TECHNICAL PAPERS ON
HEALTH AND BEHAVIOR MEASUREMENT

TECHNICAL PAPER 68

Reducing Bias in Telephone Survey Estimates of the Prevalence of Drug Use
A Randomized Trial of Telephone Audio-CASI

Charles F. Turner, Maria A. Villarroel, Susan M. Rogers ,
Elizabeth Eggleston, Laxminarayana Ganapathi, Anthony M. Roman, Alia Al-Tayyib

Reference Citation

Turner CF, Villarroel MA, Rogers SM, Eggleston E, Ganapathi L, Roman AM, Al-Tayyib AA. (2002) Reducing Bias in Telephone Survey Estimates of the Prevalence of Drug Use: A Randomized Trial of Telephone Audio-CASI. *Addictions,* 100:1432-1444.

Reducing bias in telephone survey estimates of the prevalence of drug use: a randomized trial of telephone audio-CASI

Charles F. Turner[1,2], Maria A. Villarroel[1], Susan M. Rogers[1,3], Elizabeth Eggleston[1], Laxminarayana Ganapathi[4], Anthony M. Roman[5] & Alia Al-Tayyib[1,6]

Program in Health and Behavior Measurement, Research Triangle Institute, Washington DC[1], CUNY, Queens College and Graduate Center, New York[2], Department of Maternal and Child Health, School of Public Health, University of North Carolina at Chapel Hill[3], Research Computing Division, Research Triangle Institute, Washington DC[4], Center for Survey Research, University of Massachusetts at Boston[5], Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill[6]

Correspondence to:
Charles F. Turner
Program in Health and Behavior Measurement
Research Triangle Institute
1615 M Street, NW
Washington DC 20036
USA
Tel. 202-728-2493
Fax. 202-728-2047
E-Mail: cft@rti.org

Submitted 2 August 2004;
initial review completed 21 January 2005;
final version accepted 8 April 2005.

ABSTRACT

Aim To assess the impact of telephone audio computer-assisted self-interviewing (T-ACASI) on reporting of alcohol use, alcohol problems and illicit drug use in telephone surveys of the general population. Prior research suggests that illicit drug use is underreported in traditional, interviewer-administered, telephone surveys.

Design Randomized experiment embedded in telephone survey of probability samples of populations of USA and Baltimore, MD. Survey respondents were randomly assigned to be interviewed either by human telephone interviewers or by T-ACASI after household screening, recruitment, and informed consent procedures were completed.

Setting Respondents were interviewed by telephone in their homes.

Participants Probability samples of 1543 English-speaking adults ages 18–45 residing in telephone-accessible households in USA and 744 similarly defined adults residing in Baltimore, MD, USA.

Measurements Nine questions on alcohol, marijuana, cocaine, and injection drug use adapted from 1994 NHSDA and four CAGE questions on alcohol problems. Crude odds ratios and odds ratios controlling for demographic factors calculated to test for differences between responses obtained by T-ACASI and human interviewers.

Findings T-ACASI had mixed effects on reporting of alcohol use, but it did increase reporting of one of four CAGE alcohol problems: feeling guilty about drinking (23.0% in T-ACASI vs. 17.6% in T-IAQ, OR = 1.4, $P < 0.01$). T-ACASI also obtained significantly more frequent reporting of marijuana, cocaine, and injection drug use. The impact of T-ACASI was most pronounced for reporting of recent use of 'harder' drugs. Thus T-ACASI respondents were more likely to report marijuana use in the past month (10.0% vs. 5.7%, crude OR = 1.9, $P < 0.001$), cocaine use in the past month (2.1% vs. 0.7%, crude 3.2, $P < 0.001$) and injection drug use in the past five years (1.6% vs. 0.3%, crude OR = 4.8, $P < 0.01$).

Conclusions Telephone survey respondents were more likely to report illicit drug use and one alcohol problem when interviewed by T-ACASI rather than by human telephone interviews.

KEYWORDS Alcohol use, illicit drug use, interactive voice response (IVR), measurement bias, NHSDA, NSDUH, population surveys, self-report, telephone audio computer-assisted self-interviewing (T-ACASI), telephone surveys, validity.

 doi:10.1111/j.1360-0443.2005.01196.x

INTRODUCTION

Despite the growing use of telephone surveys in studies of alcohol and drug use (e.g. Hall *et al.* 1991; Kendler *et al.* 1999; Perneger *et al.* 2001; Saxe *et al.* 2001; Midanik & Greenfield 2003; Kerr *et al.* 2004), convincing evidence exists that traditional telephone surveys – which rely upon human interviewers – obtain lower reported prevalences of illicit drug use than in-person surveys. This appears to occur because in-person surveys can employ modes of response that do not require respondents to divulge their drug use to a human interviewer. (For an overview of other divergences in the measurement characteristics of telephone and in-person surveys see de Leeuw & van der Zouwen 1988; Sykes & Collins 1988). Results for reporting of alcohol consumption are mixed, and this may reflect the lesser sensitivity of divulging information on this licit behavior to a human interviewer.

Several major studies (Aquilino 1992, 1994; Gfroerer & Hughes 1992) have compared estimates of the prevalence of licit and illicit drug use obtained from interviewer-administered telephone surveys (T-IAQ) to estimates obtained from in-person surveys using paper self-administered questionnaires (paper SAQs). These studies found significantly and substantially higher estimates of the prevalence of illicit drug use when respondents recorded their drug use on a paper SAQ rather than reporting it to a human telephone interviewer. Gfroerer and Hughes found, for example, that compared to a telephone survey, in-person administration of paper SAQs yielded estimated drug use prevalences that ranged from 33% higher for reporting any lifetime use of marijuana to 121% higher for reporting use of cocaine in the past 12 months. While Gfroerer and Hughes do not report results for the reporting of alcohol use, Aquilino (1994; p. 230) has demonstrated that the impact of interview mode (T-IAQ vs. in-person IAQ vs. in-person SAQ) is smallest for reporting of alcohol use.

Such results are usually interpreted as evidence of a negative reporting bias arising from the requirement that respondents in telephone surveys reveal their drug use to a human interviewer. While other interpretations are possible, results from the 1990 National Household Survey of Drug Abuse (NHSDA) field experiment support this interpretation (Turner *et al.* 1992). In this experiment, 3326 respondents were randomly assigned to complete NHSDA questionnaires using either paper SAQs or interviewer administered questionnaires (IAQs). Measurements made with paper SAQs yielded higher estimates of the prevalence of illicit drug use than those made using IAQs. Moreover, the relative advantage of paper SAQs in encouraging more complete reporting of drug use was a function of the expected sensitivity of the behavior being reported. Thus no advantage was found for reporting of adult alcohol use. Among illicit drugs, 'harder' drugs and reports of recent use were most affected. So, for example, the mode effect for reporting of cocaine use in the past month was 2.4 (ratio of estimated prevalences: SAQ/IAQ), but it was imperceptible for reporting of any marijuana use during the respondent's lifetime (prevalence ratio = 1.05).

