TECHNICAL PAPERS ON
HEALTH AND BEHAVIOR MEASUREMENT

TECHNICAL PAPER 8

Effects of Mode of Administration and Wording on Reporting of Drug Use

Charles F. Turner, Judith T. Lessler, and James W. Devore

Reference Citation

Turner, C.F., J.T. Lessler, and J.W. Devore. (1992) Effects of mode of administration and wording on reporting of drug use. In C.F. Turner, J.T. Lessler, and J.C. Gfroerer (eds.), *Survey Measurement of Drug Use: Methodological Studies.* DHHS Pub. No.

Survey Measurement of Drug Use

Methodological Studies

CHARLES F. TURNER AND JUDITH T. LESSLER
Research Triangle Institute

AND

JOSEPH C. GFROERER
National Institute on Drug Abuse

EDITORS

NATIONAL INSTITUTE ON DRUG ABUSE
DIVISION OF EPIDEMIOLOGY AND PREVENTION RESEARCH
5600 FISHERS LANE
ROCKVILLE, MARYLAND 20857

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
ALCOHOL, DRUG ABUSE, AND MENTAL HEALTH ADMINISTRATION

For sale by the U.S. Government Printing Office
Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328
ISBN 0-16-038065-0

7

Effects of Mode of Administration and Wording on Reporting of Drug Use

Charles F. Turner, Judith T. Lessler, and James W. Devore

From preliminary investigations reported in the preceding chapters and discussions with research staff and sponsors, several alternative designs for the National Household Survey of Drug Abuse (NHSDA) were considered for experimental testing during 1990. Major variations that were considered included

- a modified version of the NHSDA that attempted to improve question wording and format to alleviate problems noted in our cognitive assessment of the NHSDA questionnaire (see Chapters 2 and 3) and our review of faulty and missing data problems in the 1988 wave of the NHSDA (see Chapters 4 and 5);
- interviewer administration (rather than self-completion) of survey forms;
- use of a monetary incentive to increase survey participation rates; and
- collection of hair samples to test the feasibility of making (some) prevalence estimates without relying on self-reports. (Such data would provide the opportunity to validate self-reports.

Practical considerations made it impossible to test the latter two strategies in 1990. Thus, a field test was conducted to evaluate the effects of

alternative questionnaire wording and mode of administration on the reporting of drug use in the NHSDA.

This chapter describes the design of the methodological field test conducted in 1990, and it reviews estimates of the prevalence of self-reported drug use obtained by the different questionnaire versions used in this field test. In Chapter 8, we review the impact of these factors on the completeness and consistency of the reporting of drug use. Chapter 9 assesses the effects of our decomposition of complex concepts, such as "nonmedical use of psychotherapeutic drugs," on the reporting of drug use.

EXPECTATIONS REGARDING PREVALENCE

In interpreting the results of analyses presented in this chapter, one might reasonably assume that, on average, the bias in respondents' reporting of illicit drug use is negative and nontrivial in magnitude for all versions of the questionnaire: that is, the number of persons concealing past use of illicit drugs is larger than the number who falsely admit having used these drugs. Available evidence in the methodological literature suggests that the extent of this bias is, in part, a function of the degree of anonymity afforded to a respondent in admitting the use of drugs, the extent of the respondent's trust in guarantees that survey results will not be made public, and the legality of the particular drug in question (Miller et al., 1990: Ch. 6; Bradburn and Sudman, 1979; Turner and Martin, 1984).

Past research led us to expect that we would find (1) higher levels of reporting of drug use in the self-administered versus the interviewer-administered questionnaires; (2) that these differences would be greatest for illicit drugs; and perhaps (3) that these differences would be more pronounced for recent drug use than for drug use of many years ago.

We had no similar set of expectations about the effect of revised wording on the reporting of drug use. Indeed, we anticipated that, rather than affecting prevalence rates, per se, improvements in the wording and design of the questionnaire would increase the consistency and completeness of self-reports but not necessarily the frequency with which drug use was reported.

DESIGN OF THE FIELD TEST

The Sample

The 1990 NHSDA Field Test employed a multistage probability sample[1] of the household population age 12 and older, drawn from 33 purposely

[1] Janella Pantula was responsible for the design and execution of the sampling plan and for embedding the experiment within that sample design.

chosen metropolitan areas. These primary sampling units, or PSUs, were chosen from the 1988–1990 NHSDA to reflect the diversity of the survey's interview situations. Household population was defined as residents of housing units, according to the standard census definition of housing unit (U.S. Bureau of the Census, 1982). Residents of group quarters or institutions (e.g., rooming houses, dormitories, military installations) were not included.

The 1990 NHSDA Field Test sample was designed to yield approximately 4,000 interviews. The basic plan involved several stages: the selection of subareas (segments) within the chosen PSUs, the selection of sample households within subareas, and the selection of eligible individuals within sample households. The sample was designed to yield equal numbers of respondents from four age groups: 12–17, 18–25, 26–34, and 35 or older.

A randomized factorial experiment was embedded in the sample design for the study PSUs. Two alternative questionnaires (the 1990 NHSDA and a new questionnaire) were crossed with two modes of administration (self and interviewer). Each interviewer was trained in all four resultant treatments, which were randomly assigned to successive groups of four households. This design provided for two replicates of the factorial experiment in each segment.

Questionnaire Versions

Four versions of the questionnaire were included in the field test. Two versions were entirely interviewer-administered, and two used a combination of self-administered and interviewer-administered components (as is done in the current survey). One of the questionnaires was exactly the same as that used for the 1990 NHSDA. We will refer to that version of the questionnaire as Version A or by the shorthand title: *current wordings in self-administered format.* A copy of that questionnaire is reproduced in Appendix E. An interviewer-administered version of this questionnaire was also developed. This version used, as much as possible, question wordings that were identical to those used in the 1990 survey. This version will be called Version B or *current wordings in interviewer-administered format.* (Readers may also consult Appendix E for the wordings used in Version B.) The remaining two versions (C and D) tested different wordings (and different structures) for the questionnaire. Version C used these *new wordings in a self-administered format,* and Version D used the same *new wordings but in an interviewer-administered format.* A copy of Version C is included in Appendix F; it may be referred to for the question wordings used in both Versions C

and D. Below we describe in more detail some of the major features of each version of the questionnaire.

Version A
Current Wordings in Self-Administered Format

Version A is identical to the 1990 NHSDA questionnaire,[2] which mainly used self-completed forms. (Questions about cigarette use are interviewer-administered.) Respondents fill out answer sheets for alcohol, sedatives, tranquilizers, stimulants, analgesics, marijuana and hashish, inhalants, cocaine, hallucinogens, heroin, drug dependencies during the past year, special topics (e.g., use of ice and needles), drinking experiences, drug problems, and perceived risks of using drugs. While almost all of the survey is self-administered, it should be noted that the alcohol, marijuana, and cocaine questions are read aloud by the interviewers as the respondents read along and then privately mark their answer sheets. The respondent does have the option of asking the interviewer to read aloud the questions in the other sections of the survey. In addition, some respondents request that the interviewer mark the answers; however, interviewers are trained to do this only as a last resort.

Version B
Current Wordings in Interviewer-Administered Format

Version B was entirely interviewer-administered, and it made use of wordings that were virtually identical to those in the 1990 questionnaire. In altering this questionnaire for interviewer administration, it was, however, necessary to make extensive use of skip instructions. Thus, if a respondent indicated that he or she never used a particular substance, then the remaining questions on the use of that substance were skipped. (Doing otherwise would have been awkward for the interviewer who would have to repeat numerous inapplicable questions.)

Version C
New Wordings in Self-Administered Format

Version C was largely self-administered and it used a variety of new wordings in an attempt to overcome problems identified in the cognitive assessments (see Chapters 2 and 3), and in our analyses of missing and faulty data (Chapters 4 and 5). Among the innovations were the following:

[2]Chapter 2 presents a detailed cognitive assessment of the 1988 NHSDA questionnaire. That questionnaire was identical in most respects to the 1990 questionnaire.

- Before beginning the interview, the interviewer attempted to have the respondent's anchor the reference periods that would be used to recall their drug use. To do this, interviewers first wrote the boundary dates of the 30-day and 12-month reference periods on a calendar. They then asked the respondents to recall a personally experienced event that occurred at the boundary of each reference period. Those events were used as the anchors for the reference periods.
- Extensive use of branching (i.e., skip) instructions was incorporated in the new questionnaire as well as tests of the respondents' ability to implement those instructions.[3] In the new questionnaire, respondents first completed the cigarette section and then the interviewer reviewed their answers to determine if they had understood the skip instructions. Respondents then completed sections on use of alcohol, pain killers, tranquilizers, stimulants, sedatives, and inhalants; in these sections respondents were instructed to skip most questions which were not appropriate (given their prior answers).[4] The use of skip instructions in these sections allowed us to remove most of the "implicit questions" identified in our cognitive assessment of the questionnaire (see Chapter 2).
- Question wordings were revised to remove inconsistencies and ambiguities. For example, in Version A some of the answer categories were ordered from high to low and others from low to high. In Version C, they were all ordered from low to high. Similarly, question stems for the "recency questions" were changed to match their answer categories. Versions A and B asked "When was the most recent time that you. . . ", and the answer categories were phrased in terms of duration since last use ("More than 1 month ago but less than six months ago"). In Version C the question stem was changed to read, "How long has it been since you last. . . ".

[3]The marijuana and cocaine sections were similar in structure to Version A in that respondents were, in general, required to answer every question. Branching was used in the sections on hallucinogens, heroin, drinking experiences, and drug problems.

[4]There were some exceptions to this rule in order to test the reliability of answers. We did this by having respondents answer a question on 12-month use even though they had previously indicated that it had been more than a year since they had used the particular substance.

- Questions on nonmedical use of psychotherapeutic drugs were decomposed into three components of nonmedical use, and separate questions were asked about each.[5]
- Questions on drug dependencies during the past 12 months were also decomposed. In Version A, respondents were asked to review a list of substances and to mark those that they had tried to cut down on during the past year; needed larger amounts of in order to get high; and so on. In the new version of the questionnaire, these questions were asked separately for each substance if the respondent indicated any use of the substance during the past 12 months.
- The complex question on problems related to drug use was also decomposed. In Version A, respondents were instructed to perform a three-part task[6] to answer each of 11 questions about personal problems[7] that might have been induced by drug use. Respondents in Version A had to (1) indicate whether they had the problem *in the past 12 months*, (2) decide whether the problem was caused by their drug use, and (3) indicate which drug "probably caused the problem." In Version C, this complex task was decomposed. All respondents were asked whether they had experienced the particular problem during the past 12 months. Those who indicated that they had experienced the problem were then asked whether they thought the problem was caused by their use of drugs. If so, respondents were then asked to identify the substance that caused the problem.

In addition to these major changes, many smaller changes were made in Version C. For example, we changed the order of some questions, combined the answer sheets in a single booklet, and so forth.

Version D
New Wordings in Interviewer-Administered Format

Version D was entirely interviewer-administered. It included all of the features of Version C. (There was, of course, no testing of respondent's

[5] Respondents were also asked a global question on nonmedical use in the past 12 months to allow us to assess response consistency across the different questioning strategies.

[6] At the outset, respondents also had to determine if they had ever used *any* substance covered in this section of the questionnaire.

[7] Problems included becoming depressed, having arguments or fights with family or friends, finding it difficult to think clearly, needing to get emergency medical help, etc.

ability to implement skip instructions since skip instructions were implemented by the interviewer.)[8]

Field Operations

Household Contact

Household interviewing was conducted from October through December of 1990. Initial contact with potential respondents was made by a letter from the project director at Research Triangle Institute (RTI). Interviewers visited the sample households with their RTI identification badges clearly visible and introduced themselves as soon as an adult member of the household had been identified. This introduction referred to the prior letter and gave the household member a "confidentiality card," which explained the purpose of the data collection effort, assured the respondent of the confidentiality of all gathered information, and identified the time required to complete the questionnaire.

When a teenager (age 12 to 17) was involved, the interviewer was responsible for obtaining verbal consent from a parent or guardian first. A paragraph directed at parental concerns was included in the introduction to the questionnaire, and interviewers were prepared to answer any questions. Once consent had been obtained, the parent was asked to leave the interview setting to ensure the confidentiality of the youth's responses.

Callbacks

At least five callbacks were made to each household to obtain screening information. These calls were made at different hours on different days of the week to increase their effectiveness. When no household member could be contacted after repeated visits, the interviewer was permitted to take general information from a knowledgeable neighbor to determine if anyone in the household might be eligible or whether additional visits were justified.

