OFFICE OF APPLIED STUDIES

DEVELOPMENT OF COMPUTER-ASSISTED INTERVIEWING PROCEDURES FOR THE NATIONAL HOUSEHOLD SURVEY ON DRUG ABUSE (M-3)

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Substance Abuse and Mental Health Services Administration

Acknowledgments

This publication was developed for the Substance Abuse and Mental Health Services Administration (SAMHSA), Office of Applied Studies (OAS), by the Research Triangle Institute (RTI) under Contract No. 283-97-9007. Significant contributors at RTI include, in alphabetical order, Jeffrey Ancheta, Gayle S. Bieler, Brian J. Burke, Rachel A. Caspar, George H. Dunteman, Teresa P. Edwards, G.G. Frick, L. Lynn Guess, Brenda E. Gurley, Michael L. Hubbard, Donna M. Jewell, Larry A. Kroutil, Judith T. Lessler, Paul R. Moore, James O'Reilly, Patricia P. Parker, Michael A. Penne, Gilbert Rodriguez, Babu V. Shah, Richard S. Straw, Kelly Wayne, and Michael B. Witt. Peggy Barker at OAS also assisted in this report's preparation. Judith T. Lessler was the task leader for the production of the report, and Thomas G. Virag was RTI's project director for the NHSDA contract. Joseph C. Gfroerer, Director, Division of Population Surveys, OAS, was the project officer for the contract and reviewed the report.

Obtaining Additional Copies of Publication

Copies may be obtained, free of charge, from the National Clearinghouse for Alcohol and Drug Information (NCADI), a service of SAMHSA. Write or call NCADI at

National Clearinghouse for Alcohol and Drug Information
P.O. Box 2345
Rockville, MD 20847-2345
(301) 468-2600 1-800-729-6686 TDD 1-800-487-4889

Originating Office

SAMHSA OAS
5600 Fisher's Lane, Room 16-105
Rockville, MD 20857

January 2001

Table of Contents

Table of Contents (continued)

Chapter **Page**

Table of Contents (continued)

Chapter **Page**

List of Exhibits

List of Exhibits (continued)

Exhibit **Page**

List of Exhibits (continued)

List of Exhibits (continued)

List of Exhibits (continued)

Executive Summary

The National Household Survey on Drug Abuse (NHSDA) is the primary source of statistical information on the use of illegal drugs by persons in the United States. Sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA), the survey collects data by administering questionnaires to a representative sample of the population through face-to-face interviews at their place of residence.

The survey is designed to produce yearly cross-sectional estimates of substance use and abuse for the U.S. civilian, noninstitutionalized population aged 12 years or older. Prevalence estimates are produced of the use of alcohol, tobacco, marijuana and hashish, cocaine, heroin, inhalants, and hallucinogens, as well as of the nonmedical use of psychotherapeutic drugs. From the NHSDA's inception in 1971 until 1998, data were collected using paper-and-pencil interviewing (PAPI) methods with a combination of interviewer-administered questions and self-administered questionnaire (SAQs; i.e., respondent-completed answer sheets). In 1999, the NHSDA underwent a redesign that involved the implementation of computer-assisted interviewing (CAI) data collection methods. This report describes the development of this new methodology.

Background

Throughout the history of the NHSDA, there has been a continual focus on evaluating and improving the methodology of the survey. Methodological research and improvement has focused on content, sample design, questioning strategies, editing methods, and estimation procedures. The PAPI methodology had a number of limitations and problems that were identified and studied in this research. The advent of CAI procedures during the early 1990s, particularly audio computer-assisted self-interviewing (ACASI), clearly offered an opportunity for improving the measurement methods. Numerous studies have demonstrated that computer-assisted personal interviewing (CAPI) improves the quality of survey data and that ACASI results in increased reporting of sensitive issues.

The ACASI methodology allows the respondent to listen to questions through a headset and/or to read the questions on the computer screen. Respondents also key their own answers into the computer. Thus, greater privacy can be assured for the respondent even in interview settings that might not otherwise be considered sufficiently private. Programming the questionnaire allows for a more complex contingent questioning structure and strategy in a format where the routing is less visible to the respondent and makes possible the incorporation of consistency checks during the interview.

In addition to incorporating the use of CAI instruments for collecting data from respondents, the use of electronic screening of households was implemented in 1999. Prior to 1999, NHSDA interviewers used complex paper forms to conduct the 5-minute interview to determine the housing unit composition and to select the sample persons. This procedure was difficult to manage, prone to error, expensive to process, and limiting in terms of the sample selection algorithms that could be implemented.

In 1995, SAMHSA decided to initiate development and testing of CAI in the NHSDA. The development was accomplished primarily under a contract awarded to the Research Triangle Institute (RTI) in early 1996. SAMHSA carefully considered the shift to computer-assisted screening and interviewing, requiring extensive testing and proof of the feasibility of the new technology for the NHSDA. The testing protocol included a small initial feasibility experiment in the fourth quarter of 1996,

cognitive laboratory testing, a large (n = 1,982) field experiment in the fourth quarter of 1997, and a final pretest conducted in August 1998.

1996 CAI Feasibility Experiment

The 1996 CAI feasibility experiment was designed to assess the operational feasibility of using an electronic version of the NHSDA, the impact on perceptions of privacy, the length of the interview, the effect of CAI on the interviewing environment, and the quality of data provided. This study compared two versions of the CAI to the 1996 PAPI with a sample of 435 respondents in 20 purposively selected NHSDA primary sampling units (PSUs). The main results of this study were as follows:

1. CAPI reduced the time it took for interviewers to complete the personal interview component.

2. Respondents were, in general, able and willing to complete the extended ACASI interview with little help from an interviewer.

3. Interviewers were much less likely to know what the respondents' answers were when ACASI was used than when PAPI/SAQ was used.

4. ACASI appeared to increase reporting of past year and past month marijuana and cocaine use.

Cognitive Laboratory Testing

Following the CAI feasibility experiment, in the spring of 1997 several rounds of cognitive testing were conducted in RTI's Laboratory for Survey Methods and Measurement in order to further develop the ACASI methodology to be used in the 1997 field experiment. Fifty respondents were recruited. Three topics were explored in this lab testing: the voice used in the ACASI portion of the interview, a new method for asking the frequency of use question, and a method for resolving inconsistent responses during the interview.

A voice for the ACASI instrument was selected on the basis of lab respondents' ratings of eight voices that varied on several voice characteristics.

In the NHSDA prior to 1999, the question asking the number of days a respondent used a substance in the past 12 months had nine response categories that combined number of days with a periodicity estimate (e.g., " at least 12 but not more than 24 days [1-2 days a month])." This was particularly confusing for respondents who, for example, only drink alcoholic beverages every day for a 2-week vacation period each year (i.e., 14 is the appropriate number of days, but "1-2 days a month" is not appropriate). A new method was tested, in which the respondent selects the unit for reporting (i.e., days per month, days per week, or total days during the year). This procedure worked well and was incorporated into the 1997 field experiment.

