association between demographic characteristics and survey attrition.

First, bivariate associations between each predictor and non-response at wave 2 were assessed. T-tests and chi-square (or Fisher's Exact) tests were used for continuous and categorical variables, respectively, comparing baseline data for responders and non-responders. The Wilcoxon rank sum test was used for variables with distributions which were non-normal.

A general linear mixed model (GLMM) containing all relevant institutional and individual level variables was used. Results were expressed as adjusted odds ratios (OR) with the corresponding 95% confidence intervals (CI). Responders at wave 2 were used

as the referent group. An unconditional model with no variables entered was first used to estimate the intra-cluster correlation coefficient (ICC). This measure is used to describe the variability in non-response at wave 2 due to the nesting of students within institution. Students attending the same institution are assumed to be more similar due to the characteristics of that institution, and this variability due to institution may also be explored in addition to the effect of individual characteristics. Models containing only individual level variables, only institutional level variables, and both individual and institutional variables were then constructed. Individual characteristics were entered into the model based on a priori considerations. Model fit statistics were calculated and likelihood ratio tests were used to compare the change in deviance of nested models. Lastly, the reduction in level-2 (school level) variance compared to the unconditional model was calculated for each subsequent model. All analyses were conducted in SAS 9.4 (SAS Institute, Cary, NC).

Results

Descriptive statistics

Analysis of final sample demographic variables indicated that the majority of participants in the baseline survey were female (64%), white (62%) and non-Hispanic (92.5%). The mean age of all participants was 20.6 years with a standard deviation (SD) of 1.97. The majority of the 3418 participants came from private schools (39%) and public schools (27%), but students from technical colleges (22%) and a historically black university (12.0%) were also represented. Over a quarter (30%) of all participants indicated that they had used at least one tobacco product within the last 30 days and over half reported past 30-day alcohol use (63%).

Bivariate associations: wave 2 response

450 participants (12.9%) were lost to follow up at wave 2. Distributions of various individual and institutional characteristics in wave 2 responders and non-responders are presented in Table 1. The two groups were similar with respect to age ($p=.858$) or gender ($p=.763$). The proportion of blacks was highest in persons who did not respond to wave 2 ($p<.001$). Consistent with previous findings, non-responders were less likely to live on campus ($p<.001$), were less likely to seek a more advanced degree ($p<.001$), and included a greater proportion of persons whose parents had lower education levels ($p<.001$). Although the two groups reported very similar 30-day alcohol use ($p=.123$), non-responders were more likely to have used tobacco products within the past 30 days ($p<.001$). This pattern was observed for all tobacco products except for smokeless tobacco, a product with very low prevalence of use in college students (37-39).

Clear differences in non-response were also found at the institutional level. Non-responders were significantly more likely to attending schools in urban areas ($p<.001$). On the other hand, institutions with populations greater than 10,000 students ($p<.001$) or classified as private schools ($p<.001$) were overrepresented among respondents.

Multilevel modeling of non-response

Results from the model building process are found in Table 2. First, the unconditional model containing no predictors was fit. Covariance parameter estimates indicated an ICC of 10.1%, indicating that 10.1% of the variability in non-response was due to school level characteristics. Model selection criteria and likelihood ratio tests

using deviance statistics indicated that model fit increased significantly when comparing the model containing individual predictors only ($p<.001$) and institutional predictors only ($p<.001$) to the unconditional model. The full model containing all predictors was tested against the individual level model and had significantly better fit ($p<.001$). AIC and BIC fit statistics were consistent with these findings. Following these results, the full model was used for parameter estimates.

Results of the multivariable analyses are presented in Table 2. After controlling for covariates, age was still unrelated to non-response; odds of non-response, however, were significantly higher for males (OR=1.41; 95% CI=1.10, 1.79) and blacks (OR=1.74; 95% CI=1.23, 2.46). Indicating "other" race was not associated with non-response (OR=1.43; 95% CI=0.99, 2.05). Parental education of a bachelor's degree (OR=0.97; 95% CI=0.71, 1.31) or an associate's degree or lower (OR=0.97; 95% CI=0.72, 1.32) was not associated with non-response when compared to those having a parent with an advanced degree. Consistent with bivariate analyses, the odds of non-response were lower for those residing on campus (OR=0.62; 95% CI = 0.46, 0.84). Academic achievement was significantly associated with non-response, with odds of non-response higher for those who sought bachelor's degrees as opposed to advanced degrees (OR= 1.41; 95% CI=1.09, 1.83)

Analyses of substance use variables were also consistent with bivariate findings. Most importantly, the odds of non-response was significantly higher for those who used at least one tobacco product within the past 30 days (OR=1.42; 1.12, 1.80). Alcohol use was unrelated to non-response.