Household Screening

Screening was performed in each household. The eligibility criterion for the Field Test (which was the same as the criterion for the 1990 NHSDA) was based on the age and race/ethnicity of household members. Each

[8] Version D did not make *maximum* use of branching. Some repetitious gathering of information was done to assess the reliability of responses.

household type was assigned a predetermined selection probability, and, by referencing a grid appropriate to the household type, the interviewer could decide which member of the household, if any, was eligible for the interview. On average, one out of every three households had a member eligible for the study.[9] The interview was conducted by using the questionnaire treatment randomly assigned to it in the Assignment Control Form (provided by the central office).

Fieldwork Results

Implementation of these data collection procedures[10] resulted in the completion of 10,726 screenings from the 11,257 housing units selected (95.3 percent). These screenings identified 4,358 eligible housing unit members of whom 3,326 completed interviews (76.4 percent).[11] In constructing a data file for subsequent analysis, the data were weighted and imputations were made when key demographic or socioeconomic variables were missing.[12]

Outcome Variables

For each drug in the NHSDA, two (or more) alternative questions might typically be used as indicators of use of that drug during the specific period in question. For example, to determine marijuana or hashish use,

[9] Unlike the 1990 NHSDA, in the Field Test it was impossible for more than one member of the household to be selected for an interview.

[10] Interviewers were required to make at least five callbacks to try to complete screening and interviewing at a household. In reality, unlimited callbacks were made as long as, in the opinion of the interviewer's supervisor, there was a chance that the screening or interview could be completed. In some cases, more than 12 callbacks were made to complete screening/interviewing.

[11] This response rate for the Field Test is lower than that typically obtained for the NHSDA. In 1990 the NHSDA had a response rate of 82 percent (NIDA, 1991). The lower response rate for the Field Test reflects the fact that sampled PSUs were all metropolitan areas in which it is typically more difficult to obtain respondents' cooperation.

[12] Briefly, sampling weights were calculated as the inverse of the probabilities of selection at each stage of the process, adjusted for nonresponse. Since the household roster formed the basis for subsequent selection activities, failure to complete it constituted the first type of nonresponse encountered in the study. A weighting class adjustment procedure was used in which classes were based on whether the segment overlapped a central city of the MSA in which it was located. A weighting class adjustment factor was also calculated to account for person-level nonresponse within responding households. Race/ethnicity category, age group, sex, central city status, and census region were used to define the weighting classes. We examined these adjustment factors to determine if any were large enough to produce undue loss of precision because of their unequal weighting and found that further adjustments were unnecessary. Variance estimation strata were formed by grouping segments together according to their order of selection (i.e., their implicit stratification). Each group of segments defined a pseudo-stratum with two or three replicates. Finally, the small amounts of missing data for key demographic and socioeconomic variables were logically imputed for completed interviews.

Version C of the Field Test questionnaire might employ the recency-of-use question ("How long has it been since you last used marijuana or hashish?") or the direct question about marijuana use in the past 12 months ("Have you used marijuana at any time in the past 12 months?"). Because the resultant estimates might vary and because questions used on different versions of the questionnaire were not identical, we performed extended analyses, using both the most direct questions on each questionnaire version and a composite measure derived from several indicators available on the forms.

For each drug, we have compared estimates of self-reported drug use derived from the direct questions and the composite indicators (both for the total population and for demographic subgroups). Overall, we found little difference between the two sets of estimates; consequently, our subsequent analyses employed prevalence estimates derived from responses to the most direct questions on drug use available in each version of the questionnaire. It should be noted that employing the most direct questions to produce estimates of drug use differs from usual NHSDA estimation procedures. Our estimates thus are not directly comparable to published NHSDA estimates.

OVERVIEW OF RESULTS

Prevalence Rates by Experimental Treatment

Table 7-1 presents the prevalence of self-reported use of two illicit drugs, marijuana and cocaine, according to the mode of administration of the questionnaire and the wording of the questionnaire. Figure 7-1 plots the ratios for reported use of these drugs in self- versus interviewer-administration. The comparisons indicate that in all six measurements of the prevalence of marijuana and cocaine use, the self-administered questionnaire yielded higher estimated prevalence rates. As predicted, examination of the ratios indicates that the advantage of the self-administered questionnaire increases with the presumed sensitivity of the drug in question. The advantage of self-administered versions is also greatest for drug use in the past 30 days. On the self-administered form, respondents were 2.4 times more likely to report cocaine use, and 1.6 times more likely to report marijuana use, during the previous 30 days. The advantage of the self-administered form is somewhat reduced for use in the past year (ratios of 1.5 for cocaine and 1.3 for marijuana). For use at any time in the respondent's life, the advantage of the self-administered form is vanishingly small (ratios of 1.04 and 1.05, respectively).

TABLE 7-1 Estimates of Prevalence of Reported Drug Use (Percent) By Questionnaire Wording and Mode of Administration

		BY WORDING AND MODE							
		Current Wording		New Wording		BY WORDING		BY MODE	
DRUG USE		Self Admin.	Int. Admin	Self Admin.	Int. Admin	Current	New	Self Admin.	Int. Admin
ALCOHOL									
Lifetime	Est.	87.93	87.05	84.17	86.34	87.50	85.31	86.12	86.69
	s.e.	1.64	1.31	1.74	1.77	1.19	1.30	1.18	1.21
Past year	Est.	74.29	68.81	71.32	71.51	71.61	71.42	72.87	70.20
	s.e.	2.03	2.22	2.31	2.05	1.58	1.67	1.50	1.66
Past month	Est.	56.54	49.82	52.85	53.77	53.26	53.33	54.78	51.86
	s.e.	2.38	2.40	2.50	2.29	1.74	1.78	1.69	1.76
MARIJUANA									
Lifetime	Est.	37.66	33.17	35.68	36.64	35.47	36.18	36.71	34.96
	s.e.	2.46	2.46	2.12	2.48	1.90	1.66	1.65	1.75
Past year	Est.	8.71	6.68	8.57	6.59	7.72	7.53	8.64	6.63
	s.e.	1.15	1.02	1.25	1.08	0.79	0.85	0.93	0.76
Past month	Est.	4.83	3.00	5.22	3.21	3.94	4.16	5.02	3.11
	s.e.	0.87	0.63	0.95	0.81	0.52	0.62	0.68	0.51
COCAINE									
Lifetime	Est.	13.63	10.80	12.69	14.12	12.24	13.44	13.18	12.51
	s.e.	1.48	1.30	1.42	1.72	0.97	1.05	1.07	1.09
Past year	Est.	3.88	1.73	2.13	2.11	2.82	2.12	3.04	1.93
	s.e.	0.83	0.41	0.52	0.57	0.47	0.36	0.55	0.31
Past month	Est.	1.83	0.59	0.59	0.41	1.22	0.50	1.23	0.50
	s.e.	0.67	0.27	0.23	0.21	0.36	0.16	0.38	0.13

NOTES. Standard errors for estimates are shown on rows labelled s.e.

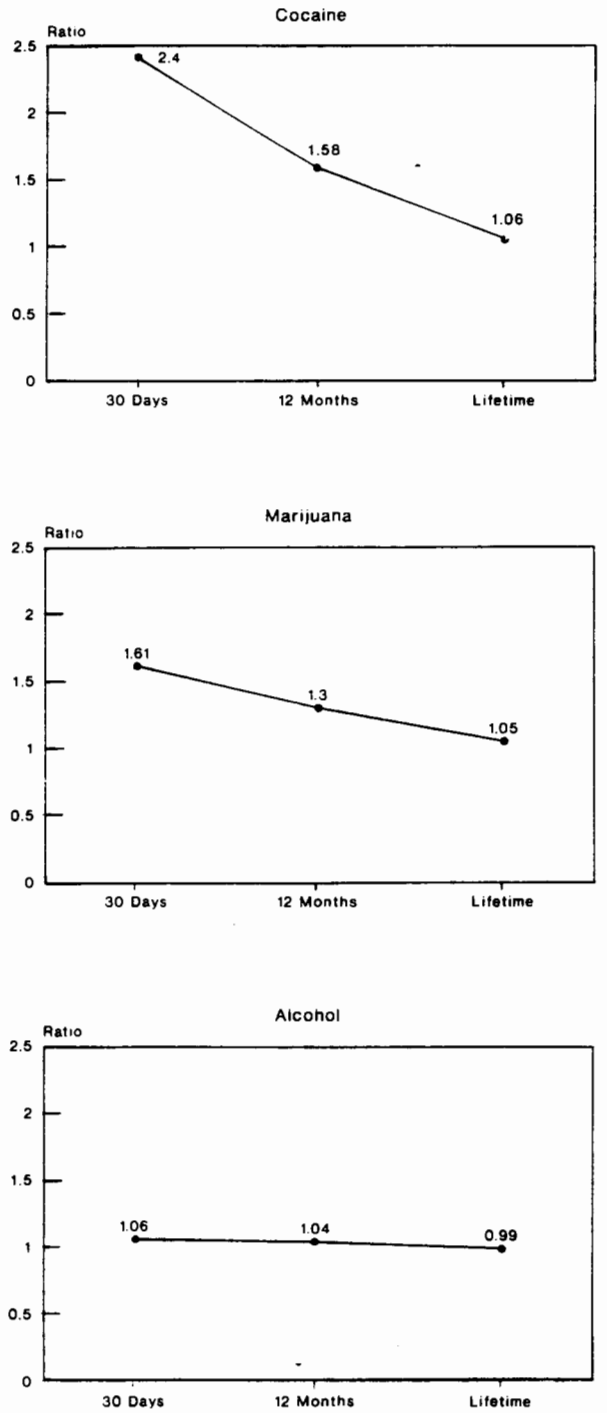


FIGURE 7-1 Ratio of estimates of prevalence of self-reported drug use: Estimate from self-administered versions divided by estimate from interviewer-administered versions.

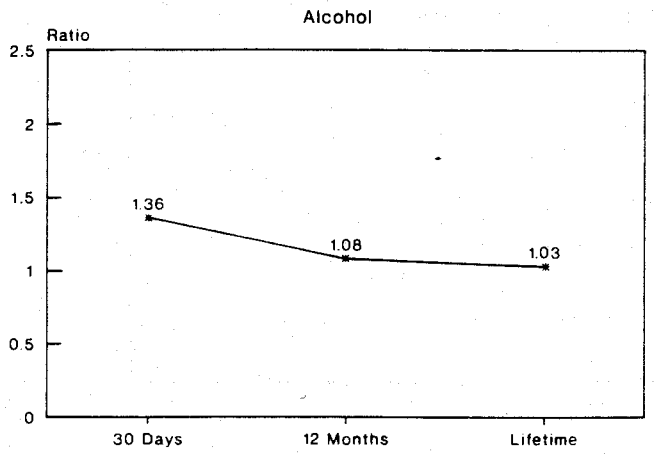
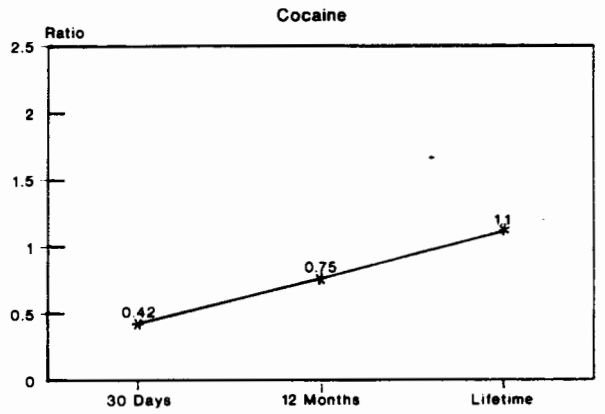


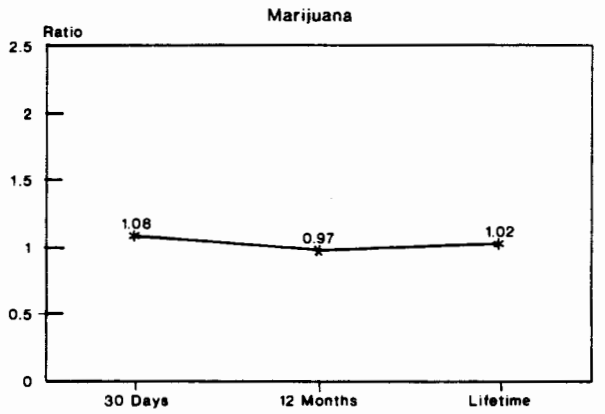
FIGURE 7-2 Ratio of estimates of prevalence of self-reported alcohol use by 12- to 17-year-olds: Estimate from self-administered versions divided by estimate from interviewer-administered versions.