Until recently, telephone surveys could not provide privacy similar to the paper SAQs used for measurement of sensitive topics during in-person surveys. But today, telephone audio computer-assisted, self interviewing (T-ACASI) offers researchers the option of providing a similarly private interview context for telephone survey measurements of sensitive behaviors. In a T-ACASI survey, a human interviewer screens and recruits eligible subjects. However, once the subject has been recruited and has completed nonsensitive portions of the survey, the phone call is transferred to the T-ACASI system and computer-controlled, pre-recorded questions are read to the subject. The subject provides responses by pressing keys on a touch-tone telephone. (For further technical details and discussions of the feasibility and costs of T-ACASI, see Turner *et al.* 1998a; Cooley *et al.* 2000; Corkey & Parkinson 2002).

T-ACASI technology may offer a unique and important opportunity to reduce the biases that afflict telephone survey measurements of drug use. By providing complete privacy to respondents, we hypothesize that T-ACASI could reduce the demonstrated biases in reporting of illicit drug use in telephone surveys. Findings from a small pilot study and a large specialized study of men who have sex with men have been encouraging (Turner *et al.* 1996; Gribble *et al.* 2000), but they do not provide convincing evidence of the impact of this technology on reporting of drug use in surveys of the general population.

NSBM experiment

The 1999–2000 National STD and Behavior Measurement Experiment (NSBME) surveyed probability samples of U.S and Baltimore adults ages 18–45 and randomly assigned respondents to answer a wide range of sensitive questions asked either by a human telephone interviewer or by T-ACASI. In this article we contrast NSBME T-ACASI and T-IAQ measurements of the prevalence of licit and illicit drug use, and we compare these estimates to comparable measurements made by the in-person 2000 National Household Survey of Drug Abuse.

NSBME sample strata

The NSBME applied the same protocol to two separate samples of US and Baltimore residents. [The Baltimore

sample strata was drawn to permit extensions of our prior work on the 1997–98 Baltimore STD and Behavior Survey (Turner *et al.* 2002)]. Statistical analyses of the NSBM experiment treat the combined national and Baltimore sample strata as a population that has been randomly assigned to one of two conditions (T-IAQ or T-ACASI). Our analyses assess the probabilities that differences in the measurements obtained in each condition could have arisen by chance from the random assignment of subjects to the two experimental conditions. Our *experimental* analyses do not weight the Baltimore sample strata to reflect its relative size compared to the national population (<1%). Thus the crude odds ratios we derive to represent the impact of survey mode could be disproportionately influenced by the Baltimore sample strata—if the impact of survey mode was substantially different for Baltimore households than for other US households. As a protection against this possibility, all crude odds ratios representing mode effects were tested for heterogeneity across the two sample strata, that is, we explicitly tested the null hypothesis that the estimated effect of survey mode on reporting of drug use was equivalent across survey strata.

METHODS

The protocol for this research was approved and supervised by Institutional Review Boards for the Protection of Human Subjects at the Research Triangle Institute (RTI) and the University of Massachusetts.

Sample design

The NSBME experiment was embedded in a telephone survey of a probability sample of women and men aged 18–45 years residing in U.S. households with working land line telephones. The survey was conducted between September, 1999 and April, 2000. Two sample strata were recruited for this survey measurement experiment: (1) a sample of the telephone-accessible U.S. household population aged 18–45 (national stratum), and (2) a parallel sample of the telephone-accessible population of the city of Baltimore, MD (Baltimore stratum). A list-assisted random digit dialed (RDD) sample was drawn for each stratum using the Genesys Sampling System (2002; see also, Kulp 1994; Brick *et al.* 1995). Sampling procedures are briefly described below; additional details of the NSBME sample design, execution, and weighting have been presented by Villarroel *et al.* (in press).

Sample execution

For the national stratum, 14,250 telephone numbers were generated, and 12,322 telephone numbers (86.5%) were successfully screened for eligibility. 2183 of these screened telephone numbers were found to be residential numbers with one or more eligible English-speaking respondents aged 18–45. One eligible household member of these households was randomly selected for participation in the survey (without substitution). 1452 of the 2183 target respondents in the national strata completed interviews (66.5%), and 91 respondents (4.2%) completed partial interviews that included at least one substantive questionnaire section. Up to a maximum of 91 calls (mean = 11.3) were made to screen households and complete an interview in the national stratum.

The second sample stratum was drawn to represent the adult population of Baltimore, MD. For the Baltimore stratum, 7498 telephone numbers were generated and 6326 (84.4%) were successfully screened for eligibility. Screening identified 1072 households with an eligible respondent, and 697 of these eligible respondents completed interviews (65.0%). An additional 47 respondents (4.4%) completed partial interviews. Up to a maximum of 82 calls (mean = 12.3) were made to screen households and complete an interview in the Baltimore stratum.

Interview modes

Telephone numbers were randomly assigned to the T-IAQ or T-ACASI conditions prior to their release to the telephone survey unit. Following screening and recruitment into the study, telephone interviewers at the Center for Survey Research (University of Massachusetts, Boston) conducted the survey either by asking the respondent questions and recording their answers (T-IAQ condition) or by transferring the respondent to a T-ACASI system developed at RTI by Cooley (see Cooley *et al.* 2000).

Measurements

Survey questions on use of alcohol, marijuana, cocaine and injection drugs were adapted from the 1994 National Household Survey of Drug Abuse (version B; SAMSHA 1996; ICPSR 2004). The CAGE questions (Ewing 1984) were administered to assess problems with alcohol consumption. (Question wordings are presented in Appendix 1).

Ensuring sample equivalence across interview modes

While randomization should ordinarily yield approximately equivalent groups, the T-ACASI condition had a higher survey break-off and lower response rate than the T-IAQ condition—particularly in the national sample strata. In addition, some minor divergences in sample

composition might be expected because the T-IAQ sample included approximately three percent of cases that did not have a touchtone phone. While these cases are excluded from the T-ACASI sample in our analysis, they could not be excluded from the T-IAQ sample. Tabulations from the 2000 NHSDA indicate, however, that estimates of the prevalence of drug use in telephone-accessible households vary almost imperceptibly depending on whether non-touchtone (non-TT) households are included or excluded from the sample [e.g. NHSDA estimates of drug use in the past 30 days for 18–49 years olds are: alcohol, 55.2% (excluding non-TT households) and 55.1% (including non-TT households); marijuana, 6.4% in both samples; and cocaine, 0.7% in both samples.]

We have previously tested the demographic equivalence of the NSBME's T-IAQ and T-ACASI samples and found small and statistically insignificant variations across experimental conditions (Villarroel *et al.* in press). Nonetheless, to control for the potential impact of fluctuations in sample composition across experimental conditions our analyses use both crude and adjusted odds ratios (ORs). The adjusted ORs are estimates of mode effects derived from logistic regression models incorporating controls for nine sociodemographic factors (sample strata, gender, age group, educational attainment, marital status, living with children in the household, race-ethnicity, region of country, and level of urbanization; Table 1 provides additional details about these control variables.)

