The Impact of Survey Mode on U.S. National Estimates of Adolescent Drug Prevalence: Results from a Randomized-Controlled Study

Richard A. Miech, Ph.D., MPH
University of Michigan

Mick P. Couper, Ph.D.
University of Michigan

Steven G. Heeringa, Ph.D.
University of Michigan

Megan E. Patrick, Ph.D.
University of Minnesota

Running Head: The Impact of Survey Mode

Word count: 3,680

Corresponding author:

Richard Miech
Super Professor
Institute for Social Research
University of Michigan
426 Thompson
Ann Arbor, MI 48104

Email: ramiech@umich.edu

Declaration of interest: This project funded by the National Institute on Drug Abuse, part of the U.S. National Institutes of Health, by grant # DA001411. All authors declare no conflicts of interest.

Abstract

Background and Aims: Increasing numbers of school-based drug surveys are transitioning data collection to electronic tablets from paper-and-pencil, which may produce a survey mode effect and consequent discontinuity in time trends for population estimates of drug prevalence. This study tested whether (a) overall, self-reported drug use prevalence is higher on electronic tablets v. paper-and-pencil surveys, (b) sociodemographics moderate survey mode effects, and (c) levels of missing data are lower for electronic tablet v. paper-and-pencil modes.

Design: A randomized-controlled experiment.

Setting: Results are nationally-representative of students in the contiguous United States.

Participants: 41,866 8th, 10th, and 12th grade students who participated in the 2019 Monitoring the Future school-based survey administration.

Intervention and comparator: Surveys were administered to students in a randomly-selected half of schools with electronic tablets (intervention) and with paper-and-pencil format (comparator) for the other half.

Measurements: Primary outcome was the total number of positive drug use responses. Secondary outcomes were percent of respondents completing all drug questions, percent of drug questions unanswered, and mean number of missing drug items.

Findings: The relative risk for total number of positive drug use responses for electronic tablets v. paper-and-pencil surveys were small and 95% confidence intervals included for reporting intervals of lifetime (RR=1.03; 95% CI, 0.93-1.14), past 12-months (RR=1.01; 95% CI, 0.91-1.11), past 30-days (RR=1.05; 95% CI, 0.93-1.20), and for heavy use (RR=1.10; 95% CI, 0.93-1.29). Multiplicative interaction tests indicated no moderation of these relative risks by race (white v. nonwhite), population density, census region, public/private school, year of school

participation, survey version, or non-complete drug responses. Levels of missing data were significantly lower for electronic tablets v. paper-and-pencil surveys.

Conclusions: Adolescent drug prevalence estimates in the US differed little across electronic tablet v. paper-and-pencil survey modes and showed little to no effect modification by sociodemographics. Levels of missing data were lower for electronic tablets.

INTRODUCTION

More than fifty countries use school-based surveys to track national trends in adolescent substance use and to evaluate national policy targeted at its reduction (1, 2). Paper-and-pencil surveys have long been the prevailing survey mode for such studies, although in recent years electronic tablets have replaced paper-and-pencil surveys in a small but growing number of countries.

Whether a transition to electronic tablets will produce a mode effect and consequent discontinuity in trends for national estimates of adolescent drug use is an open question. In general, electronic as compared to paper-and-pencil surveys tend towards slightly higher prevalence estimates of sensitive outcomes such as drug use. Evaluating the magnitude of any such mode effect for school-based studies is important to identify and adjust drug prevalence estimates.

We conducted a randomized controlled experiment to examine the potential impact of electronic tablet (7'' Samsung Tab A) vs. paper-and-pencil survey mode on national drug prevalence estimates. As described in more detail below, this experiment involved surveying a school-based, nationally-representative sample of 41,866 U.S. adolescents, in schools randomly assigned to respond using electronic tablets or paper-and-pencil.

BACKGROUND

Prevalence levels of socially undesirable behaviors such as drug use are typically higher in computer-assisted self-interviewing (CASI) compared with paper-and-pencil self-administered questionnaires (SAQ). Early studies found that computer-based surveys resulted in higher reports of adolescent drug use, sexual partners, and violence (3, 4), and computer-based surveys

also reduced the negative effect of bystanders on adolescent reporting of drug use relative to paper SAQs (5). Since these early studies, CASI or audio-CASI has become the standard for collecting data on sensitive topics in in-person surveys. This is based on the argument that computer administration affords greater privacy, both during the completion of the survey and afterwards (6).