Figure 7-2 shows a parallel analysis for reporting of the one licit drug, alcohol.[13] This plot indicates only trivial differences between prevalence rates obtained from the self-administered and interviewer-administered questionnaires (ratios ranged from 0.99 to 1.04). If, however, the analysis is restricted to the subgroup of respondents age 12 to 17 for whom alcohol is legally prohibited, the self-administered form again yields more frequent reports of use. Figure 7-2 presents this analysis: the ratios obtained range from 1.36 for alcohol use in the past 30 days to 1.03 for ever having consumed alcohol.

Figure 7-3 compares prevalence estimates obtained by wording the questions differently (these analyses combine responses obtained in self- and interviewer-administered versions of the same form). Given that question wording was expected to have an effect depending on whether the words were read or spoken, it is not surprising that this aggregation of responses yields no general finding. Note, however, the two instances in which the reported prevalence of drug use varied markedly between the current and new wordings of the questionnaire; both involve the Cocaine Form. The new question wording was only 0.41 times as likely as the old to obtain reports of cocaine use in the past 30 days, and 0.75 times as likely to obtain reports of use in the past year.

[13]Results for the Cigarette Form are not presented because it was interviewer administered in three versions of the questionnaire (A, B, and D). Reported rates of cigarette use for these versions were 25.7 (A), 21.8 (B), and 24.9 (D) percent. In the one self-administered version (C), the rate was 26.2 percent.





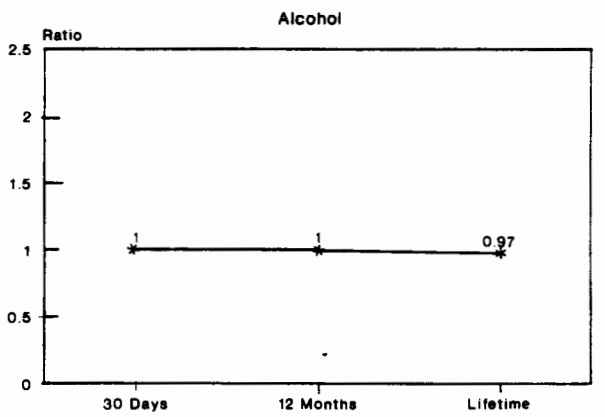


FIGURE 7-3 Ratio of estimates of prevalence of self-reported drug use: Estimate from versions using new wordings divided by estimate from versions using current wordings.

Variation in Prevalence Rates by Other Factors

The experimental variables (mode of administration, and question wording) are, by design, the major foci of interest in the Field Test. However, prevalence rates for drug use also vary by a number of other factors, including the social and demographic characteristics of the respondents. The methodological nature of this investigation led us to examine the variation in prevalence rates in terms of two characteristics of the interview situation: (1) whether the interview was conducted in private and (2) whether the interviewer had previous experience in administering NIDA-RTI drug surveys.

Figures 7-4 and 7-5 plot the variation in rates of reported use of cocaine and marijuana during the past year by such social and demographic factors as age, gender, race, education, region, and residence in the central city of a Metropolitan Statistical Area (MSA). These figures show large variations in drug use for some of the demographic subgroups: thus, drug use declines markedly with age, is higher in central cities, and is roughly twice as high among males as among females. Although none of these factors, in themselves, are surprising, questions arise about their possible interaction with the experimental treatment to produce treatment rate variations whose patterns do not represent a simple addition of the overall effects of treatment variables (wording and mode of administration) and demographic factors (e.g., age). Such effects might be expected if, for example, the responses of certain age groups in the population were more sensitive to the anonymity afforded by the self-administered questionnaire. These questions are considered in our multivariate statistical analyses.

Figure 7-6 plots the variation in prevalence rates for cocaine and marijuana use by privacy of the interview situation (i.e., presence of other persons) and the interviewer's previous experience in administering NIDA-RTI surveys. The results in Figure 7-6 are surprising; although we anticipated a negative bias in drug use reporting, because of the presence of other persons, the magnitude and direction of the observed association were unexpected.

Even more surprising was the apparent effect of interviewer experience on the reporting of past drug use. Over all of the experimental conditions, interviewers without previous experience in conducting NIDA-RTI surveys obtained a prevalence rate for use of marijuana in the past year that was 1.5 times higher than that of experienced interviewers (9.0 versus 6.3 percent). For cocaine use in the past 12 months, the rates obtained by inexperienced interviewers were almost twice as high as those obtained by interviewers with prior experience (3.3 versus 1.7

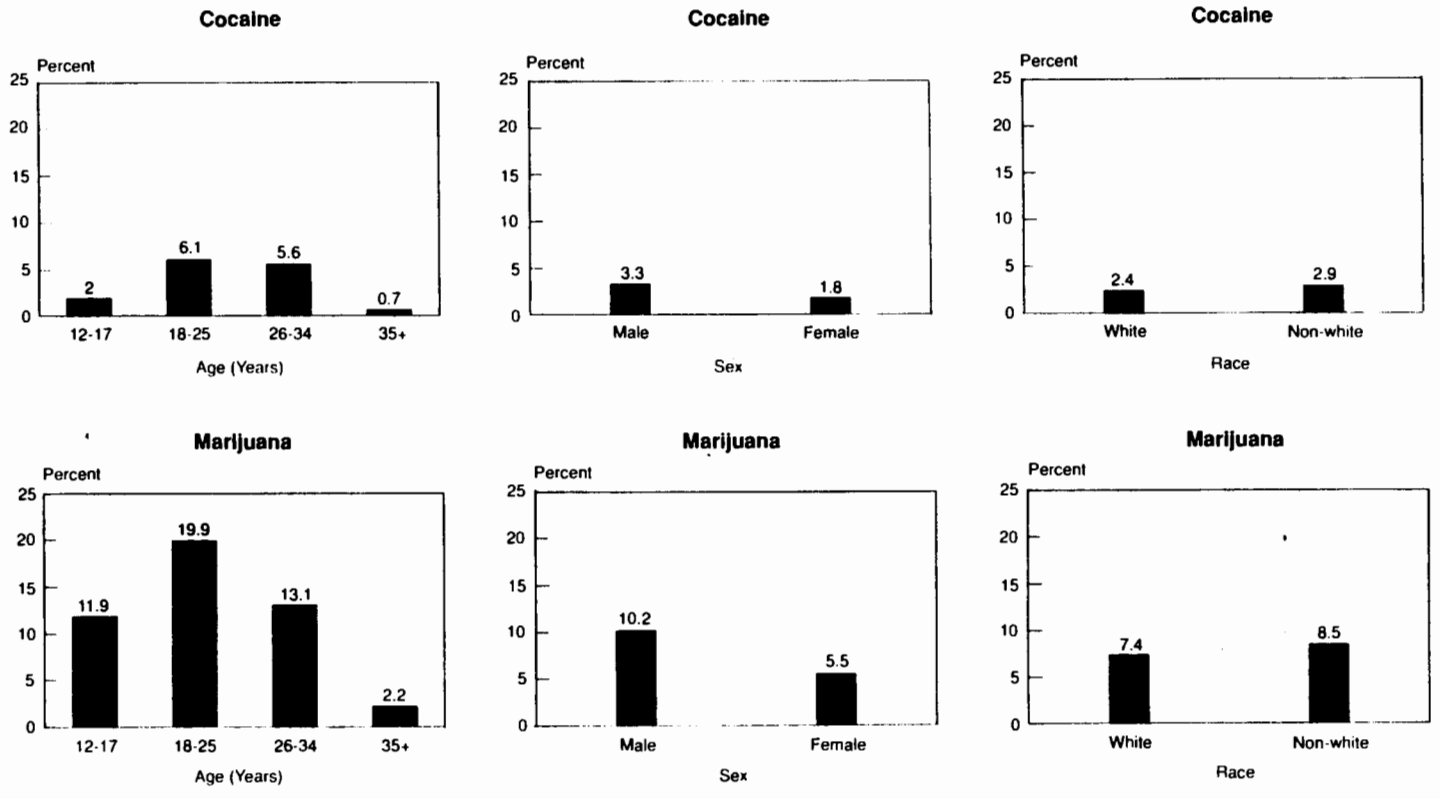


FIGURE 7-4 Estimated percentage of persons 12 years and older reporting cocaine and marijuana use in the past 12 months by age, sex, and race for all questionnaire versions combined.

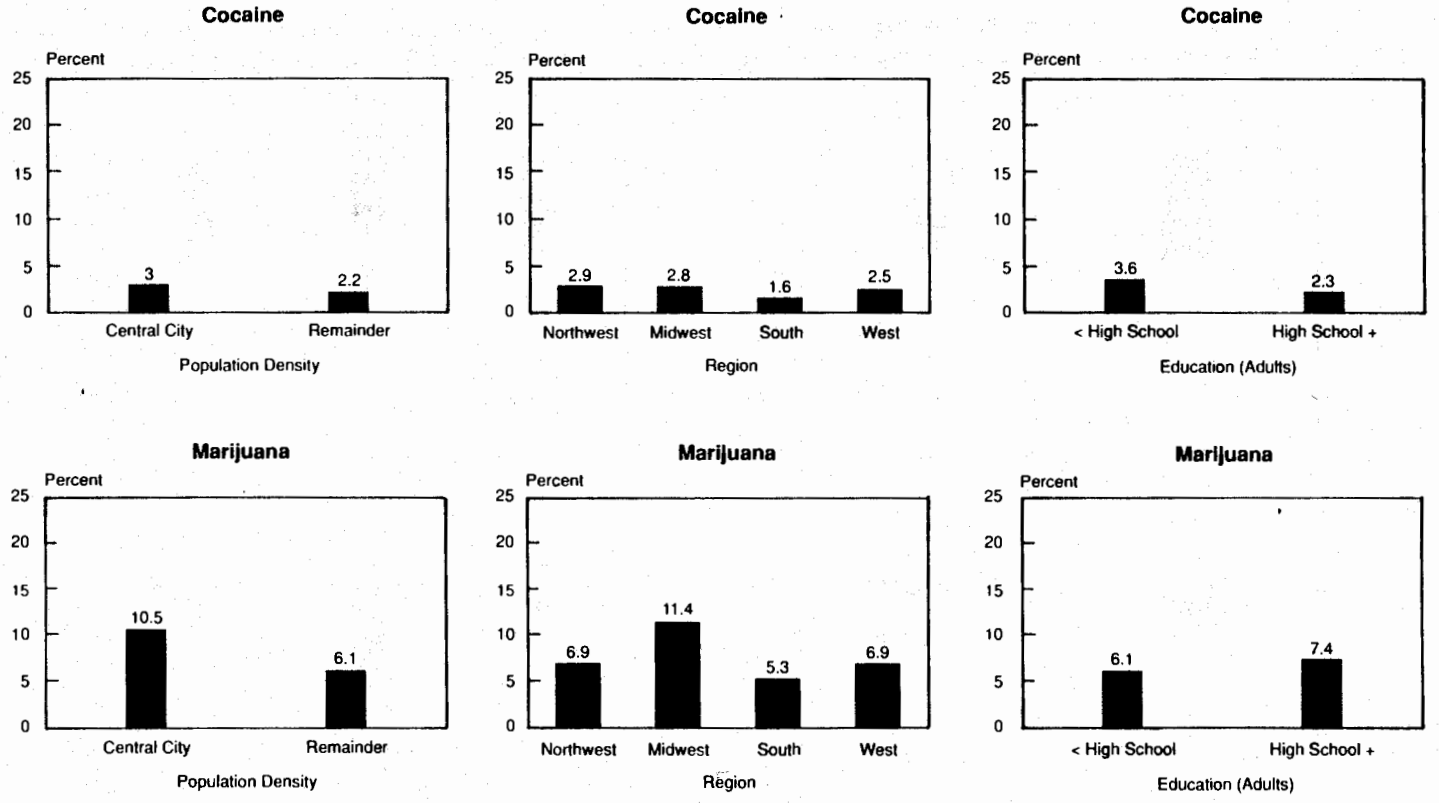


FIGURE 7-5 Estimated percentage of persons 12 years and older reporting cocaine and marijuana use in the past 12 months by population density, region, and education for all questionnaire versions combined.