A two-stage resolution methodology was developed for the field experiment. When the computer detected that a response was inconsistent with an earlier response, the respondent was first asked to **verify**

whether the second response was correct, and then to **resolve** the inconsistency by correcting one or both of the responses.

1997 Field Experiment

The 1997 field experiment, conducted during the fourth quarter of 1997, evaluated alternative versions of an ACASI NHSDA using a 2x2x2 factorial design with a sample of 1,982 respondents including 1,117 youths 12 to 17 years of age. The overall goal of the field experiment was to identify an ACASI version of the NHSDA that produced higher quality data without adversely affecting respondent burden or increasing breakoffs (i.e., incomplete interviews). Random halves of the sample were assigned to one of two levels within three experimental factors, described below.

1. **Factor 1: Structure of the contingent questioning in the CAI interview.** In the single gate question version, respondents were first asked if they had ever used a substance and were routed immediately to the next section if they had not. Under the multiple gate question version, every respondent answered three gate questions for each substance: use in the past 30 days, use in the past 12 months, and lifetime use. Only those respondents who answered "no" to each of the three questions were routed to the next section.

2. **Factor 2: Data quality checks within the ACASI interview.** For a random half of the respondents, the ACASI program included additional questions that followed up on inconsistent answers and questionable reports, such as a suspiciously low age of first use for a substance.

3. **Factor 3: Number of chances to report 30-day and 12-month use.** This factor was included at two levels: a single opportunity to report use and multiple opportunities to report use. Under the single opportunity to report use, respondents were only asked once about use during the past 30 days or during the past 12 months. Under the multiple opportunity version, respondents who indicated at least lifetime use of a substance were routed through additional follow-up questions even though they had not indicated use in the particular time period.

In addition to the experimental treatments, interviewer and respondent debriefing questions were included. A subsample (n = 3,105) of the Quarter 4 national NHSDA, which used a combination of interviewer-administered PAPI and SAQ, comprised the control group for the study. This comparison group was restricted to those 1997 NHSDA respondents who were in the same PSUs that contained the field experiment sample.

Effect of experimental factors on prevalence. Contrary to expectations, the single gate version rather than the multiple gate version resulted in increased reporting of drug use, particularly for the illicit substances (i.e., marijuana, cocaine, any illicit, and any illicit but marijuana). This was especially true for use during the past 30 days. In addition, any illicit

drug use showed higher prevalence ratios for all three reference periods in all three age categories, except for 12-month use of any illicit drug but marijuana for 12 to 17 year olds.

On the whole, when consistency checks were present, respondents gave somewhat higher reports of drug use across all drugs, for all reference periods, and all age groups. In addition, 12 to 17 year olds showed an overall tendency toward higher reporting in the version with a single opportunity to report use in a reference period. The analysis for the total sample and persons 18 years old or older did not show any pattern in favor of either treatment group.

ACASI versus PAPI. There was an overall tendency for ACASI to yield higher reports of drug use. This was especially evident among the 12 to 17 year olds where the differences were quite dramatic.

For the total sample, the difference between the ACASI and the PAPI rates of use was significant at the 0.1 level for cocaine and any illicit drug. For the 12 to 17 year olds, rates of use of every drug but cocaine showed some significant differences at the 0.05 level. The ACASI versus PAPI difference for rates of use of any illicit drug were significant at 0.05 for all three references periods.

Respondent's ease of answering questions. Respondents were asked to rate their ability to record their answers using their particular interview mode, ACASI versus PAPI, without the help of the field interviewer. Overall, a large percentage of both groups reported not needing the interviewer's help when entering answers; however, there was a tendency for a larger percentage of the ACASI respondents to indicate they needed no help (88.3% vs. 73.5%). This difference was even larger for the youths with 20% fewer of the youth ACASI respondents indicating that they required help from the interviewer. Adults with less than a high school education found the ACASI interview easier to complete on their own (83.2% vs. 68.1%).

Level of comfort in answering questions. Respondents also were asked about their level of comfort with the interviewing environment. Overall, ACASI respondents were more likely to report that they were comfortable (73.9% vs. 62.3%). Under both modes, youths were less comfortable than adults, but showed an increase of 15% in ACASI over PAPI. Additionally, about 65% of ACASI respondents who reported any illicit drug use in the past 30 days indicated that they felt comfortable using the computer. This compares to 59.6% of PAPI illicit drug users who said they felt comfortable using paper and pencil, indicating a preference for the ACASI interview mode among illicit drug users.

Recorded voice. The recorded voice was helpful to those respondents who could not read well. In one of the debriefing questions, respondents were asked to rate their own reading ability. Those who could not read well found the recorded voice more useful. Only 14.7% of the respondents with excellent reading ability felt that the voice helped them a lot, whereas 48.7% of the fair to poor readers indicated that it helped them a lot.

Privacy. ACASI was a major factor in increasing the respondents' perception that the interview was private. In all demographic groups, nearly twice as many of the ACASI

respondents reported that the interviewer saw none of their answers—overall 82.6% for ACASI versus 41.3% for PAPI. Nearly 40% of PAPI respondents indicated that the interviewer saw some of their answers, whereas only 13.1% of the ACASI respondents indicated such.

When asked which interview mode provided the best privacy protection, approximately 40% to 50% of the Quarter 4 PAPI comparison group favored the computer, 10% to 13% favored the answer sheet, and nearly 25% indicated that either method would protect equally. Furthermore, users of illicit substances were more likely to say that ACASI provided better privacy protection.

Results from the revised 12-month frequency-of-use items. For every drug except heroin (where sample sizes were quite small), CAI respondents reported greater frequency of use than did the PAPI respondents. With the exception of heroin, a greater percentage of CAI respondents provided an estimate of 12-month use that placed them in the upper half of the frequency scale (51 days or more) as compared to the PAPI respondents. This was especially true for youths, who consistently reported higher frequencies of use in CAI than in PAPI. This trend was not entirely consistent for adult respondents. The data indicate that the majority of respondents appeared to be responding as intended. Most of the respondents who selected the monthly or yearly reporting periods were infrequent users.

Effect of ACASI on mental health questions. Questions about adult mental health syndromes were included in the NHSDA instrument from 1994 to 1997. In the PAPI instrument, these questions were interviewer administered. In the 1997 field experiment, the adult mental health questions were included in the ACASI interview.