Statistical analysis

Our analyses are intended to answer four questions. First, does the mode of interview affect the likelihood that a respondent will report alcohol or illicit drug use? Second, is the impact, if any, of interview mode homogeneous across subpopulations defined by gender, ethnicity, place of residence, etc., or are some subpopulations particularly sensitive to interview mode? Third, is the pattern of reported drug use in each measurement condition consistent with expectations based upon past research, e.g. do the vast majority of 'hard' drug users also report experience with 'softer' drugs? and finally, if T-ACASI appears to provide more complete reporting of drug use, how does the estimated prevalence of drug use in the USA obtained in the NSBME compare with national estimates obtained by more expensive in-person surveys such as the National Survey of Drug Use and Health (NSDUH; formerly called the National Household Survey of Drug Abuse).

To address our first research question, we combine the national and Baltimore sample strata as described previously. Data in these analyses are unweighted, and our statistical analyses assess the likelihood that observed fluctuations in survey responses across the two experimental conditions (T-ACASI vs. T-IAQ) arose by chance from the random allocation of respondents to one of these conditions. When this null hypothesis is rejected, we conclude that the mode of interview had an effect on survey response.

Log-linear modeling (Goodman 1968, 1978) was subsequently used to test whether the impact of T-ACASI on the reported prevalence of drug use was equivalent across subpopulations defined by sample strata (Baltimore vs. National), gender, race, region of residence and other social and demographic factors. Our analyses fit log-linear models to 3-way tables of: Drug Use (*D:* Yes or No) by Mode of Interview (*M:* T-IAQ or T-ACASI) by Sociodemographic Factor (*F*, for example, Male or Female). Log-linear models were constrained to fit all 2-way marginals [(D*M*) (DF) (MF)] in these 3-way tables, and likelihood ratio chi-square statistics were calculated for the fit of these models to the observed data. Failure to obtain a statistically adequate fit indicated the presence of a statistically significant 3–way interaction. This is to say that the impact of T-ACASI on reporting of drug use was not homogeneous across subpopulations defined by the sociodemographic factor (e.g. gender). Models that include sample strata provide a test for the equivalence of the estimated mode effects across the Baltimore and national sample strata.

As a validity test, we subsequently calculate the proportion of self-reported injection drug users who report experience with cocaine and marijuana and the proportion of cocaine users who report experience with marijuana. If our measurements are equally valid, we would expect that in both measurement conditions, substantial majorities of self-reported users of 'hard' drugs would also report use of 'softer' drugs.

In our final analysis, we calculate weighted national estimates of the prevalence of drug use for US adults interviewed in the 1999–2000 NSBME (national sample stratum only) using T-ACASI and T-IAQ modes. These prevalence estimates are compared to estimates derived from the 2000 NHSDA public use dataset, restricting the NHSDA sample to persons who reported that (1) they had a touchtone telephone they received calls on, and (2) they did not reside in group quarters. For these analyses, sample weighting is applied to all observations. NHSDA tabulations use the weighting variable 'analwt_c' from the NHSDA public use dataset (DHHS 2004). This weighting variable incorporates adjustments for varying probabilities of selection and non-response plus poststratification adjustment to match Census state-level estimates of the target population by age, race, gender, and Hispanic origin (Chen *et al.* 2002). NSBME tabulations use weights constructed to reflect the varying probabilities of selection, non-response and poststratification adjustments to

Table 1 Response distributions and reported prevalence of alcohol use obtained by Telephone Audio-CASI and by human telephone interviewers (T-IAQ) in the 1999–2000 NSBME.

Response distribution				*Alcohol use prevalence*				
Measurement	*T-IAQ*	*T-ACASI*	*Time period*	*T-IAQ*	*T-ACASI*	*CRUDE OR*[a]	*P Strata by Mode Interaction*	*Adjusted Or*[a,b]
Ever drank alcohol	93.2%	91.8%	**Lifetime Use**					
Unweighted n	1,214	1,024	*Unweighted n*	93.2%	91.8%	0.81	>0.5	0.75
P for mode effect	0.19			1,214	1,024			
Drinking in past 30 days			**Drinking in past 30 days**					
0 days[c]	41.3%	40.0%						
1– 4 days	33.3%	37.1%	1+ days	58.7%	60.0%	1.05	0.28	1.09
5–9 days	13.7%	12.3%	5+ days	25.4%	22.9%	0.87	0.20	0.87
10–15 days	5.5%	5.3%	10+ days	11.7%	10.5%	0.89	0.26	0.91
16+ days	6.2%	5.3%	16+ days	6.2%	5.3%	0.84	>0.5	0.89
Unweighted n	1,213	1,024	*Unweighted n*	1,213	1,024			
P for mode effect[e]	0.48							
Heavy drinking in past 30 days			**Heavy drinking in past 30 days**					
0 days[c]	76.5%	80.0%						
1–4 days	15.5%	14.6%	1+ days	23.5%	20.0%	0.81*	0.44	0.79*
5–9 days	5.0%	3.5%	5+ days	8.0%	5.5%	0.66*	0.06	0.65*
10–15 days	1.8%	1.5%	10+ days	3.0%	2.0%	0.65	>0.5	0.63
16+ days	1.2%	0.5%	16+ days	[d]	[d]	[d]	[d]	[d]
Unweighted n	1,211	1,023	*Unweighted n*	1,211	1,023			
P for mode effect[e]	0.011							

Analysis of unweighted data from combined national and Baltimore sample strata of the 1999–2000 NSBME. Analysis excludes 39 respondents assigned to T-ACASI who were interviewed in T-IAQ because they did not have a touchtone telephone.

*$P < 0.05$.

[a]ORs contrast prevalence in T-ACASI condition to prevalence in T-IAQ condition.

[b]Adjusted ORs for mode effect derived from logistic regression model controlling for sample strata (National; Baltimore), gender (male; female); age group (18–25; 26–35; 36–45), educational level (< H.S.; H.S. graduate or equivalent; some college or trade school; 4-year college graduate and higher) , marital status (married; cohabiting; divorced, separated or widowed; never married), living with children in the household (yes; no), race-ethnicity (Hispanic; non-Hispanic black, non-Hispanic white; non-Hispanic other) region of country (North East and Middle Atlantic states; North Central; South Atlantic; South Central; Mountain; Pacific), and level of urbanization (21 largest MSAs; and county sizes: <20,000 households; 20,000 to 85,000 households; >85,000 households). All respondents in Baltimore sampling strata were coded as residing in the South Atlantic and in one of the 21 largest MSAs [Baltimore-Washington MSA].)

[c]Includes all respondents who provided a substantive answer. So, for example, persons who reported that they had *never* drank an alcoholic beverage are included in the denominators of both estimates even though they were not asked the specific questions about alcohol consumption in the past 30 days and heavy drinking.

[d]Not calculated because of sparse number of respondents reporting this level of heavy drinking. Baltimore sample strata, for example, had no cases in either interview condition reporting drinking five of more drinks on 16+ days.

[e]*P* value for Mantel-Haenszel test for linear association.

match Census estimates of population distribution by age, race, and gender (Villarroel *et al.* in press).

Statistical analyses were carried out using STATA, versions 6.0 and 8.0 (StataCorp 2000; StataCorp 2003) and SPSS, version 6.0.1 (1993). Our national estimates of the prevalence of drug use (Table 6) employ weighting and use the 'svy' STATA algorithms appropriate for weighted survey data. Other statistical analyses estimate the impact of survey mode by treating the unweighted sample (pooled Baltimore and national strata) as a population that has been randomly assigned to one of two experimental conditions. Tables 1–3 include tests of the homogeneity of estimates of the impact of survey mode across the two sample strata. As a convention in our discussion of statistical results, we use the phrases 'statistically reliable' and 'not statistically reliable' to identify instances in which a null hypothesis can or cannot be rejected with $P < 0.05$.