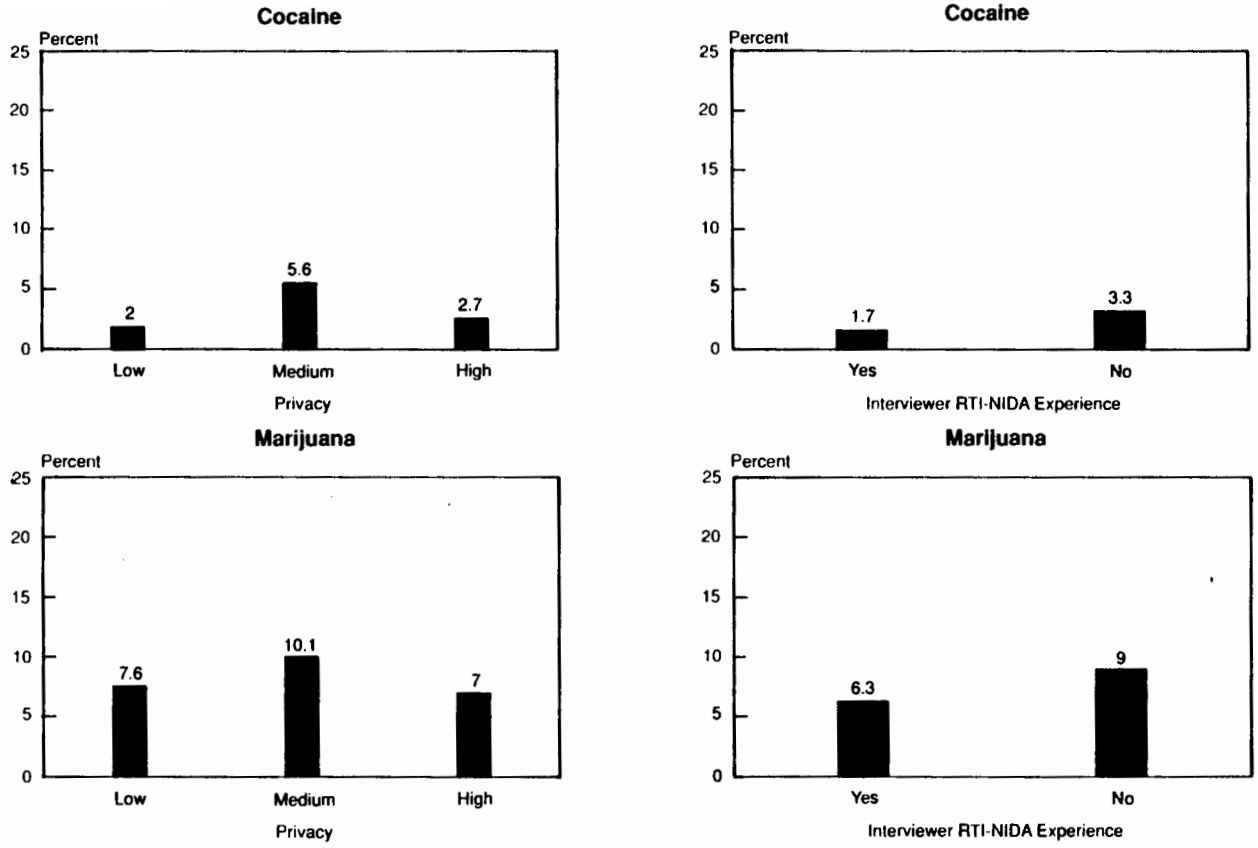


FIGURE 7-6 Estimated percentage of persons 12 years and older reporting cocaine and marijuana use in the past 12 months by interview privacy and interviewer RTI-NIDA experience for all questionnaire versions combined.

TABLE 7-2 Estimated Percentage of Persons 12 Years and Older Reporting Drug Use by Interviewer Experience in Conducting Prior NIDA-RTI Surveys

	NIDA-RTI Experience		
Drug Use	Yes	No	Ratio[a]
Illicit Drugs			
Cocaine, 30 days	0.6	0.9	1.5
Cocaine, 12 months	1.7	3.3	1.9
Cocaine, lifetime	10.7	15.7	1.5
Marijuana, 30 days	3.6	4.2	1.2
Marijuana, 12 months	6.3	9.0	1.4
Marijuana, lifetime	33.6	38.7	1.2
Licit Drugs			
Alcohol, 30 days	51.6	56.0	1.1
Alcohol, 12 months	69.5	74.9	1.1
Alcohol, lifetime	86.2	86.8	1.0
Cigarettes, 30 days	25.1	23.3	0.9
Cigarettes, 12 months	28.6	27.9	1.0
Cigarettes, lifetime	71.5	70.6	1.0

[a] Ratio of prevalence estimates: rate for interviewers without NIDA-RTI experience divided by rate for interviewers with prior experience.

percent). This pattern was found consistently across all reporting of illicit drug use, as shown in Table 7-2. For licit drugs, however, the results are less dramatic:[14] the prevalence rates obtained by inexperienced interviewers were higher for alcohol use, and for use in the past 30 days or past year, by 4 to 5 percentage points. For cigarette use the differences are smaller and mixed in direction.

Although these differences are interesting, they cannot be viewed in as straightforward a manner as differences across experimental conditions. For the variables manipulated in the experiment (i.e., mode of administration and question wording), the expected value of drug outcome variables is equivalent across experimental conditions under the null hypothesis of no treatment effect, because of random assignment of experimental conditions. However, the prior experience of the survey interviewer and the privacy of the survey were not under experimental control. Thus, other factors may lie behind the patterns observed. Table 7-3 indicates, for example, the presence of nontrivial differences in the composition of samples assigned to and completed by interviewers with and without prior NIDA-RTI experience. For instance, interviewers without prior experience were more likely to work in the Midwest. Given

[14] Reports of alcohol use for 12- to 17-year-olds yielded estimated prevalence rates for lifetime use of 50.7 percent for interviewers without NIDA-RTI experience and 43.8 percent for interviewers with such experience. For alcohol use in the past year, the rates were 42.3 percent (no experience) versus 38.7 percent (experienced); for use in the past 30 days, the rates were 16.6 percent (no experience) versus 17.6 percent (experienced).

TABLE 7-3 Distribution of Social and Demographic Characteristics of Respondents (percent) by Previous Experience of Interviewers in NIDA-RTI Surveys

Characteristics of Respondents	NIDA-RTI Experience	
	No	Yes
Reside in central city		
No	72.4	59.9
Yes	27.6	40.1
Region		
Midwest	32.7	19.2
South	14.0	26.2
West	32.7	22.8
Northeast	20.6	31.8
Race		
Nonwhite	18.0	18.1
White	82.0	81.9
Age		
12–17	9.4	9.2
18–25	15.3	12.5
26–34	20.1	18.0
35+	55.2	60.4
Sex		
Female	57.8	53.3
Male	42.2	46.7
Education[a]		
Less than high school	11.8	17.0
High school or more	88.2	83.0

[a] Education is coded only for persons age 18 and older.

that reporting of marijuana use is also higher in the Midwest than in other regions, it is possible that regional variations in drug use might be incorrectly attributed to differences in interviewer experience (or vice versa).

Such considerations lead naturally to more formal multivariate analyses, which permit some control for counter-hypotheses, such as that noted above.

STATISTICAL TESTS

Our statistical analysis proceeded through a series of steps that reflected increasingly more appropriate assumptions for inferring effects likely to be found in the population at large and that, in turn, involved increasingly expensive computational procedures. Analysis at each stage was informed by the results of prior stages so that the complexity and number of models to be tested could be reduced. Thus, for example, we did not attempt to estimate model parameters for interaction effects unless prior analyses (assuming simple random sampling) rejected the hypothesis of no effect at the .10 significance level.

TABLE 7-4 Test of Main Effect of Mode of Administration on Prevalence of Self-Reported Drug Use in NHSDA Field Test

	In Sample[a]			In Population[a]		
Drug Use	L^2	df	p	J^2	df	p
Alcohol Use						
Past 30 days	2.80	1	<.10	0.63	1	ns
Past year	2.85	1	<.10	0.41	1	ns
Lifetime	0.22	1	ns	-0.83	1	ns
Marijuana Use						
Past 30 days	7.65	1	<.01	2.31	1	.01
Past year	4.64	1	<.05	1.22	1	.06
Lifetime	1.09	1	ns	-0.40	1	ns
Cocaine Use						
Past 30 days	5.30	1	<.05	1.46	1	.04
Past year	4.18	1	<.05	1.19	1	.07
Lifetime	0.32	1	ns	-0.81	1	ns

NOTE: Tests were performed by fitting log-linear models to the three-way distribution of reported drug use {D} by question wording {Q} and by mode of administration {M}. Tests used the procedures of Goodman (1967, 1978) as implemented and adapted to complex sample designs by Fay (1982, 1985, 1989). Tests for the main effect of mode of administration upon reporting of drug use assessed the improvement in fit for a model that included a term representing this main effect (model: {DM}, {QM}) compared to one that did not (model: {D}, {MQ}). See text for discussion of tests of effects found in sample and in population. ns: not significant or borderline (i.e., $p > .10$).

[a] L^2: likelihood ratio chi square; df: degrees of freedom; J^2: jacknifed likelihood ratio chi square.

In most of our analyses we have fit models estimating the reported use of four substances (cigarettes, alcohol, marijuana, and cocaine) in three time periods (30 days, 12 months, and lifetime). The models varied in the effects they included and the methods used to estimate these effects. The sequence of analyses reported here begins with a test of the main effects and the interaction of the two experimentally manipulated variables—mode of interview and question wording—on the reported prevalence of drug use; our analyses employ no covariate adjustments because these two factors were fully randomized (see Figures 7-1 and 7-3 for alcohol, marijuana, and cocaine use). We then report a series of analyses designed to explore more complex patterns of main effects and interactions involving variables that were not randomized in the experiment.

Effects of Experimentally Manipulated Variables

Tables 7-4 and 7-5 presents statistical tests of the main effects shown in Figures 7-1 and 7-3, plus those for reports of cigarette use. Tests of main

TABLE 7-5 Test of Main Effect of Questionnaire Wording on Prevalence of Self-Reported Drug Use in NHSDA Field Test

	In Sample[a]			In Population[a]		
Drug Use	L^2	df	p	J^2	df	p
Alcohol Use						
Past 30 days	.00	1	ns	-0.35	1	ns
Past year	.02	1	ns	-0.84	1	ns
Lifetime	3.37	1	<.10	0.32	1	ns
Marijuana Use						
Past 30 days	0.11	1	ns	-0.90	1	ns
Past year	0.02	1	ns	0.55	1	ns
Lifetime	0.18	1	ns	-0.91	1	ns
Cocaine Use						
Past 30 days	5.15	1	<.05	1.75	1	.03
Past year	1.65	1	ns	0.44	1	ns
Lifetime	1.05	1	ns	-0.28	1	ns

NOTE: Tests were performed by fitting log-linear models to the three-way distribution of reported drug use {D} by question wording {Q} and by mode of administration {M}. Tests used the procedures of Goodman (1967, 1978) as implemented and adapted to complex sample designs by Fay (1982, 1985, 1989). Tests for the main effect of questionnaire wording upon reporting of drug use assessed the improvement in fit for a model that included a term representing this main effect (model: {DM}, {QM}) to one that did not (model: {D}, {MQ}). See text for discussion of tests of effects found in sample and in population. ns: not significant or borderline (i.e., $p > .10$).

[a] L^2: likelihood ratio chi square; df: degrees of freedom; J^2: jacknifed likelihood ratio chi square.

and interaction effects were performed by contrasting the fit of alternative log-linear models to the three-way cross-classification of reported drug use *by* mode of administration *by* question wording. Models were fit using the procedures of Goodman (1967, 1978) as implemented in algorithms developed by Fay (1982, 1985, 1989). (Table 7-1 presented estimated prevalence rates and standard errors for the reporting of use of each drug by mode of survey administration and wording.)

In fitting these models, we employed two complementary strategies. The first treats the sample of respondents as a population that has been randomly assigned to a set of experimental conditions; it ignores the fact that the data also represent a complex sample of a larger population. This analysis seeks to determine whether the observed differences in the response distributions are sufficient to reject the null hypothesis that they reflect only random variations in the assignment of individuals to different experimental conditions.

The second strategy takes account of the fact that the data are derived from a complex sample of the population and seeks to determine whether

the observed differences in response distributions are sufficient to reject the null hypothesis that they reflect only random variations both in the assignment of individuals to different experimental conditions and in the sampling of individuals from the study population. The second analysis estimates the likelihood that in a very large number of replications of the combined sample and experimental designs, differences in response distributions as large (or larger) than those actually observed would be obtained.

Table 7-4 indicates, as expected, that the observed differences in prevalence in Figure 7-1 are reliable for reports of use of the most sensitive drugs. Thus, for example, the increased reporting of recent use of marijuana or cocaine in the self-administered condition is statistically reliable for intrasample comparisons. When the complex sample design is accounted for, the differences remain reliable for use of these drugs during the past 30 days. The differences (by mode of administration) for use of these drugs during the past 12 months are borderline ($p < .07$). In contrast, neither the reported prevalence of cocaine and marijuana use during the respondent's lifetime nor the reporting of alcohol use in the full sample appears to be affected by use of a self-administered form.