Analysis of school level characteristics showed that the association with urban/rural status was no longer present when included along with other variables in the full model (OR=1.05; 95% CI=0.75, 1.47). The odds of non-response at wave 2 for those attending private institutions was 0.48 times the odds of non-response of those at non-private intuitions (95% CI=0.33,0.71). In addition, the odds of non-response were significantly lower for those attending institutions with populations of greater than 10,000 students (OR=0.56; 95% CI=0.43, 0.73).

Discussion

This study investigated individual and institution-level predictors of dropout from a longitudinal survey in a diverse sample of students enrolled in universities of various types. Previous literature showed that dropout was related to a few key factors relevant to this study, including survey salience, tobacco use, and participant sociodemographics. Our main aim was to assess whether past 30-day tobacco use was associated with non-response controlling for relevant covariates. .

The results of the multilevel model confirmed the hypothesis that 30-day tobacco use predicted non-response at wave 2. This finding is consistent with survey non-response research conducted in the general population and young adults, although these studies' subject matter were not primarily concerned with tobacco use or studied unique populations like young adult military personnel (23-26). It is unique in its relevancy to college students participating in a study primarily concerned with tobacco use, and it rejects the idea that student dropout is largely due to a lack of interest or relevance with regards to the survey topic. Dropout instead could be related to feelings of stigmatization as a tobacco user in a less socially acceptable setting (40) or some other unmeasured

factor. Moreover, estimation of bias due to tobacco use will allow for the correction of tobacco use estimates and associations in future waves. Techniques for correction of selection bias such as inverse probability weighting can inflate the weights of groups with elevated dropout to preserve representation and prevent skewed estimates

At the individual level, many of our findings were consistent with previous literature. For instance, higher odds of non-response was seen in blacks and other minorities, males, those with lower academic motivation, and those living off campus, results that have been replicated in many different samples of college students (17-20). Surprisingly, a few variables which had often been linked to a lack of participation, such as age, socioeconomic status, and alcohol use, were not associated with non-response in this study (20, 21, 27). This observation, however, could be explained by measurement methods used, such as the use of parental education as a proxy for student socioeconomic status as opposed to financial aid status (20). Also influential could be the restriction of other student populations to very limited age groups, such as college freshmen, not seen in the general population studies of young adults (23-26).

Institutional variables also accounted for varying levels of non-response. For example, private schools exhibited lower odds of non-response compared to public schools, as seen in study conducted by Porter et al (2006). School urban/rural status was not associated with non-response. Worth noting was the higher odds of non-response for those attending schools with larger populations, a finding that contradicts previous findings (16, 28). This, however, could be due to the relatively small number of schools and should be investigated with a more representative sample.

This study exhibits many strengths. Selecting from a diverse group of schools and participants provided a heterogeneous sample relative to many other studies of student populations, including students of different race/ethnicity, urban/rural status, and socioeconomic status. Furthermore, the use of individual and institutional level variables has been rare in studies of cross-sectional and longitudinal non-response, particularly in studies of student tobacco and substance use, and our analysis of both levels is important for minimizing selection bias and obtaining accurate prevalence estimates and associations.