Two recent meta-analyses support higher prevalence of reporting sensitive behaviors for computer-based as compared to paper surveys. Gnambs and Kaspar (7) found that computerized surveys led to significantly more reporting of socially undesirable behaviors than comparable paper-and-pencil surveys among adolescents, with the effect strongest for highly sensitive behaviors and surveys administered individually to respondents. More recently, Gomes et al. (8) focused more narrowly on reports on offending. Of the ten studies comparing paper SAQs with CASI, five comparisons favored CASI, only one reached statistical significance at the $\alpha=.05$ level.

School-based surveys administered in group settings have not yet produced evidence consistent with the general finding of higher prevalence on computer as compared to paper for sensitive items. A randomized study in Germany on youth delinquency led the authors to conclude computer and paper surveys produce comparable results (1 out of 15 measures differed at $p<.05$ in a sample of 2,033 students) (9). Few differences by mode were apparent in a similar, randomized, school-based study in Switzerland asking students to self-report on delinquency items (8 out of 72 differed at $p<.05$ in a study of 1,203 students) (10). Finally, a U.S., randomized study focusing on an extensive array of adolescent drug use measures in 7th, 9th, and 11th grade concluded that computer and paper surveys yielded similar outcomes (11).

Taken as a whole, the general findings regarding computer-based reporting on sensitive behaviors would lead us to expect higher drug prevalence levels for tablet as compared to paper-based surveys, but the published evidence supporting such a mode effect in school-based drug use surveys is inconclusive. It is possible that any mode effect in school surveys on drug use may be so small that it is largely inconsequential.

To examine potential mode effects, we conducted a randomized-controlled, school-based study that we designed with three main strengths. First, the sample size of this study was more than twenty times larger than existing, randomized-controlled studies on the topic. The large sample size is important to sufficiently power the study to detect potentially small effects. Second, this study measured potential mode effects on over 350 drug prevalence outcomes across three grades. The large number of outcomes allows the analysis to potentially detect any mode effects that are too small to observe for an individual drug outcome but may become apparent when pooling them. Finally, the study is nationally-representative so that the results are widely generalizable and not specific to any region or specialized group.

The specific aims of this randomized-controlled study are to (1) compare the number of positive responses to drug questions on electronic tablets as compared to paper-and-pencil surveys, (2) test the potential moderating influence on survey mode of race (white v. nonwhite), sex, population density, census region, public/private school, 1st or 2nd year of school participation, survey version, and non-complete drug responses, and (3) to compare levels of missing data across survey mode as measured by percent of respondents who completed all drug questions, the percentage of drug questions unanswered, and the mean number of missing drug items among non-completers.

METHODS

Design

Data come from the 2019 Monitoring the Future (MTF) survey, a cross-sectional survey of 8th, 10th, and 12th grade students that is nationally-representative of the 48 coterminous states in the U.S. It uses a multistage, stratified probability sampling procedure in which stage 1 is the selection of geographic areas (U.S. Metropolitan Statistical Areas (MSAs) or non-MSA counties), stage 2 is the selection of one or more middle (8th) or high (10th, 12th) schools in each area, and stage 3 is the selection of students within each sampled school. The project was approved by the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board, approval #HUM00131235. The MTF data are available for analysis at the National Addiction & HIV Data Archive Program (12).

Survey Procedures

University of Michigan personnel conducted the questionnaire administration in each school, following standardized procedures detailed in a project instruction manual. The questionnaire administrations took place in classrooms during normal class periods whenever possible; however, circumstances in some schools required the use of larger group administrations (13). Informed consent (passive or active [written], per school policy) was obtained from parents for students younger than 18 years and from students aged 18 years or older. Students were instructed to skip any questions they did not feel comfortable answering, and could change answers at any point (by erasing responses on paper-and-pencil surveys or going back and overriding answers on tablets). All responses on tablets were encrypted and then transmitted to the University of Michigan at first internet connection opportunity, which was typically immediately after survey completion. A University of Michigan staff member provided

technical assistance for tablet users, including immediately replacing faulty tablets with replacements from a reserve.