With one exception, there are no statistically noteworthy differences in the reported prevalence of drug use by questionnaire wording (see Table 7-5). That exception is the reporting of cocaine use during the previous 30 days. There, the new question wording yielded substantially lower rates of reported use (0.5 versus 1.2 percent).

Effects of Nonexperimental Variables and Interactions

As previously noted, bivariate analysis of the association between such factors as privacy of the interview situation and reporting of drug use can be misleading. If, for example, poorer households are both more crowded and more likely to contain a person who had once used crack, an entirely artifactual association between the privacy of the interview and reported crack use might be observed. For this reason, we investigated the effects of the two nonexperimental variables of interest (privacy and interviewer experience) as well as interactions between experimental and nonexperimental variables. Again, we have fit models that estimated the reported use of four substances (cigarettes, alcohol, marijuana, and cocaine) in each of three time periods (30 days, 12 months, and lifetime). The four additional stages of this analysis are described below.

Stages of Analysis

Stage 1. General linear modeling procedures were initially used to fit models[15] to test for the main effects of the experimentally manipulated and survey execution variables. These models predict the reporting of drug use as a function of

a. the two treatment variables (mode of administration and wording);
b. two operational characteristics of the interview: privacy and interviewer experience in previous NIDA-RTI surveys;
c. three characteristics of respondents: age, sex, and race;[16] and
d. 11 interaction effects consisting of 5 interactions reflecting the joint effects of wording and each of the five variables listed under b and c above; 5 interactions for mode of administration and these same variables; and an interaction of the respondent's age with the privacy of the interview.

Stage 2. Because our education codes were not meaningful for young respondents,[17] separate analyses were performed for persons age 18 and older. These analyses tested whether our experimental manipulations (i.e., wording and mode of administration) had different impacts depending on the educational level of the respondent. The procedures employed were the same as those described for Stage 1, except that education and its interactions with wording and mode of administration were included as independent variables, whereas other interactions were deleted.

Stage 3. The results of Stage 1 were employed to select a reduced model by using procedures that yield appropriate estimates of variance and statistical tests for inferences to populations when a complex sample design has been used.[18] Generally, the interactions listed under Stage 1(d) were included in these analyses only if we could reject the null hypothesis that the parameter was zero, with p the assumption of simple random sampling.

Stage 4. Finally, because the dependent variables in all of these analyses are binary and because some variables have means close to zero

[15]To accommodate the design's randomized assignment of treatments within sample segments, the blocking factor (segments) was also incorporated.

[16]With the exception of the 9-point privacy scale, all independent variables in these models were categorical and fit accordingly. Categories are those used in the figures and tables already presented (i.e., age: 12-17, 18-25, 26-34, 35+; race: nonwhite, white).

[17]Because the majority of 12 to 17 year olds will not have completed their educations.

[18]In Stage 1, we estimated models by using methods that do not take account of the complex sample design.

(e.g., cocaine use in past 30 days has an estimated prevalence of less than 1 percent), we repeated selected analyses from Stage 3 by using logistic rather than linear regression models.

Since the foregoing model estimation exercises were typically repeated for each of the 12 dependent variables and dozens of parameters were estimated in every instance, the results of all the analyses cannot be presented in detail. Instead, we begin by summarizing statistically significant findings, that is, by listing instances in which parameter estimates for effects involving the experimental or survey operation variables were found to be reliably different from zero.

Because the focus of this analysis is methodological, we do not discuss the covariation of drug use by the demographic and social variables included in our models. We do, however, discuss interactions in which the estimated effects of the experimental and operational variables in the Field Test were found to vary within demographic or social subgroups of the population.[19]

Stage 1: Linear Models—Basic Analysis

Main Effects. Table 7-6 presents the main effects on drug use reporting of two variables related to survey execution: the privacy of the interview and the prior experience of the interviewer. (Privacy of the interview was coded by the interviewer on a 9-point scale;[20] interviewer experience is a binary variable.)[21] Where the effects are estimated to be different from zero, with p less than .15, we show the parameter estimates for the main effects of the survey execution variables on the prevalence of reported drug use.

Table 7-6 indicates that the privacy of the interview is estimated to have reliable main effects on the reporting of alcohol use during all three time periods. Privacy was also found to have a main effect of borderline

[19]In addition, readers may note that analyses of reported cigarette use are not included. We have excluded them because the interpretation of the impact of the experimental variable does not parallel that for other drugs: questions on cigarette smoking were administered by interviewers in the nominally "self-administered" treatment that used current question wordings (Version A).

[20]Privacy was coded so that high values corresponded to low privacy. Values were assigned using a 9-point scale completed by the interviewer in response to the instruction: "Indicate on this scale of 01 through 09 how private the interview was." The following verbal labels were attached to scale points: 1—Completely private (no one was in the room or could overhear any part of the interview); 3—Minor distractions; 5—Person(s) in the room or listening about one-third of the time; 7—Serious interruption of privacy more than half of the time; 9—Constant presence of other person(s).

[21]That is, it was coded as 0 if the interviewer had no previous experience in NIDA-RTI drug surveys; it was coded as 1 if the interviewer had prior experience. For 109 respondents, information was not available on the prior experience of the interviewer. For these cases a separate parameter was fit in our models; estimates for this "missing interviewer information" parameter are not included in the tables.

TABLE 7-6 Parameter Estimates and Tests of Significance from Stage 1 Analysis of Main Effects of Non-Experimental Survey Execution Factors on Reporting of Drug Use in the NHSDA Field Test

	Main Effects	
Drug and Time Period	Privacy[a]	Interviewer Experience[b]
Cocaine		
Past 30 days	ns	ns
Past year	ns	-.037 (p = .038)
Lifetime	ns	ns
Marijuana		
Past 30 days	ns	ns
Past year	ns	ns
Lifetime	-.006 (p = .076)	ns
Alcohol		
Past 30 days	-.012 (p = .001)	ns
Past year	-.007 (p = .025)	ns
Lifetime	-.009 (p < .001)	ns

NOTES. Results are derived from Stage 1 using general linear model analyses without taking account of complex, differentially weighted sample design. The table presents p-values and estimates where $p < .15$. Models were estimated without interaction terms; table shows only those main effects estimated to be different from zero with $p < .15$. Entry of ns in table indicates: not significant or borderline (i.e., $p > .10$).

[a]Privacy is coded so that high values correspond to *low* privacy. Values were assigned using a 9-point scale completed by the interviewer in response to the instruction: "Indicate on this scale of 01 through 09 how private the interview was." The following verbal labels were attached to scale points: 1—Completely private (no one was in the room or could overhear any part of the interview); 3—minor distractions; 5—person(s) in the room or listening about one-third of the time; 7—serious interruptions of privacy more than half of the time; 9—constant presence of other person(s).

[b]Parameter estimates are the estimated main effect for interviewers with *no prior experience* in NIDA-RTI drug surveys (parameter estimates for the 109 of 3,284 instances in which data on the interviewer's prior experience were unavailable are not shown here).

significance on reporting of lifetime marijuana use. In interpreting the main effects of this variable, bear in mind that privacy was analyzed as a continuous variable assessed on a 9-point scale and was entered as a linear term in our models. The parameter estimates shown in Table 7-6 are coefficients that would be multiplied by the respondent's score on the 9-point scale used to describe the privacy of the interview. It should also be carefully noted that the coding of this variable assigns higher numbers to interviews with less privacy. Thus, the negative parameter estimates

shown in Table 7-6 indicate that the main effects for privacy take the form of a decline in reporting of drug use when the interview becomes less private. Table 7-6 also indicates an estimated reliable main effect of previous interviewer experience with NIDA-RTI drug surveys, but only for reporting of cocaine use in the past year.

Overall, this analysis and the prior analysis of experimentally manipulated variables suggest that two variables, mode of administration and privacy of the interview, had pervasive *main* effects on reported prevalence of drug use. Both variables operated in the expected direction in every instance. Less public contexts for responding encouraged greater frequency of reports of drug use.

Interaction Effects. It is reasonable to suspect that experimental and survey execution variables will not have consistent effects on reporting of drug use in all situations for all classes of respondents. We therefore tested for a variety of interaction effects involving experimental and survey execution variables.

Table 7-7 presents the results of an expanded analysis that incorporated 11 potential interaction effects into the model used in the preceding section. This table indicates the significance levels for tests of the null hypothesis that the interaction effect was zero. Interaction effect estimates for the reference cell parametrization used in our analyses are not easily interpreted and are not included in Table 7-7. Our interpretation of the statistically significant interactions instead revolves around examination of the patterns of raw prevalence rates that give rise to the finding of significant interaction.

Table 7-7 indicates that the most common interaction across models was between mode of interview and whether the interviewer had experience in NIDA-RTI surveys. For reports of the use of cocaine and marijuana, this interaction effect was estimated to be reliably different from zero ($p < .05$) in four of six instances. In the other two instances, the results are suggestive (p's $< .12$) of the same interaction effect.

Figure 7-7 plots the prevalence rates obtained for each measure of marijuana and cocaine use by interviewer experience and mode of administration. These plots make clear the nature of the interaction that is occurring. When self-administered versions of the survey were employed, interviewers with no prior NIDA-RTI experience obtained more reports of drug use than experienced interviewers. In contrast, there is no consistent difference in prevalence rates for interviewer-administered surveys conducted by interviewers with and without previous NIDA-RTI experience. This finding, which is admittedly perplexing, was analyzed further to test a potential explanation of this effect.

TABLE 7-7 Tests in Stage 1 Analysis of Interaction Effects Involving Experimental and Survey Execution Factors on Reporting of Drug Use in the NHSDA Field Test

	Reporting of Drug Use During		
Drug and Interaction Effect	30 Days	1 Year	Lifetime
Cocaine			
Wording by mode		$p = .083$	
Privacy of interview[a] by age	[b]	[b]	
Mode of administration by age			$p = .063$
Mode of administration by interviewer experience	$p < .001$	$p = .109$	$p = .015$
Wording by race	$p = .019$		
Wording by sex			$p = .016$
Wording by interviewer experience			$p = .032$
Marijuana			
Wording by mode			$p = .103$
Privacy of interview[a] by age		$p = .040$	$p = .050$
Mode of administration by age		$p = .075$	
Mode of administration by privacy[a]		$p = .042$	
Mode of administration by interviewer experience	$p = .001$	$p = .119$	$p = .048$
Wording by age	$p = .065$		
Alcohol			
Wording by mode		$p = .079$	
Privacy of interview[a] by age		$p = .119$	
Mode of administration by age		$p = .007$	$p = .001$
Mode of administration by privacy[a]		$p = .094$	$p = .097$
Wording by age			$p = .001$
Wording by sex	$p = .035$		
Wording by privacy[a]	$p = .004$	$p = .108$	

NOTES. Results are derived from Stage 1 using general linear model analyses without taking account of complex, differentially weighted sample design. The table shows only interaction effects estimated to be different from zero with $p < .15$. Survey execution and experimental factors and interactions included in these models were wording, mode of administration, wording *by* mode of administration, privacy of interview, interviewer experience with previous NIDA-RTI surveys, mode of administration *by* age, mode of administration *by* race, mode of administration *by* sex, mode of administration *by* privacy, mode of administration *by* interviewer experience, wording *by* age, wording *by* race, wording *by* sex, wording *by* privacy, and wording *by* interviewer experience.

[a]Privacy is coded so that high values correspond to *low* privacy. Values were assigned by using a 9-point scale completed by the interviewer in response to the instruction: "Indicate on this scale of 01 through 09 how private the interview was." The following verbal labels were attached to scale points: 1—completely private (no one was in the room or could overhear any part of the interview); 3—minor distractions; 5—person(s) in the room or listening about one-third of the time; 7—serious interruptions of privacy more than half of the time; 9—constant presence of other person(s).

[b]Probability level of less than .05 was found but is not shown because interaction involved one or more estimated prevalence rates of zero.