RESULTS

Alcohol use

Table 1 presents response distributions and the reported prevalence of different levels of alcohol use obtained in the NSBME's T-ACASI and T-IAQ interview modes. It will be seen that there are few statistically reliable differences in reported prevalences across modes. No statistically reliable differences are found for reported frequency of drinking in the past 30 days. Reports of 'heavy drinking'

(i.e. 5 or more drinks on one occasion) were more common when human interviewers were conducting the interviews rather than T-ACASI (adjusted $OR_s = 0.63$–0.79). The response distributions obtained in the two conditions were significantly different ($P = 0.011$ for test of linear trend), and two of the differences in reported prevalences were statistically reliable ($P < 0.05$).

Table 2 presents parallel results for the reported prevalence of alcohol problems obtained using the CAGE questions. These questions elicit a binary response indicating whether the respondent had ever experienced a particular problem with their alcohol use. It will be seen that two of the four CAGE questions produce notable differences when administered by T-ACASI rather than a human interviewer. The odds that a respondents would report feeling guilty about drinking were 1.42 times higher (adjusted OR; $P < 0.01$) when the respondent was interviewed by T-ACASI rather than a human interviewer. A parallel result is obtained for respondents reporting of taking 'a drink first thing in the morning to steady your nerves or get rid of a hangover' although the result is not statistically reliable (adjusted OR = 1.48; p = 0.093). Summing the number of CAGE symptoms reported in the two conditions, we find that slightly more symptoms were reported in the T-ACASI condition, although the result is not statistically reliable (means = 0.59 vs. 0.51, p = 0.13). If we use the criterion that reporting two or more CAGE symptoms is clinically noteworthy, no difference is found between results obtained by the two interview modes.

Illicit drug use

Table 3 presents the response distributions and the reported prevalence of illicit drug use over various periods of time. For marijuana and cocaine use, we have calculated prevalences for use in the past month, past year, past three years, and lifetime. Because injection drug use was very infrequently reported, we calculated prevalences for use in the past year, past five years, and lifetime.

These tabulations show statistically reliable and frequently large differences in the reporting of illicit drug use in the two survey conditions. In all instances, respondents were more likely to report drug use in the T-ACASI condition than when responding to a human telephone interviewer (T-IAQ). The adjusted odds ratios indicate that T-ACASI was 2.0–3.8 times more likely to elicit a report of marijuana or cocaine use in the past 30 days and 10.3 times more likely to elicit a report of injection drug use in the past five years than human interviewers. It is noteworthy that the advantage of T-ACASI—as reflected in the ORs—increases with the presumed seriousness of the type of drug use (i.e. Injection > Cocaine > Marijuana) and with the recency of use reported (e.g. past month > past year > lifetime).

Subpopulation variation in impact of T-ACASI

Previous NSBME research has suggested that the impact of private, automated interviewing technology may vary across subpopulations. Villarroel *et al.* (2002, in press) found, for example, that the impact of T-ACASI on report-

Table 2 Prevalence of CAGE problems reported to Telephone Audio-CASI and to human telephone interviewers (T-IAQ) in the 1999–2000 NSBME.

CAGE Alcohol Problems[a,b]	*T-IAQ*	*T-ACASI*	*CRUDE OR*[c]	*P Strata by Mode Interaction*	*Adjusted OR*[c]
Felt you should cut down on drinking	22.5%	22.2%	0.98	>0.5	0.97
Annoyed by criticism of your drinking	6.9%	7.7%	1.13	0.08	1.25
Felt guilty about drinking	17.6%	23.0%	1.40**	>0.5	1.42**
Drink in morning	4.3%	5.9%	1.39	>0.5	1.48
Unweighted n	*1,202–1,208*	*1,020–1,023*			
CAGE scores of 2+	14.0%	16.1%	1.19	>0.5	1.22
Mean CAGE Numerical Score	0.51	0.59			
Unweighted n	*1,199*	*1,019*			
P for mode effect[d]	0.13				
P for mode by strata interaction[d]	>0.5				

Analysis of unweighted data from combined national and Baltimore sample strata of the 1999–2000 NSBME. Analysis excludes 39 respondents assigned to T-ACASI who were interviewed in T-IAQ because they did not have a touchtone telephone.

**$P < 0.01$.

[a]Each CAGE question is a dichotomy (*yes* or *no*) so—unlike recency of reported alcohol use—response distribution is identical to prevalence estimate, i.e., *yes* = prevalence estimate, and *no* = 100% minus prevalence estimate.

[b]Persons who reported *never* drinking alcohol were not asked the alcohol problems questions and have been presumptively coded as 'no' for each question on alcohol problems.

[c]ORs contrast prevalence in T-ACASI condition to prevalence in T-IAQ condition. See Table 1 for description of control variables used in calculating the adjusted ORs.

[d]*P*-Values were derived from 2-way analysis of variance with interview method and sample strata as factors; the dependent variable was the number of CAGE problems reported (0 to 4). Analysis used anova procedure from STATA version 8 using partial sums of squares option.

Table 3 Response distributions and reported prevalence of illicit drug use for different time periods obtained by Telephone Audio-CASI and by human telephone interviewers (T-IAQ) in the 1999–2000 NSBME.

Response distribution				*Alcohol use prevalence*				
Measurement	*T-IAQ*	*T-ACASI*	*Time period*	*T-IAQ*	*T-ACASI*	*CRUDE OR*[d]	*P Strata by Mode Interaction*	*Adjusted OR*[d]
Marijuana use			Marijuana use					
In last 30 days[a]	5.7%	10.0%	Past month	5.7%	10.0%	1.86***	>0.5	2.03***
1 month to 1 year ago	5.7%	8.2%	Past year	11.3%	18.3%	1.75***	>0.5	1.97***
1–3 years ago	5.5%	6.3%	Past 3 years	16.8%	24.5%	1.61***	>0.5	1.81***
>3 years ago	31.5%	29.6%	Lifetime	48.3%	54.2%	1.27**	>0.5	1.26*
Never	51.7%	45.8%						
Unweighted n	*1203*	*1019*						
P for mode effect[c]	<0.001							
Cocaine use			Cocaine use					
In last 30 days[a]	0.7%	2.1%	Past Month	0.7%	2.1%	3.15**	>0.5	3.75**
1 month to 1 year ago	2.2%	2.7%	Past year	2.8%	4.8%	1.74*	0.38	1.79*
1–3 years ago	1.6%	3.5%	Past 3 years	4.4%	8.3%	1.98***	>0.5	2.08***
>3 years ago	13.5%	13.8%	Lifetime	17.9%	22.1%	1.31*	>0.5	1.31*
Never	82.1%	77.9%						
Unweighted n	*1208*	*1021*						
P for mode effect[c]	<0.001							
Drug injection			Drug injection					
In last year[a,b]	0.1%	1.3%	Past year	0.1%	1.3%	15.59**	0.09	22.2**
In last 5 years	0.2%	0.3%	Past 5 years	0.3%	1.6%	4.80**	>0.5	10.3**
In lifetime	1.5%	1.7%	Lifetime	1.8%	3.2%	1.80*	0.11	2.1*
Never	98.2%	96.8%						
Unweighted n	*1212*	*1023*						
P for mode effect[c]	<0.01							