There was a definite trend for ACASI to yield higher estimates of these mental syndromes compared with PAPI. In particular, the rate of a likely major depressive episode based on ACASI (14.6%) was nearly double the rate for PAPI (7.4%). Similarly, the ACASI estimate for generalized anxiety disorder (5.8%) was nearly four times the PAPI rate (1.6%). The effects were particularly pronounced for males; the ACASI estimate for generalized anxiety disorder among males (5.1%) was more than 4.5 times the PAPI rate for females (1.1%). Similarly, ACASI estimated that 12.7% of adult males were probable cases for major depressive episode, or more than twice the PAPI rate of 5.5% for females.

Reporting of nonmedical use of psychotherapeutic drugs. There was a single version of the ACASI program for questions on the nonmedical use of the psychotherapeutic drugs, which include analgesics, tranquilizers, stimulants, and sedatives. For these drugs, there were significant differences in question structure between the CAI and PAPI, which resulted in higher reported rates of use in CAI.

Overall, reporting of lifetime nonmedical use of analgesics increased by 300%, from an estimated 4.9% of the population to 14.8%. Both youths and adults showed dramatic increases, with the youth lifetime prevalence rate being 3.7 times higher under ACASI and the adult rate 3 times higher. Use during the past 12 months of analgesics showed a similar dramatic increase. For the other psychotherapeutic drugs, similar but less dramatic results were obtained. For example, overall, the reported lifetime prevalence of nonmedical tranquilizers use was 2.7 times higher under ACASI, stimulant use was 1.8 times, higher and sedative use 2.5 times higher.

Ability to resolve inconsistencies. Approximately 28% of the respondents assigned to receive an interview that would require resolution of inconsistencies triggered at least one such check item.

Overall, the consistency resolution methodology was successful. The methodology improved the quality of the data collected without adversely affecting respondent cooperation or burden. Using this methodology in future implementations of the NHSDA will allow SAMHSA to capitalize on the numerous benefits of the ACASI technology while minimizing one of the potential pitfalls—that respondent errors and inconsistencies are not identified and corrected at the time of interview.

Final Pretest

The final pretest was conducted in August 1998 and consisted of a field test and concurrent cognitive laboratory interviews. Laboratory interviews were conducted with adolescents to test a new set of tobacco questions, including usual brand of each tobacco product, and a new question on "month of first use" for better incidence data on persons recently initiating substance use. Interviews were conducted with drug treatment clients to test updated "pill cards" and revised questions on nonmedical use of prescription drugs, as well as new questions to estimate *withdrawal symptoms* related to use of specific drugs. Primarily, the field test served as a final test of all procedures for the 1999 NHSDA, incorporating all of the best features from prior testing.

Problems identified in the field test with the hardware, software, and procedures were identified through interviewer debriefings and corrected as much as possible. Lab testing identified problems with question wording, and modifications to questions were made prior to fielding the 1999 survey.

Development of an Electronic Screener

The electronic screener was developed in three phases. It was first used in the 1997 NHSDA field experiment. It was revised and tested on a small scale in the spring of 1998. A final test was included in the August 1998 pretest.

Benefits of the electronic screener include the following:

1. elimination of interviewer errors in the selection process, including accidental errors and intentional tampering with the roster by the interviewer to achieve certain selections;

2. the capability to include more variables in the respondent selection algorithms;

3. reduction of data editing, data entry, and shipping costs; and

4. more detailed and timely information on field activities.

1. Introduction

The National Household Survey on Drug Abuse (NHSDA) is designed to yield early cross-sectional estimates of substance use and abuse for subpopulations defined by age and racial/ethnic status for the U.S. civilian, noninstitutionalized population aged 12 or older. The annual survey is sponsored by the Office of Applied Studies (OAS) within the Substance Abuse and Mental Health Services Administration (SAMHSA) and is currently conducted by the Research Triangle Institute (RTI). Prevalence estimates are produced of the use of alcohol, tobacco, marijuana and hashish, cocaine, heroin, inhalants, and hallucinogens, as well as of the nonmedical use of psychotherapeutic drugs. Prior to 1999, national samples were used to control the relative standard error of estimates for four subgroups: (a) the population as a whole; (b) four age groups: 12 to 17, 18 to 25, 26 to 34, and 35 or older; (c) three racial/ethnic groups: non-black/non-Hispanic, black/non-Hispanic, and Hispanic; [1] and (d) the cross-classification by age group and racial/ethnic group. In general, the sample was selected in four stages consisting of (a) primary sampling units (PSUs), (b) area segments within PSUs, (c) dwelling units within the area segments, and (d) persons within occupied dwelling units. To select the sample persons, interviewers visited each sample dwelling unit and screened the dwelling unit to determine the number of people in each of the age groups and the household's racial/ethnic status. Sample selection tables, printed on paper, were then consulted to determine who, if anyone, was selected for the survey. In most households, no one was selected; in the others, either one or two persons were selected.

From the inception of the NHSDA until 1999, the data were collected face to face using paper-and-pencil interviews (PAPIs) with a combination of interviewer-administered questions and respondent-completed answer sheets. Interviewers collected the data in the respondents' homes and attempted to obtain a private location for the interview. The respondents' anonymity and privacy of their responses were protected by separating the identifying information from their survey responses. Respondents were also assured that their identities and responses would be handled with the strictest confidence. The more sensitive questions were self-administered either by having individual respondents complete answer sheets on their own or by having them mark their answers as the interviewer read the questions aloud to them. Parental permission was required before interviewing a respondent under 18 years of age.

In 1999, the NHSDA sample was expanded and redesigned to permit using a combination of direct and model-based small area estimation procedures that allow SAMHSA to produce estimates for all 50 States and the District of Columbia (DC). In addition, computer-assisted data collection procedures were adopted for both screening and interviewing. This report summarizes the research to develop these computer-assisted screening and interviewing procedures. Information on the redesigned sample and details regarding the small area estimation techniques are not covered in this report, but may be found in other reports available from SAMHSA (2000).

[1] In the interest of readability for this report, "white" is used to indicate "non-black/non-Hispanic" and "black" to indicate "black/non-Hispanic."

The computer-assisted applications were developed for a number of reasons. First, during the screening, interviewers made both intentional and unintentional errors. It was possible to influence who was selected for the survey by manipulating the order in which respondents in the household were listed, and interviewers sometimes made mistakes when they consulted the sample selection tables. For example, in the 1997 survey, 75 cases could not be used because of incorrect sample selections. To prevent these errors and to permit more complex sampling algorithms for selecting sample persons, an electronic screening application was developed using a handheld Newton computer.[2] Second, interviewers and respondents made errors when completing the questionnaire, and the privacy afforded by the self-administered answer sheets was less available to those respondents who did not read well and had to depend upon the interviewer for assistance in completing the self-administered portions of the interview. To alleviate these difficulties, a computer-assisted interviewing (CAI) application was developed that employed both computer-assisted personal interviewing (CAPI) for sections of the interview that had traditionally been administered by interviewers, and audio computer-assisted self-interviewing (ACASI), for sensitive questions that had been asked using self-administered answer sheets.