Randomization Procedure

Prior to randomization the 2019 MTF sample was hierarchically sorted in the order of grade, $1^{st}/2^{nd}$ year of school participation of target schools (each school participates in the survey for two years), and then public/private status of the target school. Within each of these groups target schools were listed by the nine U.S. Census divisions of the U.S within four levels of urbanicity. For each of these ordered lists a random start was selected (tablet or paper) and mode assignment then alternated down the list. If a target school was not successfully recruited then its replacement inherited the original school's survey mode assignment.

Participants

The final analysis pool consists of 41,866 students in 397 schools who answered at least one survey question on drug use. Within the sample of schools response rates for eligible, enrolled students were 80% in 12th grade, 86% in 10th grade, and 89% in 8th grade, with the great majority of non-response due to student absence. Response rates did not significantly differ at the $p<.05$ level for tablet- and paper- assigned schools, and were, respectively, 80.1% and 79.9% in 12^{th} grade, 86.4% and 85.1% in 10^{th} grade, and 89.0% and 89.1% in 8^{th} grade.

Measures

The analyses focus on the same set of drug prevalence outcomes that are presented in the Monitoring the Future annual volumes (14). These comprise a total of 377 drug use prevalence estimates across reporting intervals and grades. These include 90 lifetime use estimates, with 32 in 12^{th} grade and 29 in each of 10^{th} and 8^{th} grade, 127 estimates of past 12-month use, with 49 in 12^{th} grade and 39 in each of 10^{th} and 8^{th} grade, 108 estimates of past 30-day use, with 38 in 12^{th}

grade and 35 in each of 10th and 8th grade, as well as 52 estimates of heavy use, with 30 in 12th grade and 11 in each of 10th and 8th grade.

All measures, including the project's 377 drug prevalence outcomes, are described in detail in the project's annual reports and documentation (13, 14).

Sample size varies across substances. MTF uses multiple questionnaire versions or "forms" that include form-specific questions as well as a core set of questions that appear on all forms. Six different forms are used in 12th grade and four are used for the 10th and 8th grade samples. This procedure increases the number of questions that the survey can include and therefore the scope of the issues and substances covered. Each form is distributed to a randomly-selected subset of respondents within schools, so that responses to each specific form and each covered substance are nationally-representative. Table S1 in the supplemental appendix details the number of forms used to assess each drug.

Non-drug measures include self-identified white or non-white race, as measured by students who selected "White (Caucasian)" for the question "How to you describe yourself" out of a list of nine race/ethnicity options. Female was indicated by a response of "Female" to the question "What is your sex?" Population density was coded trichotomously as schools located in one of the 24 largest U.S. metropolitan statistical areas (MSA), a smaller MSA, or outside of a MSA (e.g. a rural area). Other measures include whether the school was in its first or second year of participation, census region of the country (South, Northeast, Midwest, or West), and whether a school was public or private.

Statistical Analysis

The analytic plan consists of three parts. First, the analyses used Poisson regression models (15) to examine whether the rate (proportion) of total number of "yes" responses to the

drug questions for each student was higher among those who recorded their survey answers on electronic tablets as compared to paper. The fitted Poisson regression models were stratified by the reporting intervals of lifetime, past 12-month, past 30-day, and heavy drug use. These models use the standard "exposure" variable option in Poisson regression to take into account variation in the total number of non-missing responses provided by the students, which vary by reporting interval, per each version of the survey form, and individual student levels of non-response (assuming unanswered questions are missing completely at random). Additional analyses considered potential moderating effects on survey mode using multiplicative interaction terms. Second, the analysis estimated proportions and means to compare levels of item missing data across tablet and paper modes. Finally, for each of the 377 drug prevalence outcomes the analyses present prevalence and standard error estimates for tablet-based, paper-based, and combined student responses, and also indicate whether the relative risk (RR) by survey mode of reported use of the substance differs significantly from RR=1.0 at the $p<.05$ level for a two-tailed test using a generalized linear regression model with a log link.

These analyses were not pre-registered on a publicly available platform, and results should be considered exploratory.

All analyses account for the survey's complex sample design using the "svy:" suite of commands in Stata. These commands use probability weights that generalize estimates to the U.S. national level. The Stata svy: commands provide robust estimates of standard errors of descriptive estimates and model parameters that reflect the combined effects of all levels of clustering (non-independence) of the nested sampling of students within schools and geographic areas. This population-averaged approach (16) to modeling the mode-specific relative risk of substance use reporting was chosen over the alternative generalized multi-level modeling

approach (17) because the study aims focused on estimates (and confidence intervals) for the fixed effects of survey mode and not on the random effects/components of variance associated with each specific stage of the data collection design. Results from models without probability weights differed only negligibly from weighted results and supported the same main conclusions (analyses not shown).