Analysis of unweighted data from combined national and Baltimore sample strata of the 1999–2000 NSBME. Analysis excludes 39 respondents assigned to T-ACASI who were interviewed in T-IAQ because they did not have a touchtone telephone.
*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.
[a]Respondents who reported never using a given drug (marijuana, cocaine, injected drug) were not asked questions about recency of their last use; they are coded as 'no' for use in all time periods in calculating prevalences. Percentage responding 'never used' is 100% minus the percent reporting use.
[b]Since only 1 respondent in the interviewer-administered condition reported injection drug use in preceding 30 days and none reported drug use in period 1 month to 1 year before survey, we have collapsed these two categories.
[c]*P* value for Mantel-Haenszel test for linear association.
[d]ORs contrast prevalence in T-ACASI condition to prevalence in T-IAQ condition. See Table 1 for description of control variables used in calculating the adjusted ORs.

ing of same-gender sexual contacts is attenuated in geographic locales that have a history of tolerance for gay men and lesbians.

We explored the homogeneity of estimates of the impact of T-ACASI on reporting of marijuana or cocaine use in the past 30 days and injection drug use in the past 5 years. (Exploratory analyses were also carried out with reporting of use during other time periods.) We fit log-linear models to test the hypothesis that the impact of T-ACASI was equivalent across subpopulations defined by: gender, marital status, age, educational level, race, sample strata (National vs. Baltimore), region of the country, presence of children in the household, and level or urbanization. This analysis revealed few consistent patterns of subpopulation differences in the impact of T-ACASI. The one exception occurred for residence outside of a major metropolitan area. These results are shown in Table 4.

As Table 4 shows, the impact of T-ACASI was strongest for persons residing outside the 21 largest metropolitan areas (MSAs) in the USA. Thus T-ACASI had a significantly greater impact on reporting of marijuana use in the past 30 days ($OR_s = 2.7$ vs. 1.2, $p = 0.035$) and a greater impact on reporting of cocaine use in the past 30 days outside of the 21 largest MSAs—although the latter result is not statistically reliable ($OR_s = 7.4$ vs. 1.1, $p = 0.151$). Reporting injection drug use in the past five years showed a similar pattern, but the number of injection drug users is sparse and the result is not statistically reliable ($OR_s = 7.6$ vs. 3.6, $p > 0.5$). These results could be interpreted as evidence that reporting of drug use is less sensitive for respondents residing in major metropolitan areas.

It should be noted that in all three panels of Table 4, the percent of respondents reporting drug use is higher in

the 21 largest metropolitan areas when human telephone interviewers collect the data, e.g. 7.4% vs. 4.7% for T-IAQ reports of marijuana use in the past 30 days. When T-ACASI is used to collect these data, this pattern reverses for reported marijuana and cocaine use. 12.0% of T-ACASI respondents residing outside of the 21 largest MSAs report marijuana use in the past 30 days vs. 8.5% of respondents residing within those large MSAs. Similarly, 2.9% of respondents outside these large MSAs reported cocaine use in the past 30 days vs. 0.8% of respondents residing within those large metropolitan areas.

Validity

While we do not have an independent criterion that can be used to assess the accuracy of respondents' reports of drug use, there are validity checks that can be performed. This is important since providing complete privacy in the T-ACASI interview mode leaves the measurement process vulnerable to respondents who may wish to speed through the interview by rapid and random pressing of answer keys.

Table 5 examines this threat to measurement validity by testing for one of the best-established patterns of reported drug use—most users of 'hard' drugs report trying 'softer' drugs as well. Table 5 shows the percent of: (1) injection drug users who report experience with cocaine and marijuana, and (2) cocaine users who report experience with marijuana. It will be seen from Table 5 that in both measurement conditions, few injection drug users report no experience with cocaine (6% T-ACASI; 14% T-IAQ) or with marijuana (0% T-ACASI; 5% T-IAQ). Similarly few self-reported cocaine users report having no experience with marijuana (3% T-ACASI; 6% T-IAQ). In every comparison, fewer 'hard' drug users report the unexpected pattern of no experience with the 'softer' drug when measurements were made using T-ACASI rather than a human interviewer.

National estimates

Table 6 presents population-weighted estimates of the prevalence of illicit drug use derived from the 2000 National Household Survey of Drug Abuse (NHSDA) and the 1999–2000 NSBME. NHSDA estimates are limited to

Table 4 Percentage of respondents reporting selected types of drug use by interview mode and urban residence in 1999–2000 NSBME (National Strata only).

Drug use estimate & urban residence	T-ACASI		T-IAQ		*CRUDE OR*	*Interaction test*[a] *P*
	%	*n*	*%*	*n*		
Marijuana use in past 30 days						
21 largest MSAs	8.5	236	7.4	296	1.2	0.035
Counties outside 21 largest MSA areas	12.0	410	4.7	510	2.7***	
Cocaine use in past 30 days						
21 largest MSAs	0.8	237	0.7	298	1.3	0.151
Counties outside 21 largest MSA areas	2.9	411	0.4	512	7.7**	
Injection Drug Use in past 5 years						
21 largest MSAs	2.5	238	0.7	299	3.8	ns
Counties outside 21 largest MSA areas	1.5	411	0.2	515	7.6	

Analysis of unweighted data from the national sample strata of the 1999–2000 NSBME. Analysis excludes 24 national strata respondents assigned to T-ACASI who were interviewed in T-IAQ because they did not have a touchtone telephone.
[a] *P* values derived from loglinear models testing for 3-way interaction of Response by Interview Mode by Level of Urbanization.
$^{**}P < 0.01$, $^{***}P < 0.001$.

Table 5 Validity analysis: Percent of Injection Drug and Cocaine Users who report use of 'softer' drugs by interview mode. (Tabulated from the 1999-2000 NSBME.)

	T-IAQ			*T-ACASI*		
Drug user	*n*	*Cocaine use*	*Marijuana use*	*n*	*Cocaine use*	*Marijuana use*
Injection Drug Users	22	86.4%	95.5%	33	94.0%	100.0%
Cocaine Users	215	–	93.9%	225	–	97.3%

Analysis of unweighted data from combined national and Baltimore sample strata of the 1999-2000 NSBME. Analysis excludes 39 respondents assigned to T-ACASI who were interviewed in T-IAQ because they did not have a touchtone telephone.

the population of US adults residing in non-group quarters with touchtone telephones. Because only age groupings (rather than exact ages) are available in the public dataset for the 2000 NHSDA, Table 6 compares NHSDA estimates for 18–49 years olds with NSBME estimates for 18–45 years olds.