Several limitations should be considered. First, a small number of schools was used to predict school characteristics associated with non-response given that this was a secondary analysis of data. Although simulation studies have shown that inferences can still be drawn with a low number of clusters (41), caution should be used in drawing conclusions from these results. Second, certain covariates utilized measurement scales available for a secondary data analysis, such as ordinal age measurement instead of date of birth and parental education as a proxy for socioeconomic status, and alternative methods should be used to explore these associations further. Lastly, this study analyzed dropout between two subsequent waves of a longitudinal study, and future studies should attempt to assess dropout between and throughout all waves

Future studies should replicate these results in a broader sample of students and colleges not confined to one state. Although we obtained sufficient samples from each of our 7 schools, replication of these findings in in large samples and across other areas of the United States could potentially add to the knowledge of what predicts non-response in tobacco use studies in contextual settings. Superior methods may then be developed for

targeting those at highest risk of non-response and preventing bias due to systematic dropout in longitudinal studies. These efforts will also assist researchers in decreasing study bias and developing best practices to decrease smoking and other substance use behaviors in college students, a vulnerable population to these behaviors.

References

1. Groves RM. Nonresponse rates and nonresponse bias in household surveys. Public Opinion Quarterly. 2006;70(5):646-75.
2. Olson CA. Survey burden, response rates, and the tragedy of the commons. Journal of Continuing Education in the Health Professions. 2014;34(2):93-5.
3. Porter SR. Raising response rates: What works? New Directions for Institutional Research. 2004;2004(121):5-21.
4. Sheehan KB. E-mail Survey Response Rates: A Review. Journal of Computer-Mediated Communication. 2001;6(2):0-.
5. Curtin R, Presser S, Singer E. Changes in Telephone Survey Nonresponse over the Past Quarter Century. Public Opinion Quarterly. 2005;69(1):87-98.
6. Korkeila K, Suominen S, Ahvenainen J, Ojanlatva A, Rautava P, Helenius H, et al. Non-response and related factors in a nation-wide health survey. European journal of epidemiology. 2001;17(11):991-9.
7. Van Loon AJM, Tijhuis M, Picavet HSJ, Surtees PG, Ormel J. Survey Non-response in the Netherlands. Annals of Epidemiology.13(2):105-10.
8. Wilks R, Younger N, Mullings J, Zohoori N, Figueroa P, Tulloch-Reid M, et al. Factors affecting study efficiency and item non-response in health surveys in developing countries: the Jamaica national healthy lifestyle survey. BMC Medical Research Methodology. 2007;7(1):1-14.
9. Zhao J, Stockwell TIM, Macdonald S. Non–response bias in alcohol and drug population surveys. Drug and Alcohol Review. 2009;28(6):648-57.
10. Groves RM, Dillman DA, Eltinge JL, Little RJ. Survey nonresponse: Wiley New York; 2002.
11. Keeter S, Miller C, Kohut A, Groves RM, Presser S. Consequences of Reducing Nonresponse in a National Telephone Survey. Public Opinion Quarterly. 2000;64(2):125-48.
12. Rubin DB. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.
13. Goldstein H. Handling attrition and non-response in longitudinal data. Longitudinal and Life Course Studies. 2009;1(1).
14. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338.
15. Adams MJ, Umbach PD. Nonresponse and online student evaluations of teaching: Understanding the influence of salience, fatigue, and academic environments. Research in Higher Education. 2012;53(5):576-91.
16. Sharkness J. Why Don't They Respond? An Investigation of Longitudinal Survey Nonresponse Among College Students Attending Four-Year Institutions. 2012.
17. Cranford JA, McCabe SE, Boyd CJ, Slayden J, Reed MB, Ketchie JM, et al. Reasons for nonresponse in a web-based survey of alcohol involvement among first-year college students. Addictive behaviors. 2008;33(1):206-10.
18. Crawford SD, Couper MP, Lamias MJ. Web surveys perceptions of burden. Social science computer review. 2001;19(2):146-62.
19. Dey EL. Working with low survey response rates: The efficacy of weighting adjustments. Research in higher education. 1997;38(2):215-27.
20. Porter SR, Whitcomb ME. Non-response in student surveys: The role of demographics, engagement and personality. Research in higher education. 2005;46(2):127-52.
21. Szelényi K, Bryant AN, Lindholm JA. What money can buy: Examining the effects of prepaid monetary incentives on survey response rates among college students. Educational Research and Evaluation. 2005;11(4):385-404.