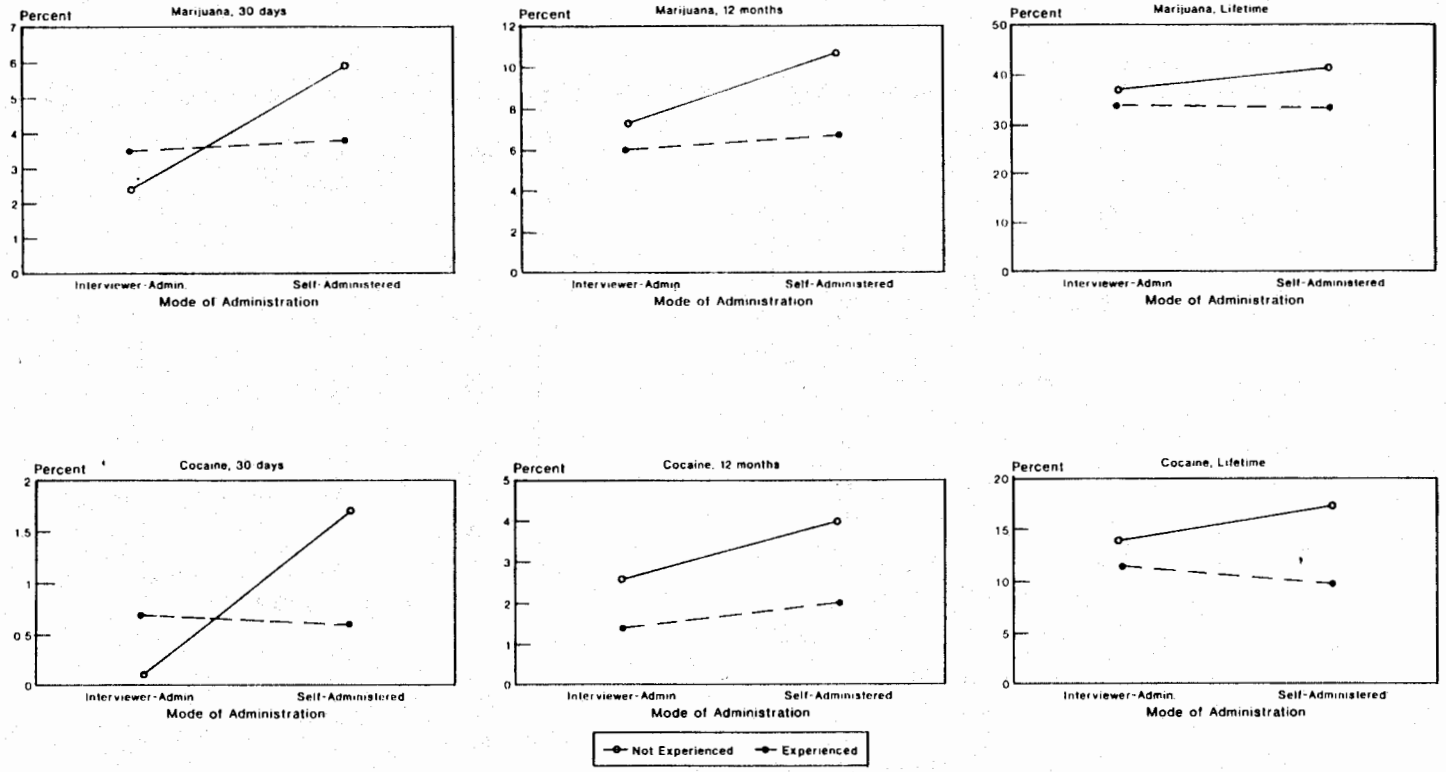


FIGURE 7-7 Estimated percentage of persons 12 years and older reporting cocaine and marijuana use in the past 30 days, past 12 months, and lifetime by mode of administration and interviewer experience.

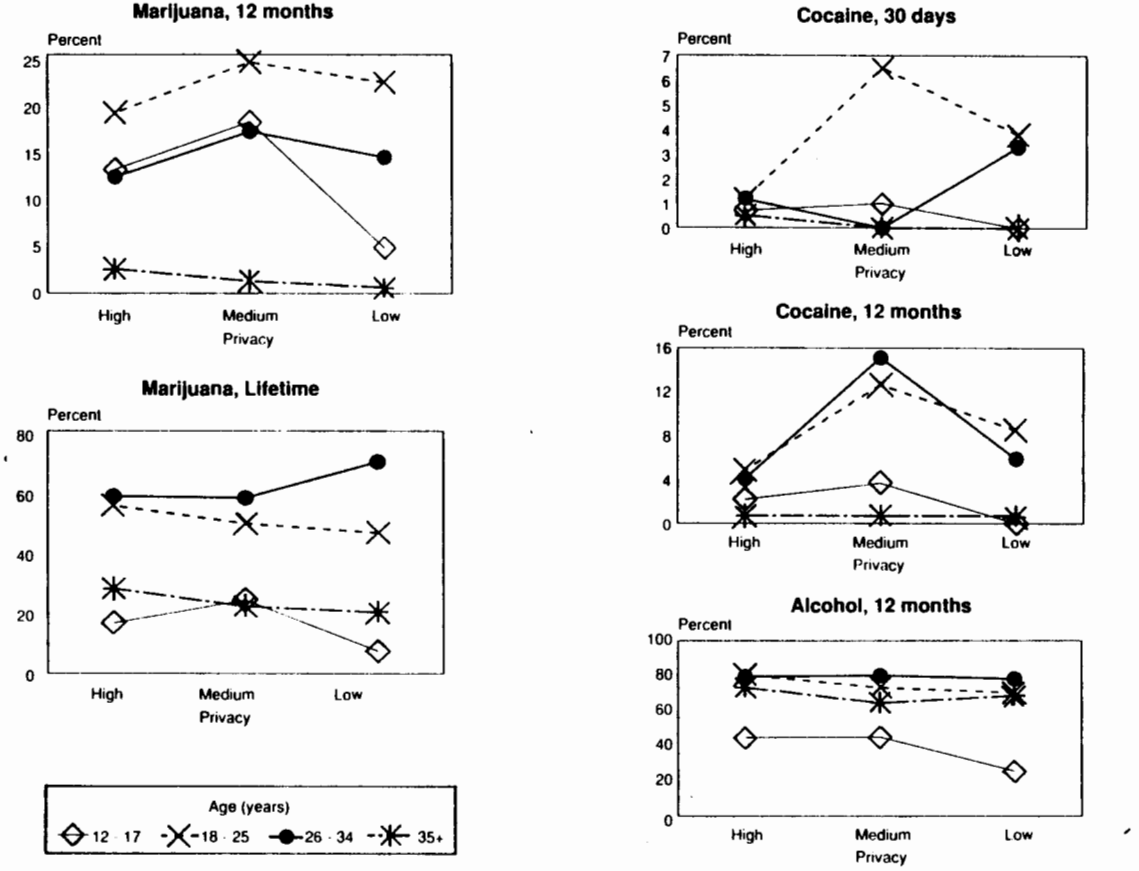


FIGURE 7-8 Estimated prevalence rates for alcohol, cocaine, and marijuana use by interview privacy and age group.

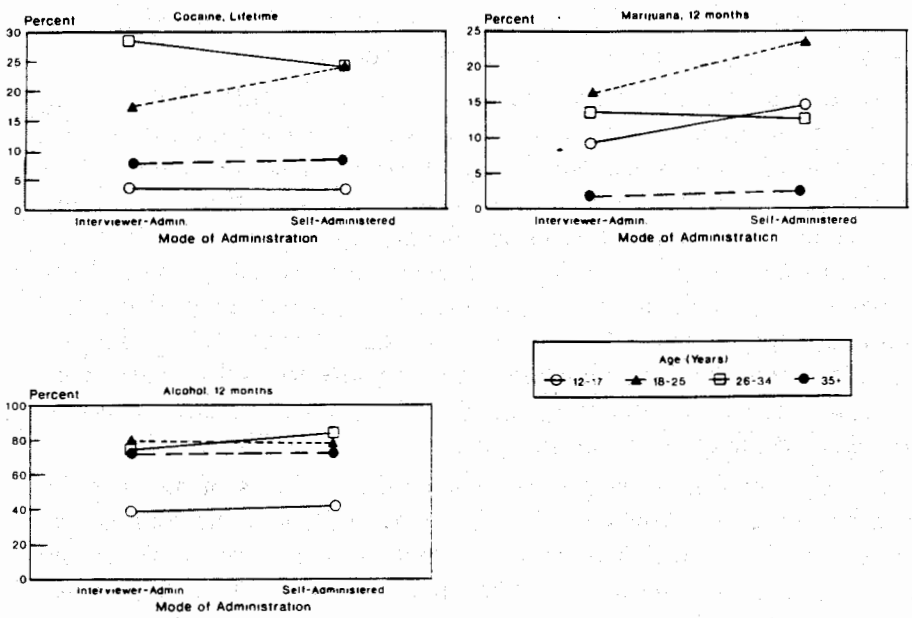


FIGURE 7-9 Estimated prevalence rates for cocaine, marijuana, and alcohol use by age group and mode of administration.

The "privacy of interview" variable also displays a number of significant interactions. For alcohol use in the past year, for example, there is an interaction of privacy with age ($p = .119$). Figure 7-8 shows prevalence rates for alcohol use in the past year and five other measures of drug use that exhibit similar interactions. These plots of prevalence of reported drug use by age and by privacy of the interview suggest that teenagers are particularly sensitive to the presence of another person and, in such circumstances, are much less likely to report drug use. Because the other person may be a parent, this result is not surprising.

A parallel phenomenon is observed for the interaction of age and mode of survey administration. Here (see Table 7-7), significant interaction occurs for the reporting of alcohol use during the past year, whereas the effects are borderline ($.06 < p < .07$) for marijuana use in that period. Again, the plots shown in Figure 7-9 suggest that this interaction reflects the greater sensitivity of younger persons to the mode of survey administration. Note, for example, that there is substantially more reporting of marijuana use during the past year on the self-administered questionnaire for persons 12 to 17 and 18 to 25 years of age. The estimated prevalence of reported marijuana use increases from 9.3 and 16.3 percent, respectively, in the interviewer-administered versions to 14.5 and 23.5 percent, respectively, when a self-administered questionnaire is used. The latter,

however, offers little advantage in obtaining reports of marijuana use during the past year from persons 26 and older.

Reporting of alcohol use shows a somewhat similar pattern. The prevalence estimate for reported use of alcohol in the past year is 38.8 percent for 12- to 17-year-olds in the interviewer-administered survey questionnaire; it increases to 42.1 percent in the self-administered version. In contrast, rates of reported alcohol use by persons 18 to 25 and 35 or older are virtually identical under the two conditions. There is, however, one exception to this pattern: estimated rates of reported alcohol use for persons 26 to 34 years of age increase from 74.5 percent in the interviewer-administered questionnaire to 84.4 percent in the self-administered questionnaire. Figure 7-10 shows prevalence rates corresponding to the question wording interactions from Table 7-7.

It should be noted that the 2 × 2 interaction effects of mode of administration and wording shown in Figure 7-10 identify the four experimental conditions. Such interactions reflect the joint effects of wording and mode of administration, and they may reflect the fact that *similar question wording* means little if the mode of administration differs. Thus, if anything, the only surprise in our results was that there were relatively few such interaction effects.

The two mode-by-wording interactions that are plotted in Figure 7-10 involve a reversal of the direction of effects that were obtained. Thus, the old wording of the questionnaire yielded slightly higher reporting of lifetime marijuana use in the self-administered format (estimated prevalences: 35.7 versus 37.7 percent), whereas the new wording yielded higher prevalence estimates in the interviewer-administered format (36.6 percent for the new wording versus 33.2 percent for the old wording). A similar reversal is noted for sex-by-wording effects on the reporting of alcohol use during the past 30 days (see Figure 7-10). For females, the estimated prevalence increased when the new wording was employed, but it declined for males.