RESULTS

Figure 1 presents the study flow chart. Before data collection began the survey's 368 target school slots were randomly assigned to tablet or paper survey administration. The project successfully recruited and surveyed a school in 89% of these slots, of which 165 were tablet administrations and 163 were paper administrations. The final analysis pool consists of 20,985 students who recorded their survey answers on tablets and 20,881 who recorded their answers on paper-and-pencil. These sample sizes exclude the 1.5% of students who did not answer any of the drug use questions on the survey, a percentage that did not significantly differ by survey mode at the $p<.05$ level.

Table 1 presents detailed information on the distribution of the school slots by survey mode. In each of the three grades the percentage of surveyed school slots assigned to tablets ranged between 49% and 51%. Survey mode was distributed between the range of 45% to 58% for the stratification factors of $1^{st}/2^{nd}$ year of school participation, public/private school, and U.S. Census region of the country.

Table 2 presents demographic characteristics of the student sample and the percentage of each demographic group that conducted the survey on tablets. The percentage of tablet

256 administrations ranged between 45% to 57% for sex, race (white v. non-white), and population
257 density of the area in which the school was located.

258 Table 3 presents relative risk of prevalence estimates by survey mode as estimated using
259 Poisson regression models. For all reporting periods the direction of the relative risk was for
260 higher estimates on tablets, although differences were small and for all reporting intervals the
261 95% confidence interval of the relative risk included one. Specifically, the relative risk favoring
262 tablets for lifetime use was 1.03, for past 12-month use was 1.01, for past 30-day use was 1.05,
263 and for heavy use was 1.10.

264 The analysis tested the possibility that sociodemographic characteristics may have
265 moderated the size of a tablet mode effect (analyses not shown). These analyses consisted of
266 separate Poisson regression models that included indicator variables for the demographic
267 characteristics of interest, tablet mode, and their multiplicative interaction. Each of these
268 separate models was run for each of the four reporting intervals. No significant interaction with
269 survey mode was present for race (white v. nonwhite), the four-category U.S. Census region of
270 the country, or population density. Tablet mode did significantly interact with sex for analyses
271 of lifetime, past 12-month and past 30-day drug use and indicated smaller differences by survey
272 mode for females; however, in models run separately by sex there was no significant difference
273 at the $p<.05$ level by survey mode for females or for males in any reporting interval.

274 In addition, no interactions with survey mode were significant at the $p<.05$ level for
275 public/private school status or for $1^{st}/2^{nd}$ year status of school participation in the survey. No
276 interactions with survey mode were present for any of the ten separate versions, or "forms," of
277 the survey questionnaire (six in 12^{th} grade and four for the 10^{th} and 8^{th} grade samples) in any of
278 the four reporting intervals. The analyses also examined whether an indicator for missing

information on at least one drug question significantly interacted with survey mode and found that it did not for any of the reporting intervals.

Table 4 presents information on levels of item missing data across survey mode. Tablet as compared to paper mode had significantly lower levels of item missing data. For all four response categories combined the percentage of students who answered all the drug questions on their survey was 80% for tablets and 68% for paper. The percentage of unanswered drug questions was 3% on tablets and 6% for paper. The mean number of missing drug questions among students who did not answer all the drug questions was 13 for tablets and 17 for paper. Table 4 shows that tablets as compared to paper have lower levels of missing data for all these measures in each of four reporting intervals. Supplement Table S1 presents all 377 nationally-representative drug prevalence estimates calculated using (a) tablet-based responses, (b) paper-and-pencil responses, and (c) responses combined across the two modes. Acknowledging that individual responses to specific substance use questions may be highly correlated, we note that in independent t-tests, the estimates significantly differed by survey mode at the $p<.05$ level for 21 outcomes, out of which 18 estimates were higher for tablets and the other 3 were higher for paper-and-pencil. The finding of 21 significant differences amounts to 5.6% of the 377 outcomes. This may be explained by chance alone but given the non-independence of the estimates the primary conclusions in this paper are based on the regression modeling that takes into account the covariance of the drug reports and other sources of non-independence in the study design and measurements.