22. Rogelberg SG, Conway JM, Sederburg ME, Spitzmüller C, Aziz S, Knight WE. Profiling active and passive nonrespondents to an organizational survey. Journal of Applied Psychology. 2003;88(6):1104.
23. Cunradi CB, Moore R, Killoran M, Ames G. Survey nonresponse bias among young adults: the role of alcohol, tobacco, and drugs. Subst Use Misuse. 2005;40.
24. Vinther-Larsen M, Riegels M, Rod MH, Schiøtz M, Curtis T, Grønbæk M. The Danish Youth Cohort: characteristics of participants and non-participants and determinants of attrition. Scandinavian journal of public health. 2010;38(6):648-56.
25. Winefield AH, Winefield HR, Tiggemann M. Sample attrition bias in a longitudinal study of young people. Australian Journal of Psychology. 1990;42(1):75-85.
26. Young AF, Powers JR, Bell SL. Attrition in longitudinal studies: who do you lose? Aust N Z J Public Health. 2006;30.
27. McCoy TP, Ip EH, Blocker JN, Champion H, Rhodes SD, Wagoner KG, et al. Attrition bias in a US Internet survey of alcohol use among college freshmen. Journal of studies on alcohol and drugs. 2009;70(4):606-14.
28. Porter SR, Umbach PD. Student survey response rates across institutions: Why do they vary? Research in Higher education. 2006;47(2):229-47.
29. Jans M, Roman A. National response rates for surveys of college students: Institutional, regional, and design factors. Ann Arbor, MI: Michigan Programing Survey Methodology, University of Michigan, Institute for Social Research, Section on Survey Research Methods. 2007.
30. Nuzzo E, Shensa A, Kim KH, Fine MJ, Barnett TE, Cook R, et al. Associations between hookah tobacco smoking knowledge and hookah smoking behavior among US college students. Health education research. 2013;28(1):92-100.
31. Sidani JE, Shensa A, Barnett TE, Cook RL, Primack BA. Knowledge, attitudes, and normative beliefs as predictors of hookah smoking initiation: a longitudinal study of university students. nicotine & tobacco research. 2014;16(6):647-54.
32. Sutfin EL, McCoy TP, Berg CJ, Champion H, Helme DW, O'Brien MC, et al. Tobacco use by college students: a comparison of daily and nondaily smokers. American journal of health behavior. 2012;36(2):218.
33. Fielder RL, Carey KB, Carey MP. Prevalence, frequency, and initiation of hookah tobacco smoking among first-year female college students: a one-year longitudinal study. Addictive behaviors. 2012;37(2):221-4.
34. Berg CJ, Haardoerfer R, Lewis M. DECOY: Documenting Experiences with Cigarettes and Other Tobacco in Young Adults. American Journal of health behavior. 2016;in press.
35. McDaniel SW, Madden CS, Verille P. Do topic differences affect survey non-response? Journal of the Market Research Society. 1987.
36. Kann L, Brener ND, Warren CW, Collins JL, Giovino GA. An assessment of the effect of data collection setting on the prevalence of health risk behaviors among adolescents. Journal of Adolescent Health. 2002;31(4):327-35.
37. Enofe N, Berg CJ, Nehl EJ. Alternative Tobacco Use among College Students: Who is at Highest Risk? American journal of health behavior. 2014;38(2):180-9.
38. Latimer LA, Batanova M, Loukas A. Prevalence and harm perceptions of various tobacco products among college students. nicotine & tobacco research. 2013:ntt174.
39. Spangler J, Song E, Pockey J, Sutfin EL, Reboussin BA, Wagoner K, et al. Correlates of smokeless tobacco use among first year college students. Health education journal. 2014;73(6):693-701.
40. Stuber J, Galea S, Link BG. Smoking and the emergence of a stigmatized social status. Social science & medicine (1982). 2008;67(3):420-30.

41. Hoyle RH, Gottfredson NC. Sample size considerations in prevention research applications of multilevel modeling and structural equation modeling. Prevention Science. 2015;16(7):987-96.