In 1996, 1997, and 1998, SAMHSA sponsored a variety of research studies that explored the usefulness of using CAI procedures to reduce errors in the NHSDA. The research included a small-scale field experiment conducted in late 1996 that examined the operational feasibility of using CAI procedures for the NHSDA. This 1996 experiment examined perceptions of privacy, the length of the interview, the effect of CAI on the interviewing environment, and the quality of the data provided. Following the field experiment, a series of laboratory studies addressed methods for resolving inconsistent responses, ways for asking the questions on the frequency of use in the past 12 months, and reactions to different voices for ACASI interviews. In late 1997, a large-scale field experiment was conducted. The 1997 experiment examined the ACASI portion of the CAI interview by looking at (a) different ways of asking contingent questions, (b) the ability of respondents to resolve inconsistent responses, and (c) the effect of giving respondents multiple chances to report use of particular substances. In addition, the 1997 study examined the use of a handheld computer to conduct the household screening and sample selection activities. Finally, in the summer of 1998, a series of cognitive laboratory studies examined procedures for asking additional questions on tobacco use, and in August 1998, a field test of a revised electronic screening application and CAI procedures was conducted. This report summarizes the 1996 field experiment and describes the procedures and results of the other research.

This report covers a variety of NHSDA field experiment topics. To start, Chapter 2 gives a brief history of research on the NHSDA, and Chapter 3 offers further background information, including a literature review and an overview of critical design and operational issues. Chapter 4 focuses on the 1996 feasibility experiment and cognitive laboratory research, while Chapters 5 through 9 delve into the 1997 field experiment. Specifically, Chapter 5 summarizes the design and conduct of the 1997 effort, Chapter 6 compares CAPI/ACASI with PAPI for selected outcomes, Chapter 7 describes the effect of ACASI experimental factors on prevalence and data quality, Chapter 8 details the development and testing of an electronic screener, and Chapter 9 describes the operation of the 1997 field experiment. The next two chapters offer insights into the willingness of NHSDA respondents to be interviewed (Chapter 10) and the effect of NHSDA interviewers on data quality (Chapter 11). Chapter 12 is devoted to further refinement of the CAI procedures during the 1998 laboratory and field testing of a tobacco module. Please note that the exhibits, which are numbered according to the sections that they follow, are grouped immediately *after* the sections in which they are discussed.

[2]The Newton computer is an Apple product (Message Pad 2000). The U.S. Government does not endorse or favor any specific commercial product or company. Trade, proprietary, or company names appearing in this publication are used only because they are considered essential in the context of the studies reported herein.

2. History of Research on the NHSDA

Since 1971, the NHSDA has provided information on the extent of substance use and abuse in the United States. The first survey was conducted by the National Commission on Marihuana and Drug Abuse and established the basic methodology of using in-person interviews combined with self-administered answer sheets to increase the respondents' privacy and their willingness to respond. The self-administered answer sheets were also designed to avoid the use of skip patterns (i.e., contingent questioning structures used in computer-administered questionnaires in recent NHSDAs) so that interviewers would not be able to infer drug use on the part of the respondent based on the time it took to complete the answer sheets. In addition, the basic sampling strategy of targeting the survey toward younger people in order to improve the precision of the estimates was established in this first survey (Gfroerer, 1992).

From 1971 through 1998, 17 rounds of the NHSDA were conducted.[1] The survey became an annual, widely followed national benchmark beginning in 1990. The sample sizes increased over the years to provide accurate estimates of trends among minorities, youths, and other groups especially affected by drug abuse. From 1991 to 1993, supplemental samples were selected for six large metropolitan areas: Chicago, Denver, Los Angeles, Miami, New York City, and Washington, DC. Periodically, rural supplements were selected. Emerging issues, such as successful ballot initiatives in November 1996 to legalize medical uses of marijuana in California and Arizona, led to increases in sample sizes in these States in April 1997.

The core content includes questions on alcohol, tobacco, marijuana and hashish, cocaine, heroin, hallucinogens, inhalants, and the nonmedical use of prescription drugs. Prior to 1994, there were changes in both the number and wording of questions that addressed this content. To provide a more constant set of core questions, a new set of core answer sheets was adopted in 1994. The core answer sheets include (a) tobacco, (b) alcohol, (c) marijuana, (d) cocaine in any form, (e) crack cocaine, (f) heroin, (g) hallucinogens, (h) inhalants, and nonmedical use of (i) analgesics, (j) tranquilizers, (k) stimulants, and (l) sedatives. This revised core included fewer basic questions, revisions of wording so that questions were asked consistently across these answer sheets, rewording to enhance understanding, and a fixed order of administration. More information on changes that were made in 1994 can be found in SAMHSA (1996a). The core questions include age at first use, lifetime use, past year use, and past month use. In addition to the core, other topics have been covered on a periodic basis, and the substantive content has broadened over time. Topics have included mental health, need for drug treatment, perceptions of risk, driving behaviors, human immunodeficiency virus (HIV) risk behaviors, and others. These changes in content and scope have been made to meet the needs of researchers and policymakers for authoritative and timely information and analysis.

There has also been a continual focus on evaluating and improving the methodology of the survey by focusing on content, sample design, questioning strategies, editing methods, and estimation procedures. Methodological improvements (Gfroerer, 1992; SAMHSA, 1996a) have included a variety of efforts:

[1] The National Institute on Drug Abuse (NIDA) sponsored the NHSDA from 1974 to 1991; the survey series has been sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA) since October 1992.

1. an extensive review of the methodology by an NHSDA advisory committee in 1984 and a subsequent revision of the questioning methods for the 1985 survey;

2. a small-scale double-blind validation study in 1986 in which a sample of people from drug treatment facilities were interviewed using the NHSDA methodology (Harrell, 1985, 1997);

3. the adoption of machine editing in 1988 to provide detailed documentation of editing and improve the consistency of edits;[2]

4. increases in sample sizes and improvements in the sampling methodology;

5. a series of methodology studies that included a cognitive appraisal of the 1988 instrument (Forsyth, Lessler, & Hubbard, 1992), a study of missing and inconsistent data and nonresponse in the 1988 survey (Cox, Witt, Traccarella, & Perez-Michael, 1992; Witt, Pantula, Folsom, & Cox, 1992), a series of laboratory studies examining questioning techniques (Hubbard, 1992), a nonresponse follow-up study (Caspar, 1992), an experimental field test of alternative questionnaires conducted in 1990 (Hubbard, Pantula, & Lessler, 1992; Turner, Lessler, & Devore, 1992a; Turner, Lessler, George, Hubbard, & Witt, 1992b; Turner, Lessler, & Gfroerer, 1992c), and a U.S. Bureau of the Census match study that examined correlates of household screening nonresponse and person-level nonresponse (Gfroerer, Lessler, & Parsley, 1997a; Parsley, 1993);

6. an evaluation of the feasibility of using a telephone survey to collect drug use data (Gfroerer & Hughes, 1991, 1992);

7. a 1992 experiment examining the use of answer sheets with skip patterns (Lessler & Durante, 1992); and

8. pretests of revised questionnaires in 1992 and 1993 in preparation of the adoption of the revised questioning and editing strategies in 1994 (SAMHSA, 1996a).