Stage 2: Linear Models Including Education

For persons age 18 and older, we undertook a parallel analysis introducing education as a model variable. Education is a binary variable indicating whether the respondent graduated from high school or not. Respondents who were 12 to 17 years old were not included in the analysis because this variable would not be meaningful for them. The models we estimated were intended to test whether the effects of experimentally manipulated factors (wording and mode of administration) varied by education. We were interested, for example, in whether the effects on prevalence rates

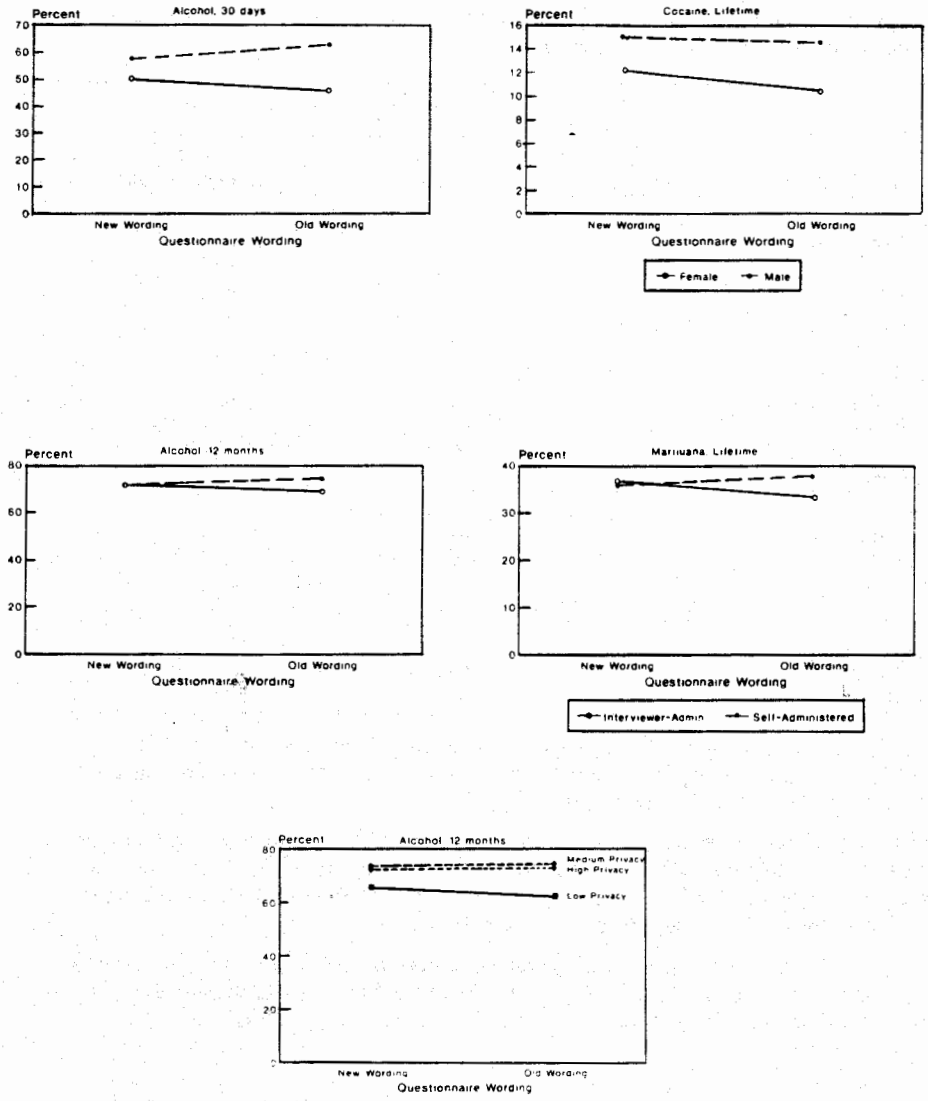


FIGURE 7-10 Estimated prevalence rates for drug use measurements showing noteworthy question wording interactions in Table 7-7.

TABLE 7-8 Tests of Significance and Parameter Estimates for Interaction Effects of Education and Wording and Education and Mode of Administration on Prevalence of Reported Drug Use (from Stage 1 Analyses)

	Interactions of Education and	
Drug and Time Period	Wording[a]	Mode of Administration[b]
Cocaine		
Past 30 days	ns	p = .065 .020
Past year	ns	ns
Lifetime	p = .140 .061	ns
Marijuana		
Past 30 days	ns	ns
Past year	p = .127 -.047	p = .153 -.045
Lifetime	ns	ns
Alcohol		
Past 30 days	ns	ns
Past year	ns	ns
Lifetime	ns	ns
Cigarettes		
Past 30 days	ns	ns
Past year	ns	ns
Lifetime	ns	ns

NOTE: The sample for this analysis was restricted to persons age 18 and older. See the text for a description of other terms included in the model. Entry of ns in table indicates: : not significant or suggestive (i.e., $p > .20$).

[a]Estimates are from a reference cell parametrization of the interaction effect and represent the difference in the two simple effects of wording (new minus old) measured separately for the less than high school (<HS) and high school or more (HS+) education levels. The analysis is restricted to persons age 18 and older.

[b]Estimates are from a reference cell parametrization of the interaction effect and contrast the two simple effects of mode (interviewer- minus self-administered) at the two educational levels. The analysis is restricted to persons age 18 and older.

of different questions were equivalent for persons of different educational backgrounds.

The model used in this analysis included the following as main effects: education, age, race, sex, privacy, interviewer experience with NIDA-RTI surveys, mode of administration, and question wording. In addition, interaction terms were included for education by mode of administration and education by questionnaire wording.

Table 7-8 presents the results of tests for the interaction effects of education and our experimental variables on the prevalence of reported drug use. From this table it can be seen that only 4 of the 24 interactions

involving education were statistically noteworthy, which suggests[22] that the mode of administration or wording of questions did not have many differential effects on persons with varying levels of education.[23] Those effects that did appear, however, all involved reporting of illicit drug use.

Figure 7-11 plots the estimated prevalence rates that correspond to the four education interactions found to be statistically noteworthy in Table 7-8. These plots indicate that the two experimental variables appear to have a more pronounced effect on reporting of drug use by persons with less than a high school education. The self-administered questionnaires yielded a markedly larger increase in reporting of cocaine use in the past 30 days, and marijuana use in the past year, by persons who had not graduated from high school. In contrast, the new questionnaire wording produced markedly decreased reporting of marijuana use in the past year for this same group.

Stage 3: Reduced Models Taking Account of Complex Sample Design

The next step in our analysis used statistical procedures that took into account the complex sample design in determining the variances of model estimates. For these analyses, information gained in previous steps was used to reduce the number of terms entered into the models. Readers can note, for example, that 11 interaction terms were estimated for each model in the first stage of analysis. Less than half of these interactions, on average, yielded parameter estimates that were reliably different from zero—even under the assumption of simple random sampling. Model complexity and computational expense were substantially reduced by using results from the first stage of analysis to prune our models.

Main Effects. Initially we began by fitting models to test for the main effects of the two experimentally manipulated variables (mode of administration and wording) and for NIDA-RTI experience and interview privacy. Table 7-9 presents the statistical tests and parameter estimates (if $p < .20$) representing the main effects of these variables on the prevalence estimates for reported alcohol, marijuana, and cocaine use. (Other independent variables included in the model were sex, race, age, region, and whether respondent resided in the central city of an MSA.)

Contrary to expectations, the effect of taking into account the complex sample design did not consistently reduce the number of instances in

[22] Clearly, the possibility cannot be ruled out that the interactions were of a higher order than those tested here.

[23] The education classification used in this analysis is quite broad and distinguishes only between persons who have graduated from high school and those who have not. Interaction effects might be detectable if a more finely graded categorization were used.

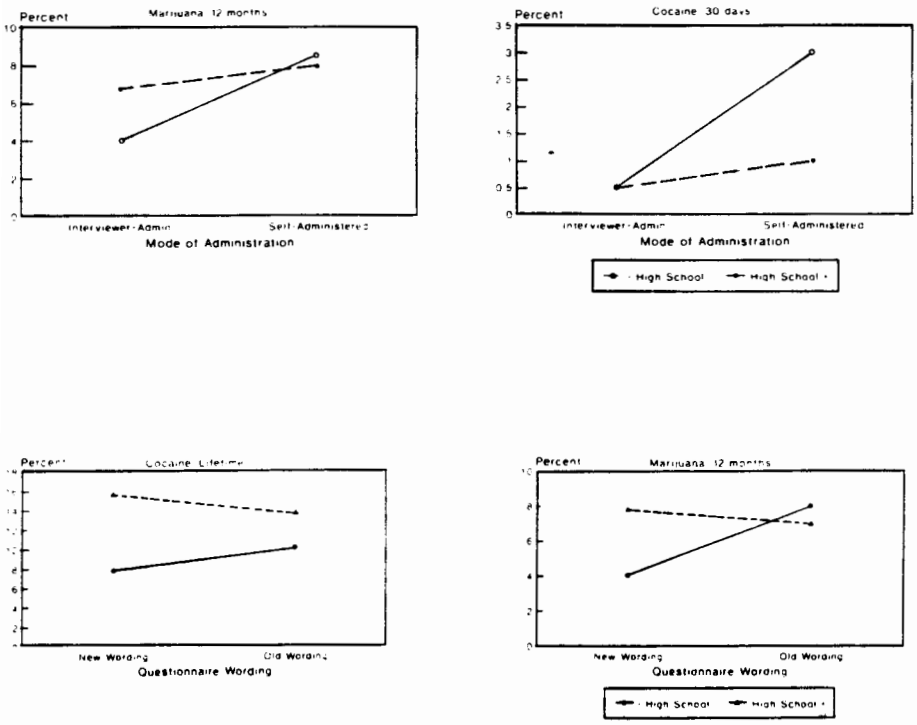


FIGURE 7-11 Estimated prevalence rates for four education interactions shown in Table 7-8.

TABLE 7-9 Experimental and Survey Execution Factors Estimated in Stage 2 Analyses to Have a Nonzero Main Effect (with $p < .20$) on Prevalence of Drug Use Reported by Respondents in the NHSDA Field Test

	MAIN EFFECTS			
Drug and Time Period	Questionnaire Wording[a]	Mode of Administration[b]	Privacy[c]	NIDA-RTI Experience[d]
Alcohol				
Past 30 days	ns	ns	$p = .002$ -.015	$p = .125$ .101 (no experience) .149 (no information)
Past year	ns	ns	$p = .043$ -.009	$p = .061$ .048 (no experience) .046 (no information)
Lifetime	$p = .108$ -.024	ns	$p = .021$ -.009	ns
Marijuana				
Past 30 days	ns	$p = .039$ -.017	ns	ns
Past year	ns	$p = .18$ -.015	ns	$p = .063$ .019 (no experience) .076 (no information)
Lifetime	ns	ns	$p = .155$ -.006	$p = .078$ .043 (no experience) .102 (no information)
Cocaine				
Past 30 days	$p = .083$ -.007	$p = .081$ -.007	$p = .111$ .001	ns
Past year	ns	$p = .185$ -.008	$p = .044$ .002	$p = .012$ .012 (no experience) .068 (no information)
Lifetime	ns	ns	ns	$p = .012$ .037 (no experience) .072 (no information)

NOTES. Model fit contained *only* main effects. Results are derived from Stage 2 analyses that take account of complex sample design employed. The table presents p-values and estimated coefficients for main effects if $p < .20$. Entry of ns in table indicates: not significant or suggestive (i.e., $p > .20$).

[a]Wording effects are estimated net deviations in reported prevalence of drug use expected for new question wording.

[b]Mode of administration effects are the estimated net deviations in prevalence of drug use expected for interviewer administration of the questionnaire.

[c]Effects of NIDA-RTI experience are the estimated net deviations in prevalence of drug use expected for interviewers who have no prior experience with NIDA-RTI surveys and for interviewers who did not provide this information. (For 109 of 3,284 respondents, information on interviewer experience was unavailable.)

[d]Privacy is coded so that high values correspond to *low* privacy. Values were assigned using a 9-point scale completed by the interviewer in response to the instruction: "Indicate on this scale of 01 through 09 how private the interview was." See the notes to Table 7-6 for the labels attached to categories. Entries for effects in the above table are coefficients representing the expected net difference in prevalence associated with a unit change on this privacy scale.

which parameters were estimated to be significantly different from zero. Thus, for mode of administration, we find a reliable ($p < .05$) effect on reporting of marijuana use during the past 30 days. We also find three borderline effects ($.19 < p < .05$) of the mode of administration on the reporting of marijuana use in the past year and of cocaine use in the past month. Parameter estimates indicate that the net effect of the use of a self-administered form is a 1.5 to 1.7 percentage point increase in the prevalence of reported marijuana use and a 0.7 to 0.8 percentage point increase in the prevalence of reported cocaine use during the past month and the past year.

The other experimental variable, wording of survey questions, yields two borderline main effects. The new question wording is estimated to have a net negative effect on reporting of lifetime alcohol use (estimated net effect = –2.4 percentage points, $p = .11$) and cocaine use in the past 30 days (–0.7 percentage points, $p = .08$).

Privacy of the survey interview shows consistently reliable ($p < .05$) effects on reporting of alcohol use. The reported prevalence of alcohol use decreases as interview privacy decreases. A 1-point difference in privacy (on the 9-point privacy scale) has an estimated net effect of between 0.9 and 1.5 percentage points on the prevalence of self-reported alcohol use. Significant main effects of privacy are also found for reporting of cocaine use in the past year; a borderline effect is found for reporting of cocaine use in the past 30 days.

Most surprisingly, estimates of the main effects on prevalence rates of interviewers' previous experience in NIDA-RTI surveys are now significantly ($p < .05$) different from zero in two instances and of borderline significance ($p < .13$) in three additional instances. All of the estimated net main effects indicate that interviewers without prior experience in NIDA-RTI surveys obtained higher prevalence rates of reported drug use.