Table 1: Bivariate associations of institutional and individual factors associated with non-response at wave 2 (n=3,418)

	Overall sample (n=3,481)		Responders (n=2,968)		Non-responders (n=450)		
	Mean (N)	SD (%)	Mean (N)	SD (%)	Mean (N)	SD (%)	p*
Individual level							
Age	20.5	1.97	20.5	1.94	20.6	2.13	.858
Race							<.001
Black	832	24.6	663	22.6	169	37.9	
White	2133	63.2	1903	65	230	51.6	
Other	411	12.2	364	12.4	47	10.5	
Gender							.763
Female	2199	64.4	1912	64.5	287	63.8	
Male	1215	35.6	1052	35.4	163	36.2	
Parental Education							<.001
No college degree	908	26.9	820	28.0	88	19.9	
Bachelor's degree	1156	34.3	1010	34.5	146	33.0	
Advanced degree	1309	38.8	1101	37.6	208	47.1	
Reside on campus							
Yes	1493	43.7	1354	45.6	139	30.9	<.001
Degree sought							<.001
Associate's or lower	232	7.1	187	6.6	45	10.5	
Bachelor's degree	668	20.5	549	19.4	119	27.7	
Advanced degree	2367	72.5	2101	74.1	266	61.9	
Past 30-day alcohol use							
Yes	2155	63.1	1893	62.2	297	66.0	.123
Past 30-day tobacco use							
Yes	1012	29.6	828	27.9	184	40.9	<.001
30-day use by product							
Cigarettes	455	13.3	369	12.4	86	19.1	<.001
E-cigarettes	372	10.9	307	10.3	65	14.4	.009
Hookah	416	12.2	342	11.5	74	16.4	.003
Little cigars/cigarillos	385	11.3	301	10.2	84	18.7	<.001
Smokeless tobacco	123	3.6	102	3.4	21	4.7	.211
Institutional level							
Urban school							
Yes	1327	38.8	1088	36.7	239	53.1	<.001
Private school							
Yes	1321	38.7	1236	41.6	85	18.9	<.001
Student population size							
>10,000	1501	34.9	1352	45.6	149	33.1	<.001

*p-value comparing responders and non-responders

Table 2: Results from a multi-level model assessing institutional and individual factors associated with non-response at wave 2

	Unconditional model		Individual level		Institutional level		Full model	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Individual level								
Age			0.94	(0.89, 1.00)			0.95	(0.89, 1.01)
Race								
White			*ref*	*ref*			*ref*	*ref*
Black			1.57 *	(1.14, 2.16)			1.74 *	(1.23, 2.46)
Other			1.30	(0.90, 1.87)			1.43	(0.99, 2.05)
Gender								
Female			*ref*	*ref*			*ref*	*ref*
Male			1.36 *	(1.07, 1.74)			1.41 *	(1.10, 1.79)
Parental Education								
No college degree			1.02	(0.75, 1.39)			0.97	(0.72, 1.32)
Bachelor's degree			1.01	(0.74, 1.38)			0.97	(0.71, 1.31)
Advanced degree			*ref*	*ref*			*ref*	*ref*
Reside on campus								
Yes			0.62 *	(0.46, 0.85)			0.62 *	(0.46, 0.84)
Degree goal								
Associate's degree or lower			1.05	(0.69, 1.60)			1.01	(0.68, 1.52)
Bachelor's degree			1.46 *	(1.22, 1.90)			1.41 *	(1.09, 1.83)
Advanced degree			*ref*	*ref*			*ref*	*ref*
Past 30-day tobacco use								
Yes			1.42 *	(1.12, 1.80)			1.41 *	(1.10, 1.78)
Past 30-day alcohol use								
Yes			1.14	(0.89, 1.47)			1.14	(0.89, 1.45)
Institutional level								
Urban school								
Yes					1.14	(0.74, 1.79)	1.05	(0.75, 1.47)
Private school								
Yes					0.37 *	(0.22, 0.60)	0.48 *	(0.33, 0.71)
Student population size								
>10,000					0.59 *	(0.41, 0.86)	0.56 *	(0.43, 0.73)

Model fit criteria							
τ_{00}	0.3694	0.2112		0.0417		0.0014	
Reduction in τ_{00}	*ref*	51.0%		89.6%		99.4%	
Deviance	2547.55	2336.54		2534.65		2321.32	
AIC	2551.55	2362.54		2544.65		2353.32	
BIC	2551.44	2361.84		2544.38		2352.46	
χ^2	*ref*	211.01	*	12.90	*	15.22	*

*p < .05