DISCUSSION

This study set out to examine if the transition to data collection by electronic tablets from traditional paper-and-pencil surveys would produce a mode effect and consequent discontinuity in trends for nationally-representative estimates of adolescent drug use prevalence. Analyses are based on a randomized controlled experiment, in which students in a randomly-selected half of schools recorded their survey answers on electronic tablets and in the other half of schools used paper-and-pencil. To our knowledge this is the largest study of this type, and the first with a nationally-representative sample.

The results support two major findings. First, differences in prevalence estimates were small across survey mode for the four reporting intervals of lifetime, past 12 month, past 30 day, and heavy drug use. The results are consistent with the existing literature, which led to the expectation of slightly higher prevalence levels for tablets as compared to paper. A contribution of this study is to show that this effect is quite small and likely has little to no effect on population prevalence or trend estimates.

A second major finding is that students using electronic tablets had significantly lower levels of missing data than did those using traditional paper and pencil. For all four reporting intervals the percentage of students who answered all the drug questions on their survey form was higher than 80% for tablet-based responses and about ten percentage points lower for paper-based responses. In addition, among students who did not answer all the drug questions on their survey, those who used tablets as compared to paper had fewer missing responses, for all four reporting intervals. We suspect tablets have higher completion rates lower missing data levels because answering questions on touchscreens takes less time and is perceived as more confidential than bubbling answers on paper-and-pencil optical scan sheets. These results indicate higher data quality for tablet as compared to paper administrations, an advantage for

tablets in addition to their potential to collect paradata (18-20) and also to include skip patterns that allow in-depth questions for subpopulations of specific theoretical or policy interest.

We draw two main conclusions from these results. First, any discontinuity in time trends introduced by a transition of an adolescent drug study to tablets from paper-and-pencil is expected to be negligible. This is a major consideration for many countries that are considering a transition to tablet data collection for their adolescent drug use surveillance systems. Of the many factors considered in the decision to transition to tablets, the loss of direct comparisons of prevalence estimates with previous years would be a major drawback. Such a loss would be particularly detrimental for analysis and policy formulation related to any drug prevalence that is changing rapidly, such as teen vaping in the U.S. (21, 22).

This conclusion applies not only to future transitions to tablet data collection but also to ones that have recently taken place. For the few school-based adolescent drug surveillance systems that have already made the transition to tablets without a full-scale comparison with paper these results provide a scientific rationale for direct comparison of tablet-based results with previous findings based on traditional paper and pencil (23).

A second, related conclusion is that tablet and paper modes would be expected to be largely interchangeable for school-based drug studies of adolescents. Countries that wish to transition to tablet data collection for adolescent drug surveillance may not be able to use tablets in all schools because of factors such as shipping logistics for remote regions or security issues related to bringing in valuable electronic equipment to dangerous neighborhoods. The results from this study suggest potentially high scientific validity for a hybrid study that uses electronic tablets for data collection where possible and paper surveys otherwise.

It is important to note caveats and limitations of this study. First, the finding of little to no survey mode effects for drug prevalence may not apply to other outcomes such as attitudes and beliefs, which this study did not analyze. Second, this study is limited to one country. Applicability of this study's findings to other nations would be bolstered by country-specific analyses also showing no survey mode effect; framed as confirmatory studies, such analyses would not necessarily need to be as large or as widely representative as this one. A third, related caveat is that studies of specific U.S. geographical regions and racial/ethnic groups are warranted to examine the extent to which these nationally-representative results apply. Finally, the results of this study do not include students who were unable to complete self-reported questionnaires as a result of low English proficiency or physical, sensory, or reading disabilities.

CONCLUSION

This randomly-controlled study found only very small differences across electronic tablet v. paper-and-pencil survey modes for school-based, nationally-representative estimates of adolescent drug prevalence. These results suggest that a transition to tablet-based surveys will not preclude direct comparison of drug prevalence results collected with paper and pencil. They also provide scientific justification for large drug surveillance projects to use both paper-and-pencil and tablet surveys, which may be required in some countries if shipping tablets to remote regions is infeasible or the use of electronic tablets raises security issues.

Acknowledgements: We thank John Haeussler, the project's sampling statistician, who designed and oversaw the study's randomization procedure. We also thank Virginia Laetz, who oversaw all data collection for this project and carried through the randomization procedures.