In addition to the methodology studies that have been conducted in the context of the NHSDA, the measurement methodology for the survey has been evaluated by comparing it to results from other studies. For example, Gfroerer, Wright, and Kopstein (1997b) compared the NHSDA to Monitoring the Future (MTF), which is a school-based survey that produces

[2] One of the key privacy-enhancing features of the NHSDA was that each answer sheet was placed in an envelope after its completion. At the end of the entire interview, this envelope was then sealed and returned for processing. Thus, there was no field editing of the respondent-completed answer sheets; moreover, because of the promises of confidentiality and anonymity, respondents were not called back to resolve inconsistencies.

estimates of drug use by adolescents for many of the same substances as the NHSDA. The two surveys show similar trends, but the rates of drug use are higher in the MTF. Although many methodological differences were noted between the two surveys, the authors speculated that the higher rates of use in the MTF may be due to greater underreporting in the household setting used in the NHSDA.

Other studies have examined the impact that the mode of interview and the privacy of the interview setting had on drug use reporting. For example, Gfroerer and Hughes (1991, 1992) compared estimates from the NHSDA to those from a random-digit dialing (RDD) survey, and Aquilino (1994) compared three modes of interviewing for a group of respondents who were randomly assigned to complete either a telephone interview, a personal interview, or a self-administered questionnaire during a personal interview. Using data from the same experiment, Aquilino (1997) examined the effect of the presence of third parties on the reporting of illicit drug use. The results from the majority of these studies have supported the use of the basic methodology of the NHSDA. Household interviews, self-administered questionnaires, and a private setting all contributed to improved reporting.

In spite of the evidence supporting the basic methodology, comparisons with other surveys of drug use (Gfroerer et al., 1997b) indicated that improvements could be made. The advent of computer-assisted interviewing (CAI) procedures, particularly audio computer-assisted self-interviewing (ACASI), clearly offered an opportunity for enhancing the privacy of the interviews and improving the measurement methods. In reviewing the results of the 1990 NHSDA field experiment, Turner et al. (1992a) noted that although the research supported the use of self-administered questionnaires (SAQs), it does pose problems for respondents with poor reading skills. Respondents who cannot read well enough to complete the questionnaire on their own must depend on the interviewer to read the questions and answers to them while they mark their answers on answer sheets. Clearly, the interviewing environment is not as private for these respondents as it is for those who can read well, and ACASI permits them to answer in complete privacy. Numerous other studies had demonstrated that computer-assisted personal interviewing (CAPI) improves the quality of survey data and that ACASI results in increased reporting of sensitive behaviors (see Chapter 3). However, as discussed in Chapter 3, moving the NHSDA to use CAI procedures entailed examining and evaluating a number of issues.

3. Background: Literature Review and Research Issues

Survey administrators at the Substance Abuse and Mental Health Services Administration (SAMHSA), and previously at the National Institute on Drug Abuse (NIDA), have strived to continually improve and refine the NHSDA's methodology, while ensuring the comparability of trend data. In doing so, changes have been carefully evaluated to ensure that it is clearly understood how the changes affect trend data. Thus, major changes in data collection and estimation procedures have been instituted only after thorough testing, evaluation, and consultation. For the NHSDA, this is especially important because of its focus on highly sensitive topics.

3.1 Prior Research

Previous research has indicated that the validity of self-reported substance abuse data is highly dependent on the methods used to collect the data (Aquilino, 1994; Aquilino & LoSciuto, 1990; Duffy & Waterton, 1984; Schober, FeCaces, Pergamit, & Branden, 1992; Turner et al., 1992a). Moreover, research showing the feasibility of audio computer-assisted self-interviewing (ACASI) and its potential for improving the reporting of sensitive behaviors was pivotal in SAMHSA's decision to actually develop and test computer-assisted interviewing (CAI) procedures for use in the NHSDA (Duffer, Lessler, Weeks, & Mosher, 1996; O'Reilly, Hubbard, Lessler, Biemer, & Turner, 1994; Turner, Ku, Sonenstein, & Pleck, 1996). Successful implementation of CAI procedures in other surveys did not necessarily mean that CAI would work in the NHSDA, due to its unique data collection method and the scope of sensitive topics covered. Thus, extensive testing of a computer-assisted NHSDA and proof of its feasibility for the NHSDA was required.

A number of benefits may accrue from using CAI methods. Simply moving to a computer-assisted personal interviewing (CAPI) mode of data collection provides several benefits:

1. Routing can be controlled by the computer rather than by the interviewer, so fewer errors are made. As a result, more complex questionnaires than can be scripted on paper are possible.

2. Question items cannot be inadvertently skipped. An answer must be recorded for each question that appears on the screen.

3. Questions are presented to all respondents in the same order.

4. Out-of-range and inconsistent responses can be identified "on the spot" and corrected.

5. Data can be processed more quickly and analysis files created more easily.

6. Customized question wordings are handled easily, and "fills" are generated automatically by the system.

7. Numeric calculations can be performed by the computer software at the time of interview.

8. On-screen question-by-question specifications can be provided to the interviewers rather than requiring them to refer to hard-copy specifications.

9. Information about the respondent (either from sampling frame data or from previous interview data) can be preloaded and used to customize question wordings and routings.

10. Timing data can be collected more easily, even down to the individual question level.

Prior research suggests that the enhanced privacy under ACASI may increase the willingness of respondents to report sensitive behaviors. Other benefits that may accrue from using ACASI include (a) enhanced ability to interview respondents who speak languages other than English or Spanish; (b) improved standardization of the question presentation, thereby decreasing the interviewer component of measurement variance; and (c) increased ability to maintain the privacy of the interview for illiterate and semiliterate respondents.[1]

Prior to 1996, three large national surveys that included questions on sensitive topics had adopted CAI methods with ACASI components. In 1995, Cycle V of the National Survey of Family Growth (NSFG), sponsored by the National Center for Health Statistics (NCHS), was converted to CAPI with a short (approximately 8 minutes) ACASI module. The sample for the Cycle V NSFG was a list sample of women between the ages of 14 and 44 residing in households included in the 1993 National Health Interview Survey. The ACASI module included questions about abortions, number of sexual partners, HIV risk behaviors, forced sex, and family violence. Results from a pretest of approximately 800 women found higher reporting of abortions in ACASI than when the questions were administered by the interviewer during the CAPI portion of the interview (Lessler, Weeks, & O'Reilly, 1994). Results from the main study of approximately 10,800 women showed that 3% of unmarried women told the interviewer they had had four or more male sexual partners in the past 12 months compared with 9% reporting four or more partners in ACASI (Abma, Chandra, Mosher, Peterson, & Piccinino, 1997).