Interaction Effects. Table 7-10 presents tests for the effects on reporting of drug use of selected interactions of experimental and survey operational variables and respondent characteristics after linear adjustments for differences in age, sex, race, region, and central city distributions of the experienced and inexperienced interviewers' work assignments. As before, models were estimated for alcohol, marijuana, and cocaine use during each of the three time periods. Selection of the interaction terms included in each model was based on the results of the models fit in Stage 2. In these analyses, however, procedures are employed that account for the complex sample design of the survey.

TABLE 7-10 Tests in Stage 2 Analysis for Interaction Effects Involving Experimental and Survey Execution Factors on Reporting of Drug Use in the NHSDA Field Test

	Reporting of Drug Use During		
Drug and Interaction Effect	30 Days	1 Year	Lifetime
Cocaine			
Wording by mode	ns	$p = .103$	ns
Privacy[a] by age	b	c	
Mode of administration by age			ns
Mode of admin. by interviewer experience	$p = .054$	ns	$p = .068$
Wording by race	$p = .141$		
Wording by sex			ns
Wording by interviewer experience			$p = .043$
Marijuana			
Wording by mode	ns	ns	ns
Privacy[a] by age		ns	ns
Mode of administration by privacy[a]		$p = .035$	
Mode of administration by age		$p = .155$	
Mode of admin. by interviewer experience	$p = .016$	$p = .189$	$p = .106$
Wording by age	ns		
Alcohol			
Wording by mode	ns	ns	ns
Privacy[a] by age		ns	$p = .182$
Mode of administration by privacy[a]		$p = .062$	$p = .182$
Mode of administration by age		ns	
Mode of admin. by interviewer experience	ns		
Wording by race			
Wording by sex	$p = .014$		
Wording by privacy	$p = .077$	ns	
Wording by age			$p = .004$

NOTES. Results are derived from Stage 2 analysis of reduced models using estimation procedures that take account of the survey's complex sample design. The table presents p-values where $p < .20$ and results for every interaction effect tested in the reduced models. The entry "ns" in the table indicates that the null hypothesis that the interaction effect was zero could not be rejected with $p < .20$. If no entry appears in the table, the specified interaction effect was not included in the reduced model for the dependent variable. The selection of parameters to include in these models was based on results of general linear models fit in Stage 1.

[a]Privacy is analyzed as a linear variable coded so that high values correspond to *low* privacy. See the note to Table 7-6 for further description of this variable.

[b]Probability level of less than .15 was found but is not shown because the interaction involved one or more estimated prevalence rates of zero.

[c]Probability level of less than .05 was found but is not shown because the interaction involved one or more estimated prevalence rates of zero.

Table 7-10 shows the results of these analyses: *n.s.* indicates that the model was fit and that the null hypothesis regarding the interaction could not be rejected. If the hypothesis could not be rejected with at least $p < .20$, then the corresponding interaction effect was deleted from the

reduced model that predicted reporting of that particular drug and time period (e.g., marijuana use in the past 30 days).

The effect of accounting for the complex sample design reduces somewhat the number of interactions that are reliably different from zero. Note that by taking the sample design into account in estimating the variances in Table 7-10, we are effectively asking whether the effects of our experimental manipulation can be reliably generalized to the population at large. In contrast, the models fit without accounting for the complex sample design (Table 7-6) ask whether there are differences between the subsamples—compared among themselves—that are too large to be attributed to the random assignment of individuals to the four experimental conditions.

The interaction that appears with greatest regularity in Table 7-10 is for the joint effect of mode of administration and interviewer experience in past NIDA-RTI surveys. This interaction was included in six of the models and was found to be of at least borderline significance in five instances, with p-values ranging from .016 to .189. Table 7-10 also indicates that for all three instances in which the interaction effects of mode of administration and interview privacy were included in the model, they were found to have at least borderline significance (p-values from .035 to .182).

Effects of the interaction of mode of administration and age were tested in three models. A weak effect on reporting of marijuana use in the past year was found ($p = .16$). In six instances we also fit terms for interactions between privacy and respondents' age. Two of these effects were significant with $p < .05$ and two were borderline ($.10 < p < .20$) when the complex sample design was accounted for.

Stage 4: Logistic Regression Models

Because the reported prevalence of cocaine use in the past 30 days and past year was less than 10 percent, model-fitting using OLS regression procedures for this variable is particularly problematic. We thus performed one final set of analyses fitting logistic regression models. These models were fit, as in the previous stage, by using computational methods that take account of the complex sample design in estimating variances.

Initially, a model was fit that included only main effects for the survey execution and experimentally manipulated variables (this model also included main effects for age, sex, race, region, and residence in a central city). For cocaine use in the past year, which has an overall prevalence of 2.5 percent, the results for the logistic model parallel those

reported in Table 7-9.[24] Thus, reliable ($p < .05$) main effects of privacy and NIDA-RTI experience were found for the reporting of cocaine use in the previous year, and a borderline ($p = .17$) effect was found for mode of administration. For cocaine use in the past 30 days, which has a reported prevalence of less than 1 percent, results for the logistic regression model diverge somewhat. Significant main effects of mode of administration ($p = .04$) and NIDA-RTI experience ($p = .02$) were found by using the logistic model. In addition, an almost significant ($p = .06$) effect was found for question wording, and a borderline effect for privacy (parameter estimates indicate that the estimated net main effect of the new wording is to reduce the reported prevalence of cocaine use in the previous 30 days).

Estimates of the interaction effects for our model of cocaine use in the past 12 months parallel those reported in Table 7-10 with one exception. When the logistic form is used, the interaction effect for privacy by age becomes statistically unreliable ($p = .23$). For cocaine use in the past 30 days, the results of the logistic regression are also similar to those from the linear model with two exceptions. The significance level for the interaction of mode of administration and NIDA-RTI experience increases to $p = .035$, and the significant privacy-by-age interaction found in Table 7-7 becomes statistically unreliable.

CONCLUSIONS AND CONUNDRUMS

Mode of Administration

Although the results described above are not always consistent for every substance examined, on balance the results indicate that having interviewers administer the questionnaire (without the use of self-administered answer sheets) reduces the reporting of drug use. This conclusion is supported by the finding that lack of privacy during an interview had a negative effect on the reporting of drug use, particularly for respondents 12 to 17 years of age for whom the person present is likely to be a parent.

Test of Alternate Explanation

It has recently been suggested to the authors that the elevated reporting of drug use by 12- to 17-year-olds on the self-administered questionnaire might be due to the fact that this age group made a large number of marking errors that resulted in their being classified as users. We examined

[24] In the former analysis, a linear regression model was fit using procedures that take account of the complex sample design.

TABLE 7-11 Estimated Percentage of Persons Age 12 to 17 Reporting Use of Alcohol, Cigarettes, Marijuana, and Cocaine by Consistency of Reporting[a]

Drug Use	Self Administered	Interviewer Administered	Ratio Self/Interviewer
Cocaine, lifetime			
Consistent respondents	3.2	3.0	1.1
All respondents	3.3	3.5	0.9
Cocaine, past year			
Consistent repondents	1.9	1.8	1.1
All respondents	2.3	1.7	1.4
Cocaine, past 30 days			
Consistent respondents	1.0	0.5	2.0
All respondents	0.7	0.5	1.4
Marijuana, lifetime			
Consistent respondents	17.4	11.3	1.5
All respondents	19.8	11.8	1.7
Marijuana, past year			
Consistent respondents	13.8	9.0	1.5
All respondents	14.5	9.3	1.6
Marijuana, past 30 days			
Consistent respondents	5.1	3.1	1.6
All respondents	5.2	3.2	1.6
Alcohol, lifetime			
Consistent respondents	39.1	43.6	0.9
All respondents	47.7	46.5	1.0
Alcohol, past year			
Consistent respondents	34.5	36.4	0.9
All respondents	42.1	38.9	1.1
Alcohol, past 30 days			
Consistent respondents	17.6	15.1	1.2
All respondents	20.0	14.7	1.4
Cigarettes, lifetime			
Consistent respondents	28.9	29.9	1.0
All respondents	33.5	33.1	1.0
Cigarettes, past year			
Consistent respondents	21.4	14.8	1.4
All respondents	20.6	17.1	1.2
Cigarettes, past 30 days			
Consistent respondents	12.0	9.2	1.3
All respondents	10.6	9.5	1.1

[a]Estimates are for all respondents and for respondents whose reports across all questions on use were consistent.

this hypothesis by calculating weighted estimates of prevalence for young people whose responses across all drug use questions were consistent. These were compared with estimates of use for all respondents 12 to 17 years old; the results are presented in Table 7-11.

There are 12 comparisons in Table 7-11. In three cases, the pattern for consistent respondents is different from that for all respondents. For

TABLE 7-12 Prevalence of Reported Drug Use by Interviewer Experience in Prior NIDA-RTI Drug Surveys for Respondents Who Both Received and Completed the Self-Administered Versions of the Questionnaire

	Prior NIDA-RTI Experience	
Drug Use	Yes	No
Cocaine		
30-day Use	0.6	1.6
Past Year	2.1	4.1
Lifetime	11.4	17.8
Marijuana		
30-day Use	3.9	6.5
Past Year	7.3	11.9
Lifetime	38.0	43.4
Alcohol		
30-day Use	53.5	56.8
Past Year	71.7	77.0
Lifetime	85.3	86.9

NOTE. Cases were excluded from this analysis if interviewers assisted respondents in filling in the self-administered forms.

lifetime use of alcohol and past-year use of alcohol, consistent respondents had lower estimates of prevalence with the self-administered form than with the interviewer-administered form. For lifetime cocaine use, the consistent respondents had higher prevalence estimates using the self-administered questionnaire, whereas for all respondents, the interviewer-administered version produced larger estimates of prevalence. In addition, nonusers can more easily be consistent in their reports since they do not have to make choices as to whether or not particular episodes of use occurred within the various time periods.

Given this analysis, our finding that self-administered forms yield more reports of drug use does not appear to be due to a greater number of marking errors.

Interviewer Experience

Although most of the analytic results presented in earlier sections conform to reasonable expectations, one consistent and perplexing finding is that interviewers without experience in prior NIDA-RTI surveys obtained more frequent reports of drug use—particularly for the self-administered versions of the questionnaire. This finding persisted even after careful checking to ensure that the variable representing NIDA-RTI experience had not been coded backward by mistake.

Our model-fitting exercises show that the introduction of various controls for differences in the composition of the samples assigned to

the different groups of interviewers did not eliminate the result that was evident in the raw marginals. We have subsequently explored one additional possibility. In the self-administered version of the questionnaire, interviewers were permitted to complete the form when respondents were unable or unwilling to do so. More experienced interviewers might have been more likely to use this option, which would thereby turn a nominally self-administered questionnaire into an interviewer-administered one, with its attendant negative bias in reporting sensitive behaviors.

To test this hypothetical explanation, we retabulated prevalence rates from the self-administered versions of the questionnaire by NIDA-RTI experience. This analysis excluded the 360 cases in which interviewers reported assisting respondents in filling out the form. Table 7-12 shows the resultant prevalence estimates. Although one might ask for refinement of this analysis, there appears to be little basis for believing that this mechanism could account for the more frequent reports of drug use obtained by interviewers lacking prior NIDA-RTI experience.

REFERENCES

Bradburn, N., Sudman, S., and Associates (1979) *Improving Interview Method and Questionnaire Design*. San Francisco: Jossey-Bass.

Fay, R.E. (1982) Contingency table analysis for complex designs: CPLX. *Proceedings of the American Statistical Association (Survey Methods Section)*. 1982:44–53.

Fay, R.E. (1985) A jackknifed chi-square test for complex samples. *Journal of the American Statistical Association*. 389:148–157.

Fay, R.E. (1989) *CPLX, Contingency Table Analysis for Complex Sample Designs: Program Documentation*. Washington, D.C.: Statistical Methods Division, U.S. Bureau of the Census.

Goodman, L.A. (1967) The analysis of cross-classified data: Independence, quasi-independence, and interactions in contingency tables with or without missing entries. *Journal of the American Statistical Association*. 63:1091–1131.

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

Miller, H.G., Turner, C.F., and Moses, L.E., Eds. (1990) Chapter 6: Methodological Issues in AIDS Surveys. In *AIDS: The Second Decade*. Washington D.C.: National Academy Press.

NIDA (National Institute on Drug Abuse). (1991) *National Household Survey on Drug Abuse: Main Findings, 1990*. DHHS Pub. No. (ADM) 91–1788. Washington, DC: U.S. Government Printing Office.

Turner, C. F., and Martin, E. (1984) *Surveying Subjective Phenomena*. 2 vols. New York: Russell Sage.

U.S. Bureau of the Census (1982) *1980 Census of Population and Housing: Users' Guide*. Washington D.C.: Government Printing Office.