References

1. AHUMADA G., ARNEDA J. C., CLARKE P., CUMSILLE F., HYNES M. Report on Drug Use in the Americas 2019, Washington, DC: Inter-American Drug Abuse Control Commission (CICAD), Organization of American States (OAS); 2019.
2. KRAUS L., NOCIAR A. ESPAD report 2015: results from the European school survey project on alcohol and other drugs: European Monitoring Centre for Drugs and Drug Addiction; 2016.
3. TOURANGEAU R., RASINSKI K., JOBE J. B., SMITH T. W., PRATT W. F. Sources of error in a survey on sexual behavior, Journal of Official Statistics Stockholm 1997: 13: 341-366.
4. TURNER C. F., KU L., ROGERS S. M., LINDBERG L. D., PLECK J. H., SONENSTEIN F. L. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology, Science 1998: 280: 867-873.
5. AQUILINO W. S., WRIGHT D. L., SUPPLE A. J. Response effects due to bystander presence in CASI and paper-and-pencil surveys of drug use and alcohol use, Substance Use & Misuse 2000: 35: 845-867.
6. COUPER M. P., TOURANGEAU R., MARVIN T. Taking the audio out of audio-CASI, Public Opinion Quarterly 2009: 73: 281-303.
7. GNAMBS T., KASPAR K. Disclosure of sensitive behaviors across self-administered survey modes: a meta-analysis, Behavior research methods 2015: 47: 1237-1259.
8. GOMES H. S., FARRINGTON D. P., MAIA Â., KROHN M. D. Measurement bias in self-reports of offending: a systematic review of experiments, Journal of experimental criminology 2019: 15: 313-339.
9. BAIER D. Computer-assisted versus paper-and-pencil self-report delinquency surveys: Results of an experimental study, European Journal of Criminology 2018: 15: 385-402.
10. LUCIA S., HERRMANN L., KILLIAS M. How important are interview methods and questionnaire designs in research on self-reported juvenile delinquency? An experimental comparison of Internet vs paper-and-pencil questionnaires and different definitions of the reference period, Journal of Experimental Criminology 2007: 3: 39-64.
11. HALLFORS D., KHATAPOUSH S., KADUSHIN C., WATSON K., SAXE L. A comparison of paper vs computer-assisted self interview for school alcohol, tobacco, and other drug surveys, Evaluation and Program Planning 2000: 23: 149-155.
12. UNIVERSITY OF MICHIGAN. National Addiction and HIV Data Archive Program: Institute for Social Research, University of Michigan; 2019.
13. BACHMAN J. G., JOHNSTON L. D., O'MALLEY P. M., SCHULENBERG J. E., MIECH R. A. The Monitoring the Future Project after Four Decades: Design and Procedures. Occasional Paper #82, Ann Arbor, MI: Institute for Social Research; 2015.
14. MIECH R. A., JOHNSTON L., O'MALLEY P. M., BACHMAN J. G., SCHULENBERG J. E., PATRICK M. E. Monitoring the Future National Survey Results on Drug Use, 1975-2019: Volume I, Secondary School Students, Ann Arbor, MI: Institute for Social Research; 2020.
15. WINKELMAN R. Poisson Regression. Econometric Analysis of Count Data, Fifth Edition, Berlin: Springer; 2008, p. 63-126.
16. HEERINGA S. G., WEST B. T., BERGLUND P. A. Applied Survey Data Analysis, 2nd Edition Boca Raton, FL: Chapman & Hall/CRC; 2017.
17. SNIJDERS T. A., BOSKER R. J. Multilevel analysis: An introduction to basic and advanced multilevel modeling: Sage; 2011.
18. COUPER M. P., KREUTER F. Using paradata to explore item level response times in surveys, Journal of the Royal Statistical Society: Series A (Statistics in Society) 2013: 176: 271-286.

19. McClain C. A., Couper M. P., Hupp A. L., Keusch F., Peterson G., Piskorowski A. D. et al. A typology of web survey paradata for assessing total survey error, Social Science Computer Review 2019: 37: 196-213.
20. Kreuter F. Improving surveys with paradata: Analytic uses of process information: John Wiley & Sons; 2013.
21. Miech R. A., Johnston L., O'Malley P. M., Bachman J. G., Patrick M. E. Trends in Adolescent Vaping from 2017 to 2019 -- U.S. National Estimates, New England Journal of Medicine 2019: 1490-1491.
22. Miech R. A., Patrick M. E., O'Malley P. M., Johnston L. D., Bachman J. G. Trends in reported marijuana vaping among U.S. adolescents, 2017-2019, JAMA 2019: 323: 475-476.
23. Cullen K. A., Gentzke A. S., Sawdey M. D., Chang J. T., Anic G. M., Wang T. W. et al. E-Cigarette use among youth in the United States, 2019, Jama 2019: 322: 2095-2103.