There are several other ways in which the NHSDA and NSBME samples cannot be made precisely equivalent. Most obviously, the NSBME and NHSDA measurements were separated slightly in time (September 1999 to April 2000 for NSBME vs. January to November, 2000 for the NHSDA). In addition, the NHSDA was an in-person survey while NSBME was a telephone survey introducing potential mode effects. There are also a variety of other technical differences between the studies. So, for example, the NHSDA included a Spanish-language version while the NSBME did not capture the non-English speaking segment of the Hispanic population. Similarly, the in-person NHSDA achieved a 68.6% survey response rate (calculated as the product of the NHSDA household screening rate of 92.8% and interview response rate of 73.9%; Epstein 2002; Appendix B). This is higher than that achieved by the NSBME telephone survey (national stratum: 62% for T-IAQ and 53% for T-ACASI; Villarroel *et al.* 2002, in press).

Because the impact of these divergences cannot be calibrated, it is not appropriate to conduct statistical hypothesis testing. Nonetheless, a striking pattern emerges from this tabulation. The national prevalence estimates derived from the 2000 NHSDA are remarkably similar to the population-weighted estimates obtained when the NSBME telephone interviewers asked the questions (T-IAQ condition). In contrast, national prevalence estimates derived from the NSBME's T-ACASI measurements are all considerably higher than both the estimates derived from the 2000 NHSDA data and the estimates derived from the NSBME T-IAQ measurements. So, for example, the estimated prevalence of marijuana use in the past 30 days is estimated to be 6.4% in both the NHSDA and the NSBME T-IAQ condition, while the NSBME T-ACASI measurements provide a population-weighted prevalence estimate of 12.6%.

Table 6 Population-weighted estimates of prevalence of illicit drug use among US population residing in telephone-accessible households derived from (1) sample ages 18 to 49 in 2000 National Household Survey of Drug Abuse, and (2) sample ages 18 to 45 in 1999–2000 NSBME T-IAQ and T-ACASI conditions. (All estimates are weighted)

Prevalence	*2000 NHSDA*	*NSBME T-IAQ*[b]	*NSBME T-ACASI*
Any marijuana use			
In last 30 days	6.4%	6.4%	12.6%
In last year	11.2%	11.7%	21.2%
in lifetime	48.2%	49.1%	57.2%
Any cocaine use			
In last 30 days	0.7%	0.8%	1.7%
In last year	2.0%	2.9%	4.4%
in lifetime	17.2%	17.8%	22.2%
Any drug injection			
In last year[a]	0.1%	0.0%	1.5%
In lifetime[a]	1.5%	1.1%	3.5%

NOTE. NHSDA sample is restricted to population who indicated they had a touchtone telephone and did not live in group quarters. See text for description of variation in design of NHSDA and NSBME. NSBME uses national sample strata only. Tabulations from NHSDA and NSBME do not impute values for missing data.

[a]Note that all NSBME estimates are for reported injection of 'heroin, speed, cocaine, or steroids'. Respondents in the NHSDA were first asked if they 'ever injected drugs not prescribed to you?' Responses to this question yielded the injection prevalence estimate of 1.7%. NHSDA questions on the timing of the last injection were only asked for heroin, cocaine, and stimulants. The lifetime injection prevalence estimate for these three types of drugs is 1.5% lifetime and 0.1% in last year.

[b]T-IAQ condition of NSBME includes approximately 3% of sample who did not have touchtone telephones.

DISCUSSION

The results of this national experiment provide strong evidence that T-ACASI increases the reporting of *illicit* drug use – at least for the population included in this survey (US adults ages 18–45). Evidence of the impact of this technology on reporting of *licit* drug use is mixed which is consistent with expectations from past research (e.g. Turner *et al.* 1992). Reporting of alcohol use, per se, appears largely unaffected by eliminating the requirement that respondents reveal their drinking behaviors to a human telephone interviewer. Indeed, our results *suggest* that respondents may be more likely to report episodes of heavier drinking (5 or more drinks on one occasion) when questioned by a human telephone interviewer. Reporting of problems with alcohol use was, however, more frequent when respondents responded to a computer rather than a human interviewer for one of four problems: feeling guilty about drinking.

While the results for alcohol use and alcohol problems are somewhat mixed, those for reporting of illicit drug use are not. Respondents are consistently and considerably more likely to report illicit drug use when they are responding to a T-ACASI computer rather than to a human interviewer. So, for example, the reporting of cocaine use in the past month increased three-fold from 0.7% to 2.1%, and the reporting of injection drug use in the past five years increased five-fold from 0.3% to 1.6%. Indeed, we found that only 1 of 1212 respondents reported injection drug use in the past year to a human interviewer while 13 of 1023 respondents reported it in the T-ACASI condition.

We cannot independently corroborate that this increased reporting of illicit drug use is a *more accurate* reflection of its true prevalence. It is, however, generally accepted by survey methodologists that biases in the reporting of illicit or stigmatized behaviors produces a *net* negative bias in estimates in the prevalence of these behaviors in the population (Sudman & Bradburn 1974; Fay *et al*. 1989; Turner *et al*. 1998b). This occurs because the number of survey respondents who deny engaging in stigmatized behaviors that they have, in fact, engaged in, is assumed to be larger than the number who falsely report behaviors that they have *not* engaged in. The increased rates of reporting of illicit drug use under more private survey conditions could thus be interpreted to reflect a reduction in the reporting bias and thereby an increase in the accuracy of the measurements.

In the present instance, some confidence in this assumption is gained by examining the patterns of response to the various drug use questions in the T-ACASI condition. The T-ACASI measurements show the expected structuring of drug use reports with the overwhelming majority of respondents who report a history of injection drug use also reporting marijuana use (100%) and cocaine use (94%). Similarly, 97% of T-ACASI respondents who report cocaine use also report some history of marijuana use. The structuring of drug use reported by T-ACASI respondents is fully comparable with that obtained when interviewers asked the question. At a minimum, we believe this result allows us to rule out random key pressing as an explanation for the increased reporting of drug use obtained in the T-ACASI condition.

When population-weighted data are used to generate national estimates from the NSBME experiment, we obtain both expected and unexpected results. As we previously saw with unweighted data, the T-ACASI condition of the NSBME national sample produces substantially higher estimates of the prevalence of illicit drug use in the USA than the interviewer-administered condition. This is an important finding, but it is one we fully expected to obtain based on our preliminary studies (Turner *et al*. 1996, 1998a; Gribble *et al*. 2000) and other recently published work (Corkey & Parkinson 2002). Initially, we were surprised that the telephone interviewer-administered (T-IAQ) condition of the 1999–2000 NSBME produced estimates of the prevalence of illicit drug use that were virtually identical to those derived from the in-person NHSDA of 2000. This would seem to be inconsistent with expectations based on Gfroerer and Hughes work in 1988 and Aquilino's *et al*.'s findings from 1986 and 1991. On reflection, however, we believe that this would be an over-interpretation of our findings. There are numerous ways in which the NHSDA and NSBME samples cannot be made equivalent, and thus the surprising concordance of the NHSDA and NSBME IAQ estimates should not be interpreted as evidence that the two measurement protocols produce equivalent results. Such a conclusion would require experimental evidence from studies in which members of a population are randomly assigned to one of these two protocols (see below).

There is little doubt, however, that the T-IAQ and T-ACASI measurements made in the NSBME fall within the range that would be expected based on the NHDSA findings, and that T-ACASI does reduce bias in reporting of illicit drug use. The latter finding may have important implications for future research on the prevalence of illicit drug use.