Operationally, the ACASI module worked well. A staff of 260 interviewers worked on the Cycle V NSFG. Interviewers reported positive reactions from respondents who completed the ACASI module. One negative aspect of the ACASI implementation was the use of an external audio box, which added additional weight for the interviewer to carry. Setting up an external voice box proved difficult for some of the NSFG interviewers. Kelly, Mosher, Duffer, and Kinsey (1997) speculated that the use of internal voice cards would reduce this operational burden associated with ACASI.

The National Survey of Adolescent Males (NSAM), sponsored by the National Institute of Child Health and Human Development (NICHD), incorporated an ACASI component. The study collects data on the sexual and HIV risk behaviors of young men in the United States. The NSAM-1 is a longitudinal study of young men who were aged 15 to 19 in 1988. These men were interviewed again in 1991. As part of the 1995 NSAM, the original sample members were interviewed a third time, and a second panel (NSAM-2) of young men aged 15 to 19 were also interviewed. In 1995, a methodological experiment was included that compared data on sensitive topics collected in a traditional, hard-copy self-administered questionnaire (SAQ) to data collected using ACASI. Comparing ACASI to the SAQ, Turner et al. (1998) reported that respondents were nearly four times as likely to report male to male sex when the ACASI methodology was used (5.5% vs. 1.5%). ACASI respondents also reported higher rates of other socially stigmatizing or illegal behaviors, including drug use, sexual contact with a prostitute, sexual activities while under the influence of drugs or alcohol, and a variety of violence measures.

[1] In the paper version of the NHSDA, these respondents typically must get help from the interviewers to complete the answer sheets, which compromises privacy.

Operationally, ACASI achieved the same positive review as in the NSFG. A total of 123 interviewers were trained for the study, and only a small number of computer problems were reported (Turner et al., 1996). As with the NSFG, however, the primary improvement suggested for ACASI was to replace the external audio boxes with internal sound cards.

The National Longitudinal Survey of Adolescent Health also included an ACASI component during an in-home interview (Bearman et al., 1997). Approximately 12,000 adolescents in grades 7 to 12 were selected from the school rosters obtained from the schools that participated in the in-school data collection. The interview was approximately 90 minutes long and included both CAPI and ACASI components. The ACASI component asked sensitive questions about contraception, romantic relationships, sexual activity outside of a relationship, motivations for birth control, and use of tobacco, alcohol, and drugs; the adolescents also completed a delinquency scale (Nicols et al., 1997). The final data collection report did not indicate that there were any problems with using ACASI to collect the sensitive information.

Although the potential benefits of CAI methods for the NHSDA seemed great, SAMHSA carefully considered its shift to the new technologies. Not only are the start-up costs large (i.e., development work and purchase of hardware), but the impact on trend data is also a major consideration. Trend data are particularly critical because the NHSDA is a key data source in tracking the effects of the changes in the welfare and health care systems on substance abuse and also in tracking the recent surge in illicit drug use among our Nation's youths.

3.2 Design and Operational Issues Addressed

A number of issues had to be successfully resolved before an automated NHSDA instrument could be considered a viable possibility.

3.2.1 Critical Design Issues

Work at Research Triangle Institute (RTI) concentrated on three issues that were critical to deciding on the CAI approach to the study, namely, interview length, use of contingent questioning, and resolution of inconsistent responses. Three key questions were asked:

1. Are respondents willing to take part in an extended ACASI interview?

2. Should skip patterns (i.e., contingent questioning) be included in the core sections of the NHSDA instrument?

3. Can we expect respondents to be able to resolve inconsistent answers in the ACASI portion of the interview?

For the NHSDA to successfully employ ACASI, respondents had to be willing to use a computer to complete the self-administered portion of the NHSDA interview. Depending on how much of the NHSDA instrument was programmed for ACASI administration and to what extent contingent questioning was employed, an ACASI interview could last 30 minutes or longer. As noted, previous studies had reported few respondent difficulties. However, these studies had significantly shorter ACASI sections. There was concern that respondents would be unwilling to take part in the interview, might break off before the interview was completed, or

might provide poor quality data as a result of having to complete so much of the interview using the computer.

Survey methodologists working on the survey were also concerned that the respondent would feel uncomfortable having the interviewer in his or her home with little, if anything, to do for such a large part of the interview. Similarly, there was a need to assess how interviewers reacted to having the interview out of their control for such a large amount of time because the typical interaction between respondent and interviewer would be significantly altered in the automated NHSDA. Thus, the research staff examined whether this change was likely to have any consequences for the quality of the data collected or for the types of individuals chosen to work as interviewers for the NHSDA in the future.

Having CAI procedures available for the self-administered portions of the interview offered the possibility of using contingent questioning in the interview. The NHSDA did not utilize contingent questioning in most of the core answer sheets. Skip patterns were incorporated, however, into the supplemental answer sheets. One reason that contingent questioning was not used in the core sections related to concerns about respondents' ability to accurately follow the routing. This was no longer a concern with an automated instrument. With proper programming and testing, computer programs will route the respondent to the next appropriate question based on the answers he or she provides.

There remained, however, unresolved issues as to how best to incorporate contingent questioning in the core sections of the NHSDA. In the paper NHSDA instrument, the practice of asking respondents every question in the core sections regardless of their prior answers provided multiple chances to assess use in certain time periods. Depending on how contingent questioning was incorporated, such multiple assessments would no longer be possible. But perhaps the most important issue to resolve was whether including contingent questioning in the core sections would alter the way that respondents answered the questions. For example, respondents might learn that answering "yes" to a question about their use of a drug results in additional questions, whereas answering "no" moves them to a new section of the instrument. To the extent that a respondent wants to complete the interview quickly, he or she might choose to answer in the negative even when that is not an accurate answer. Such response effects are well documented in the survey literature (see Dijkstra & van der Zouwen, 1982). With the special importance placed on using the NHSDA to monitor trends in drug use, response effects caused by incorporating additional contingent questioning needed to be scrutinized so that changes that affect trend data could be examined. At the same time, the researchers realized that the inclusion of contingent questioning would reduce the burden associated with the NHSDA for respondents who do not use drugs. This reduction in burden could be maintained, or new questions targeted toward nonusers could be incorporated into the NHSDA.

One further reason for excluding contingent questioning from the core sections of the NHSDA has been to protect the privacy of the respondent. Because every respondent had to answer every question on each of the core answer sheets, it was difficult for an interviewer to know whether or not a respondent was a user of the drug type. The use of contingent

questioning can cause nonusers to finish the NHSDA questionnaire faster than users.[2] Although interviewers are probably less able to identify users and nonusers in an ACASI methodology than they are under the answer sheet format, an assessment of whether respondents perceive the use of the contingent questioning to be a threat to privacy was examined during the conversion process.