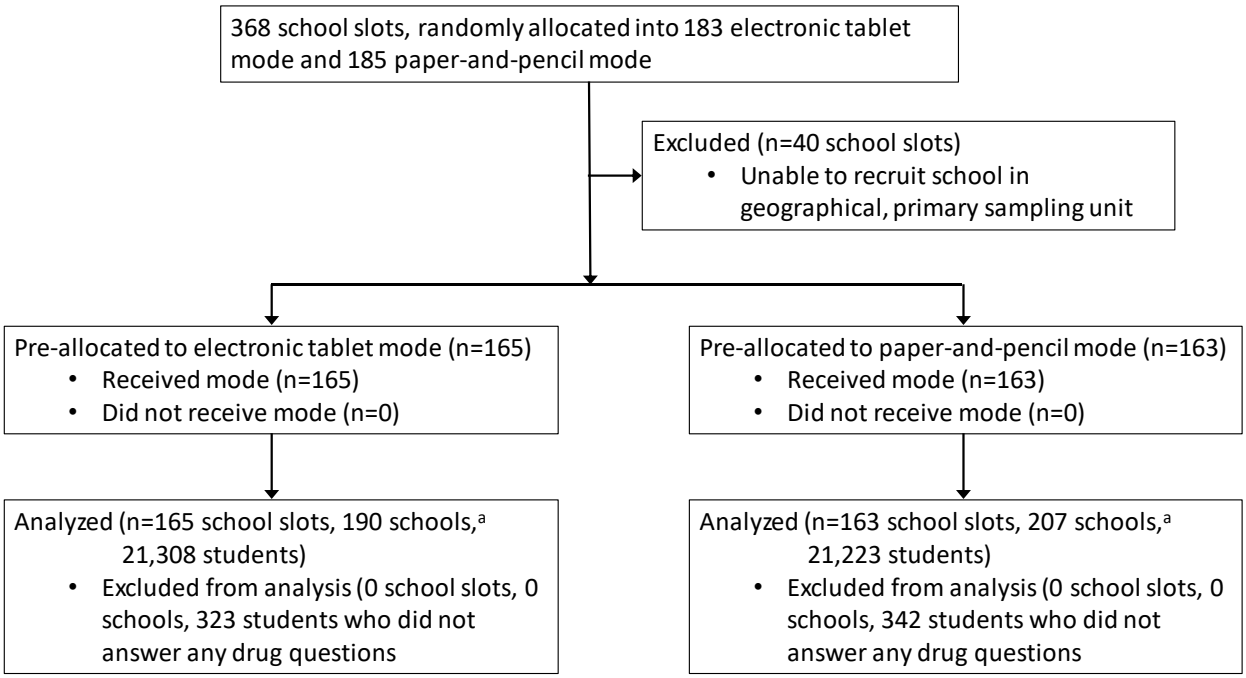

Figure 1: Study Flow Chart

[a]More than one school was surveyed in some school slots in order to meet project requirements for minimum number of students surveyed. All schools within the same school slot were surveyed using the same pre-assigned survey mode (i.e. tablet or paper-and-pencil).

441

442

Table 1: Number of Target and Survey School Slots, by Selected Characteristics (% Tablet Mode in Parentheses)

	8th Grade	10th Grade	12th Grade
Target Sample			
School slots	140 (49%)	114 (50%)	114 (50%)
Surveyed Sample			
School slots	124 (51%)	102 (49%)	102 (51%)
Stratification Factors for Randomization			
# First Year Schools	49 (49%)	47 (51%)	48 (50%)
# Second Year Schools	75 (52%)	55 (47%)	54 (52%)
# Public	103 (50%)	86 (49%)	88 (50%)
# Private	21 (52%)	16 (50%)	14 (57%)
South	42 (55%)	33 (52%)	36 (47%)
Northeast	29 (45%)	22 (45%)	20 (50%)
Midwest	31 (48%)	26 (50%)	26 (58%)
West	22 (55%)	21 (48%)	20 (50%)