Implications

If the major source of bias in well-executed telephone surveys of drug use is the requirement that respondents reveal their drug use to a human interviewer, then the NSBME results suggest T-ACASI may substantially reduce this source of bias. Given that telephone data collection is considerably less costly and also (largely) avoids loss of statistical precision due to the sample clustering required for in-person surveys, T-ACASI might be used, at a minimum, to augment data collection for large-scale surveys like the National Survey of Drug Use and Health (successor to the NHSDA). This augmentation could provide more precise and/or less expensive estimates for subpopulations or topics that were too costly to contemplate with a NSDUH design that relies exclusively on in-person surveys.

A multimode NSDUH is not, however, something to be considered lightly. A focused program of methodological research would seem appropriate to consider whether, and if so, how, such a design might be constructed and tested. Some components of such an effort might include:

- Identifying topic areas and/or subpopulation estimates that might benefit from increased sample sizes (e.g. improved local area estimates, improved precision for *key* measures of change over time within subpopulations, etc.).
- Establishing a testing regimen to evaluate the feasibility, costs, and impact of using a T-ACASI mode to augment NSDUH data collection. This testing regimen should begin with randomized experiments that compare population prevalence estimates derived using T-ACASI and the in-person Audio-CASI currently employed for the NSDUH.
- Developing statistical models for generating synthetic estimates that combine data from multiple data collection modes (i.e. in-person ACASI and T-ACASI) that have different error structures due to differing response rates, measurement artifacts, etc.

- Considering the potential impact on the viability of a multimode strategy of changes in the telecommunications landscape including the growth in the number of households without landline telephone service, 'do not call' registries, etc.

Acknowledgements

Support for this research was provided by NIH grants R01-MH56318 and R01-HD31067 to Charles Turner. The authors wish to thank James Chromy, Michael Vorburger and Harpur Gordek of RTI's Statistical Research Division for constructing sample weights for the NSBME (JC, HG) and checking our NSDUH tabulations (MV), and Joseph Catania of UCSF for collaboration in early phases of the design of this study and for access to the public use dataset from his 1996 National Survey of Sexual Health. The authors also wish to thank the many other people who made major contributions during the proposal, design, and fielding of this research.

Conflict of interest statement

Although we are unsure whether it constitutes a conflict of interest, we note that the authors are employed by the Research Triangle Institute (RTI) and the Center for Survey Research at the University of Massachusetts (UMass). Both are non-profit institutions that conduct scientific research in a variety fields including population survey research. Scientists at RTI have played a major role in the development and testing of Audio-CASI and T-ACASI automated interviewing technologies. Both RTI and UMass conduct surveys for their own scientists, and, under contract, for scientists at other institutions, and for government agencies, and commercial organizations.

References

Aquilino, W. S. (1992) Telephone versus face-to-face interviewing for household drug use surveys. *International Journal of the Addictions*, **27**, 71–91.

Aquilino, W. S. (1994) Interview mode effects in surveys of drug and alcohol use: a field experiment. *Public Opinion Quarterly*, **58**, 210–240.

Brick, J. M., Waksberg, J., Kulp, D. & Starer, A. (1995) Bias in list-assisted telephone samples. *Public Opinion Quarterly*, **59**, 218–235.

Chen, P., Emrich, S., Gordek, H., Penne, M. A., Singh, A. & Westlake, M. (2002) *Person-Level Sampling Weight Calibration for the 2000 NHSDA*. Washington, DC: SAMSHA. Available online at: http://www.oas.samhsa.gov/NHSDA/2kmrb/00SamplingWeightW.pdf [30 accessed July 30 2004].

Cooley, P. C., Miller, H. G., Gribble, J. N. & Turner, C. F. (2000) Automating telephone surveys: Using T-ACASI to obtain data on sensitive topics. *Computers and Human Behavior*, **16**, 1–11.

Corkey, R. & Parkinson, L. (2002) A comparison of four computer-based telephone interviewing methods: Getting answers to sensitive questions. *Behavior Research Methods: Instruments and Computers*, **34**, 354–363.

DHHS (US Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies). (2004) *National Household Survey on Drug Abuse*, 2000 [Computer file]. 3rd ICPSR version. Research Triangle Park, NC: Research Triangle Institute [producer], 2004. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor], 2004.

Epstein, J. F. (2002) *Substance Dependence, Abuse, and Treatment: Findings from the 2000 National Household Survey on Drug Abuse*. (*NHSDA* Series A-16. DHHS Publication no. SMA 02–3642). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

Ewing, J. A. (1984) Detecting alcoholism: The CAGE questionnaire. *JAMA*, **252**, 1905–1907.

Fay, R., Turner, C. F., Klassen, A. D. & Gagnon, J. (1989) Prevalence and patterns of same-gender sexual contact among men. *Science*, **243**, 338–348.

Genesys Sampling Systems. (2002) Genesys Methodologies. Available online at: http://www.genesys-sampling.com/refernce/genmeth.htm [accessed 24 June 2002].

Gfroerer, J. C. & Hughes, A. L. (1992) Collecting Data on Illicit Drug Use by Phone. In: Turner, C. F., Lessler, J. T. & Gfroerer, J. C., eds. *Survey Measurement of Drug Use: Methodological Issues*, pp. 277–295. (DHHS Publication Number 92–1929). Washington, DC: Government Printing Office.

Goodman, L. A. (1968) The analysis of cross-classified data: Independence, quasi–independence, and interactions in contingency tables. *Journal of the American Statistical Association*, **63**, 1091–1131.

Goodman, L. A. (1978) *Analyzing Qualitative/Categorical Data*. Cambridge, Mass: ABT.

Gribble, J. N., Miller, H. G., Catania, J. A., Pollack, L. & Turner, C. F. (2000) The impact of T-ACASI interviewing on reported drug use among men who have sex with men. *Substance Use and Misuse*, 869–890.

Hall, W. D., Carless, J. M., Homel, P. J., Flaherty, B. J. & Reilly, C. J. (1991) The characteristics of cocaine users among young adults in Sydney. *Medical Journal of Australia*, **155**, 11–14.

ICPSR (Inter-University Consortium for Political and Social Research). (2004) Description—Study no. 6949: National Household Survey on Drug Abuse, 1994. Published online at: http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/06949.xml.AccessedJuly1, 2004.

Kendler, K. S., Karkowski, L. M., Corey, L. A., Prescott, C. A. & Neale, M. C. (1999) Genetic and environmental risk factors in the aetiology of illicit drug initiation and subsequent misuse in women. *British Journal of Psychiatry*, **175**, 351–356.

Kerr. W. C., Greenfield, T. K., Bond, J., Ye, Y. & Rehm, J. (2004) Age, period and cohort influences on beer, wine and spirits consumption trends in the US National Alcohol Surveys. *Addiction*, **99**, 1111–1120.