Finally, the third critical issue that was addressed was to determine if consistency checks should be used in the ACASI portion of the automated NHSDA instrument. With sufficient time for programming, very complex consistency checks can be programmed into the computer-assisted NHSDA. However, it was not known if respondents could handle the process of resolving their own inconsistencies, particularly when the information being collected is especially sensitive. Because of the extensive inconsistent data in prior rounds (i.e., years) of the NHSDA, the methodologists felt that incorporating consistency checks would improve the quality of the data collected. However, it is possible that a respondent who is prompted once too often by the computer to resolve an inconsistency in his or her data will simply refuse to continue with the interview. At the same time, the quality of the NHSDA data would be improved if inconsistent answers could be resolved by the respondent rather than by an editing clerk after the fact.

The decision of how and to what extent to use consistency checks was closely tied to the decision regarding contingent questioning structures. The more extensive the use of these structures, the fewer the number of consistency checks that would be needed. However, if the automated NHSDA had only used contingent questioning structures as they appeared in the paper answer sheets, the total number of consistency checks could be quite large.

3.2.2 Operational Issues

A number of operational issues also needed to be resolved if an automated NHSDA was to be considered viable. Each of these issues is outlined briefly below.

Using showcards and pillcards. The paper NHSDA questionnaire included both showcards and pillcards. In the self-administered portion of the interview, showcards were used to display lists of drugs that the respondent should think about as he or she answered a particular set of questions. They were also used during the interviewer-administered portion of the interview for questions that have a large number of response categories, such as work status or income. Four pillcards were used during the self-administered portion of the interview to show the respondent pictures of prescription analgesics, tranquilizers, stimulants, and sedatives. Provisions for incorporating these showcards and pillcards were required, particularly in the portion of the NHSDA that was to be converted to an ACASI format.

Incorporating "other, specify" questions. A small number of questions in the NHSDA instrument require the respondent to write in an answer rather than simply mark a box

[2] Other factors, such as reading ability, ease of recall, and attention, can also introduce considerable variability in the time required to complete the interview.

or fill in a number. In most cases, these are "other, specify" questions where the respondent indicated an answer other than those listed and there is space designated for him or her to write in that other answer. Handling open-ended text within the ACASI portion of the NHSDA might be problematic if respondents are not familiar with a computer keyboard or have little experience with typing. This required investigation before the NHSDA could be converted to a CAPI/ACASI format.

Using long sets of response categories. A few questions in the self-administered portion of the NHSDA interview had especially large numbers of response categories. When these questions are programmed for ACASI, the categories barely fit on one screen, and the screen appears especially daunting to respondents. Similarly, there was concern that respondents relying on the recorded voice to help them select a category might be overwhelmed by the amount of information they must process. Alternatives to these long lists were needed. For example, several of the questions ask the respondent to report the number of days he or she used a particular drug during the past 12 months. For these questions, there were 11 response options. Alternatives, such as leaving the question open-ended and allowing the respondent to type in a number, or using some type of unfolding methodology, were likely to create response effects that needed to be well documented. Because the 12-month use items are important items in analysis, they required special care during the conversion to ensure that changes to these items were adequately tested and the effects well documented.

Automating the screener and other related forms. In addition to automating the NHSDA questionnaire, decisions were also required regarding whether to automate other components of the survey, such as the screener or verification form.

Adding additional questions to the NHSDA. A computer-assisted instrument can allow questions to be more carefully tailored to a particular respondent. SAMHSA has given special attention to whether new questions should be added to the NHSDA to develop a better understanding of specific subpopulations. In a PAPI/SAQ instrument, it is difficult to tailor questions to respondents who report specific drug usage patterns because it means re-asking a number of questions simply to screen the correct respondents into the series of questions. For example, asking questions of respondents who indicate using marijuana, cocaine, and heroin in the past 12 months and who also report several treatment episodes in the past 12 months requires four questions to be re-asked before the new series of questions can begin. Similarly, ACASI offers the opportunity to incorporate new questions targeted toward respondents who have never used drugs.

Training interviewers to conduct an automated NHSDA. Survey organizations are becoming increasingly skilled in training interviewers for computer-assisted data collections. Similarly, as CAI methods become more common, more and more interviewers have previous experience working with the technology. This can be both an asset and a liability. Interviewers who have worked on other CAPI/ACASI studies generally feel more comfortable working with a computer. However, to the extent that different studies use different software for programming, interviewers sometimes have to "unlearn" the procedures they used for the last study they worked on in order to learn the current procedures. This becomes especially problematic when interviewers are working on more than one computerized study concurrently and must keep track of which keystrokes are required to activate a given function for each application. During the

NHSDA conversion process, special attention was given to developing training protocols that ensured that all interviewers and supervisors working on the computer-assisted NHSDA were thoroughly trained. Wojcik, Bard, and Hunt (1991) addressed a number of these issues, including the need for sufficient electrical capacity, a sufficient number of telephone jacks (for transmission training), and secured storage for the computer equipment. Considering the large number of interviewers employed by the NHSDA (e.g., more than 300 interviewers worked on the 1998 NHSDA), these issues were not inconsequential for a full-scale computer-assisted NHSDA training session.

4. Initial Investigations: 1996 Feasibility Experiment and Cognitive Laboratory Research

The initial investigations of computer-assisted interviewing (CAI) methods included a 1996 feasibility experiment and a series of cognitive laboratory studies that investigated specific issues related to response under CAI. In this chapter, we present a summary of these preliminary studies. Additional information can be found elsewhere (Caspar & Edwards, 1997; Research Triangle Institute [RTI], 1997).

4.1 1996 Feasibility Experiment

The 1996 feasibility experiment focused on assessing the operational feasibility of using an electronic version of the NHSDA instrument. The study was conducted in the fall of 1996 and compared two CAI versions to the 1996 paper-and-pencil interview (PAPI) survey in 20 purposively selected primary sampling units (PSUs). Each CAI instrument contained a computer-assisted personal interviewing (CAPI) section that corresponded to the 1996 PAPI instrument and an audio computer-assisted self-interviewing (ACASI) section that corresponded to the 1996 answer sheets. The two versions tested differed in their approach to the ACASI component.

The NHSDA core answer sheets[1] began with a question on whether the respondent had ever used the substance, then questions on the characteristics of their use followed. In these core sections, even if the respondent had never used the substance, he or she was required to mark an answer for these subsequent questions on characteristics of use. Thus, for every question, there was a response choice labeled "I have never _____ in my life." The "answer every question strategy" was used to prevent false negative reports of substance use due to either privacy or response burden concerns on the part of the respondent and to eliminate response errors due to failure on the part of the respondent to follow the correct route through the questionnaire. Non-core answer sheets sometimes allowed the respondents to skip detailed questions that were not applicable given earlier responses.