443

444

445

Table 2: Demographic Characteristics of Sample (% Tablet Mode in Parentheses)

	8th Grade	10th Grade	12th Grade
% Female	50 (52%)	52 (50%)	51 (54%)
% Male	50 (53%)	48 (51%)	49 (55%)
% White, Non-Hispanic	47 (48%)	47 (56%)	52 (49%)
% Nonwhite	53 (55%)	53 (45%)	48 (57%)
% Large metropolitan statistical area	32 (43%)	31 (43%)	33 (50%)
% Medium metropolitan statistical area	50 (54%)	51 (56%)	48 (54%)
% Non-metropolitan statistical area	19 (57%)	18 (45%)	19 (53%)
Number of schools[a]	143 (49%)	126 (46%)	128 (46%)
Total number of respondents[b]	14,032 (51%)	14,439 (50%)	13,395 (53%)

[a]More than one school was surveyed in some school slots in order to meet project requirements for minimum number of students surveyed. All schools within a school slot were surveyed using the same pre-assigned survey mode (i.e. tablet or paper-and-pencil).

[b]Number of respondents slightly smaller for sex and race samples due to missing responses. Number of respondents with non-missing values for sex in 8th, 10th, and 12th grade was 13,495, 14,044, and 12,644, respectively. Number of respondents with non-missing values for race was 13,368, 14,022, and 12,828, respectively. The total number of students reported here does not exclude the 1.5% who did not answer any drug questions.

Table 3: Relative Risk for Drug Use by Reporting Interval, Results from Poisson Regression (95% Confidence Intervals in Parentheses)

	Lifetime (n=41,839)	Past 12 Months (n=41,720)	Past 30 Days (n=41,843)	Heavy Use (n=41,764)
Tablet v. Paper Mode	1.03 (0.93-1.14)	1.01 (0.91-1.11)	1.05 (0.93-1.20)	1.10 (0.93-1.29)

462

Table 4: Measures of Non-Response on Survey Items, by Mode

	------- Tablet -------		------- Paper -------	
	Estimate	95% C.I.	Estimate	95% C.I.
Overall				
% Completed all drug questions ($\bar{n}$=92)[a]	79.64	(78.19-81.02)	67.99	(65.92-70.00)
Percent drug questions unanswered	2.86	(2.52-3.20)	6.26	(5.59-6.92)
Mean # missing drug items, among non-completers	12.56	(11.58-13.54)	17.28	(16.14-18.41)
Lifetime Use				
% Completed all drug questions ($\bar{n}$=22)[a]	83.88	(82.54-85.14)	73.63	(71.66-75.52)
Percent drug questions unanswered	2.44	(2.15-2.73)	5.52	(4.91-6.14)
Mean # drug items missing, among non-completers	3.52	(3.27-3.76)	4.86	(4.54-5.17)
Past 12-Month Use				
% Completed all drug questions ($\bar{n}$=25)[a]	82.80	(81.46-84.07)	73.01	(71.14-74.80)
Percent drug questions unanswered	3.12	(2.70-3.54)	6.46	(5.78-7.13)
Mean # missing drug items, among non-completers	5.12	(4.68-5.55)	6.63	(6.26-7.00)
Past 30-Day Use				
% Completed all drug questions ($\bar{n}$=23)[a]	83.38	(82.06-84.62)	72.09	(70.09-74.00)
Percent drug questions unanswered	2.54	(2.23-2.85)	5.89	(5.25-6.53)
Mean # missing drug items, among non-completers	3.83	(3.55-4.12)	5.28	(4.96-5.60)
Heavy Use				
% Completed all drug questions ($\bar{n}$=22)[a]	86.02	(84.83-87.13)	77.59	(75.84-79.25)
Percent drug questions unanswered	3.15	(2.83-3.47)	6.20	(5.61-6.78)
Mean # missing drug item, among non-completers	2.88	(2.67-3.09)	3.55	(3.30-3.80)

463 Sample size=41,866 for overall analyses, 41,839 for lifetime analyses, 41,720 for past 12 month analyses, 41,843 for past 30 day analyses, and
464 41,819 for analyses of heavy use.
465 [a]Total number of drug questions varies across different versions of the survey questionnaire.
466 Note: All comparisons across table and paper are statistically significant at the .05 level, as indicated by non-overlapping 95% confidence
467 intervals.