Kulp, D. (1994) Dynamics of list-assisted random digit dialing (RDD) frame coverage. *Proceedings of the Survey Research Methods Section, American Statistical Association.* 11–18. Available online at http://www.Amstat.Org/Sections/Srms/Proceedings/Papers/1994–002.pdf [accessed 30 July 2004].

de Leeuw, E. D. & van der Zouwen, J. (1988) Data quality in telephone and face-to-face surveys: A comparative meta-analysis. In: Groves, R. M., Biemer, P. P., Lyberg, L. E. *et al.*, eds. *Telephone Survey Methodology*, pp. 283–299. New York: Wiley.

Midanik, L. T. & Greenfield, T. K. (2003) Defining 'current drinkers' in national surveys: results of the 2000 National Alcohol Survey. *Addiction*, **98**, 517–522.

Perneger, T. V., Klag, M. J. & Whelton, P. K. (2001) Recreational drug use: a neglected risk factor for end-stage renal disease. *American Journal of Kidney Disease*, **38**, 49–56.

SAMSHA (Substance Abuse and Mental Health Services Administration) (1996) *National Household Survey on Drug Abuse: Public Release Codebook 1994-B*. Rockville, MD: SAMSHA.

Saxe, L., Kadushin, C., Beveridge, A., Livert, D., Tighe, E., Rindskopf, D. *et al.* (2001) The visibility of illicit drugs: implications for community-based drug control strategies. *American Journal of Public Health*, **91**, 1987–1994.

SPSS Inc (1993) *SPSS Statistical Software*, Version 6.0.1 Chicago, IL: SPSS Inc.

StataCorp. (2000) *Intercooled STATA 6.0 for Windows 98/95/NT*. College Station, TX: StataCorp LP.

StataCorp. (2003) *STATA Release 8, Version SE for Windows*. College Station, TX: StataCorp LP.

Sudman, S. & Bradburn, N. (1974) *Response Effects in Surveys*. Chicago: Aldine.

Sykes, W. & Collins, M. (1988) Effects of mode of interview: Experiments in the UK. In: Groves, R., Biemer, P. P., Lyberg, L. E. *et al.*, eds. *Telephone Survey Methodology*, pp. 301–320. New York: Wiley.

Turner, C. F., Forsyth, B. H., O'Reilly, J., Cooley, P. C., Smith, T. K., Rogers, S. M. *et al.* (1998a) *Automated Self-Interviewing and the Survey Measurement of Sensitive Behaviors*. In: Couper, M., Baker, R., Bethlehem, J., Clark, C., Martin, J., Nicholls, W. & Reilly, J., eds. *Computer-Assisted Survey Information Collection.*, pp. 455–473. New York: Wiley and Sons.

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H. & Sonenstein, F. L. (1998b) Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science*, **280**, 867–873.

Turner, C. F., Lessler, J. T. & Devore, J. (1992) Effects of mode of administration and wording on reporting of drug use. In: Turner, C. F., Lessler, J. T. & Gfroerer, J. C., eds. *Survey Measurement of Drug Use: Methodological Issues*, pp. 177–219. DHHS Publication Number, pp. 92–1929. Washington, DC: Government Printing Office.

Turner, C. F., Miller, H. G., Smith, T. K., Cooley, P. C. & Rogers, S. M. (1996) Telephone audio computer-assisted self-interviewing (T.-ACASI) & survey measurements of sensitive behaviors: Preliminary results. In: Banks, R., Fairgrieve, J., Gerrard, L. *et al.*, eds. *Survey and Statistical Computing 1996*, pp. 121–130. Chesham., Bucks, UK: Association for Survey Computing.

Turner, C. F., Rogers, S. M., Miller, H. G., Miller, W. C., Gribble, J. N., Chromy, J. R. *et al.* (2002) Untreated gonococcal and chlamydial infection in a probability sample of adults. *JAMA*, **287**, 726–733.

Villarroel, M. A., Turner, C. F., Eggleston, E., Al-Tayyib, A. A. & Rogers, S. M. (2002) Same-gender sex in the USA: Impact of interview methodology and time on prevalence estimates. *Abstract #44782 from 130th Annual Meeting of The American Public Health Association.* Published online at: http://apha.confex.com/apha/130am/techprogram/paper_44782.htm [accessed 18 March 2005].

Villarroel, M. A., Turner, C. F., Eggleston, E., Al-Tayyib, A. A., Rogers, S. M., Roman, A. M. *et al.* (in press) Same-gender sexual contacts in the USA: Impact of New Telephone Survey Technology on Prevalence Estimates. Available online at: http://www.soc.qc.edu/staff/turner/techpdfs/60_sgs_exp.pdf [accessed 4 April 2005].

APPENDIX 1

Question Wordings

Now we have some questions about your use of alcohol and other drugs. As you answer these questions, count as a drink a can or bottle of beer; a wine cooler or a glass of wine, champagne, or sherry; a shot of liquor or a mixed drink or cocktail.

Have you EVER had a drink of an alcoholic beverage, such as wine, beer, or liquor?
(Response categories: Yes; No) (If answered: No, Don't Know, or Refused, skip to first question on illicit drug use.)

Now, please think specifically about the past 30 days. During the past 30 days, on how many days did you drink one or more drinks of alcoholic beverages?
(Response categories: none; 1–4 days; 5–9 days; 10–15 days; 16–25 days; every day or almost every day) (If answered: No, Don't Know, or Refused, skip to first question on illicit drug use.)

During the past 30 days, on how many days did you have five or more drinks on the same occasion? (By 'occasion', we mean at the same time or within a couple of hours of each other.)
(Response categories: I never had 5 or more drinks on the same occasion during the past 30 days; On 1–4 days in the past 30 days, etc.; On 5–9 days, etc.; On 10–15 days, etc.; On 16–25 days, etc.; I had 5 or more drinks on the same occasion every day or almost every day. For, etc. insert: 'I had 5 or more drinks on the same occasion'.)

Have you ever felt you should cut down on your drinking?
(Response categories: Yes; No)

Have people annoyed you by criticizing your drinking?
(Response categories: Yes; No)

Have you ever felt bad or guilty about drinking?
(Response categories: Yes; No)

Have you ever taken a drink first thing in the morning to steady your nerves or get rid of a hangover?
(Response categories: Yes; No)

Have you ever, even once, used marijuana or hashish?
(Response categories: Yes; No) (If answered: No, Don't Know, or Refused, skip to first question on cocaine use.)

How long has it been since you last used marijuana or hashish?
(Response categories: In the past 30 days; Between 1 month and 1 years ago; Between 1 years and 3 years ago; More than 3 years ago.)

The next two questions ask about cocaine, including all the different forms of cocaine such as powder, 'crack', free base and coca paste.

Have you ever, even once, used any form of cocaine?
(Response categories: Yes; No) (If answered: No, Don't Know, or Refused, skip to first question on injection drug use.)

How long has it been since you last used any form of cocaine?
(Response categories: In the past 30 days; Between 1 month and 1 years ago; Between 1 years and 3 years ago; More than 3 years ago.)

Have you EVER injected yourself with heroin, speed, cocaine, or steroids?
(Response categories: Yes; No) (If answered: No, Don't Know, or Refused, skip to next [non-drug] section of questionnaire.)

When was the last time you injected yourself with heroin, speed, cocaine, or steroids?
(Response categories: In the past 30 days; Between 1 month and 1 years ago; Between 1 years and 5 years ago; More than 5 years ago.)