The two CAI versions tested in 1996 differed in the structure of their ACASI components. In one version, the 1996 answer sheets were, more or less, exactly replicated and, for the core sections, respondents were not routed past detailed questions when they reported that they had not used the particular substance. In the other ACASI version, contingent questioning was used in the core answer sheets, and respondents who reported no use of a substance were routed past the detailed questions. The two ACASI versions were called MIRROR and SKIP. The CAPI components were the same.

This 1996 feasibility experiment was conducted in 20 of the 1996 NHSDA PSUs. Within these PSUs, additional housing units were selected and randomly assigned to receive either the CAPI/MIRROR or CAPI/SKIP version. All programming and testing of the CAI instruments was completed on an expedited basis; interviewers were trained in late October 1996, and all CAI interviews were completed in November 1996.[2] A total of 435 interviews were completed: 177 paper, 136 CAPI/MIRROR, and 122 CAPI/SKIP.

---

[1] The term "answer sheets" is used because the interview was constructed so that the respondent could mark answers in response to questions read by an interviewer. In most cases, however, the respondent was allowed to complete an answer sheet without the interviewer reading the questions.

[2] Because the PAPI cases were part of the regular NHSDA sample, interviewing of these cases continued through the end of 1996, which corresponds to the regular NHSDA Quarter 4 data collection period.

Additional information was collected using seven mechanisms, described as follows:

1. Several debriefing questions were completed by the interviewer immediately after the interview to gather information on the setting and privacy of the interview and the respondent's ability to complete the self-administered portions of the interview.

2. A brief discussion was held between the respondent and interviewer on the privacy, difficulty, and interest in the interview. These discussions were recorded, transcribed, and coded.

3. An analysis was made of a keystroke file that had recorded all keystrokes made by either interviewers or respondents during the interview.

4. Time stamps were subsequently analyzed to determine the time required to complete sections of the interview.

5. An observation form was completed by the interviewer while the respondent was working on the ACASI component of the interview.

6. Problem logs were maintained by in-house project staff during the data collection.

7. Interviewer debriefing calls were made during which operations and potential improvements were discussed.

Overall, the 1996 feasibility experiment demonstrated that a CAI approach to collection of NHSDA data was workable:

1. Respondents were able and willing to complete an extended ACASI interview.

2. Large decreases in response rates were unlikely.

3. CAPI reduces the time it takes for interviewers to complete the personal interview component.

4. The SKIP version was some 10 minutes shorter than the MIRROR version due to respondents answering fewer questions.

5. About 14% of all ACASI respondents asked the interviewer to explain a question, and about 25% asked about how to use the computer during the interview. Respondents were less likely to ask for help when using the SKIP version.

6. Relative to the answer sheets, there were large differences in the degree to which the interviewer could glean the respondents' answers to the self-administered sections under the two ACASI versions. This indicated that the ACASI administration was more private, even though the privacy of the overall setting was similar.

7. ACASI appeared to increase reporting of past year and past month marijuana and cocaine use.

8. Very few respondents gave a pattern of response that indicated that they were either unwilling or unable to complete the response task once they had begun.

Based on these positive results, several laboratory experiments were conducted in 1997 to examine alternative questioning strategies under a CAI version of the NHSDA.

4.2 Cognitive Laboratory Testing

4.2.1 Introduction

To refine the CAI NHSDA instrument prior to the 1997 field experiment, cognitive laboratory testing was conducted between April 22, 1997, and June 10, 1997. A total of 50 subjects were recruited and interviewed in RTI's Laboratory for Survey Methods and Measurement. The majority of the testing covered three specific areas: (a) the voice used for the ACASI portion of the interview, (b) a new method for collecting the 12-month frequency of use data, and (c) procedures that allow the respondent to resolve inconsistencies in his or her data at the time of interview. In addition, we tested the implementation of a "multiple use" treatment under which respondents received multiple opportunities to report use of a particular substance within the same time period.

4.2.2 Subject Recruitment and Demographics

Subjects were recruited for laboratory testing through a number of sources, including flyers placed around the Raleigh/Durham/Chapel Hill, North Carolina, area; advertisements placed in newsletters; and word of mouth. For those subjects who were included in the test of multiple use questions, we required them to have drunk alcohol during the past 12 months. Most of the subjects were interviewed at Research Triangle Institute (RTI); however, in a few cases, interviews were scheduled for other locations when participants were unable to travel to RTI. Participants were told that the interview would last no longer than 2 hours. Each participant was paid $35 for his or her time. All the laboratory participants were 18 years of age or older. Youths were not included in this round of testing because the need for parental permission extends the time it takes to conduct laboratory testing.

4.2.3 Inconsistency Resolution Procedures

One of the most significant potential benefits of converting the NHSDA to a computer-assisted format is the chance to resolve inconsistent data at the time of the interview. However, achieving the privacy benefits of the ACASI component of the interview requires that the respondent be able to resolve inconsistencies for many items on his or her own. Thus, one of the goals of the laboratory testing was to develop a method for resolving inconsistent data that the respondent could easily understand and complete without significant intervention by the interviewer.

We developed a resolution methodology that combines two components. First, respondents were asked to **verify** that an answer they entered was, in fact, correct. So, for example, when a 20-year-old respondent indicated that she was 51 the first time she drank alcohol (a clearly inconsistent answer), the computer was programmed to verify that this information was correct. This step was included to help eliminate inconsistencies that may be due to keying errors. If the respondent indicated that this information was incorrect, she was then routed back to answer the question again (perhaps this time entering her age when she took her first drink as "15"). A second component incorporates the **resolution** of inconsistent answers. For example, a respondent who indicated drinking alcohol on 15 days in the past 12 months, but then reported drinking alcohol on 25 days in the past 30 days would first be asked to verify the last entry keyed. If he indicated that the entry was correct, then he was routed to a question that identified the inconsistency and was provided with an opportunity to fix the incorrect entry.

In developing the actual text of these verification and resolution screens, we sought to incorporate several important features:

1. The original responses that were inconsistent were provided to the respondent to enhance recall and comprehension.

2. The resolution screens were worded so as to not explicitly place the responsibility for the inconsistency on the respondent (e.g., "your answers"), but rather to imply that the computer may be incorrect (e.g., "the computer recorded").

3. Respondents were asked to identify the incorrect response when two items were inconsistent to facilitate the flow of questioning.

4. When a respondent indicated that both answers were incorrect, he or she was routed back to the two items in the same order they were presented the first time in order to maintain a consistent flow through the instrument.

5. Respondents were explicitly notified when their answers were inconsistent. This was done rather than attempting to resolve the inconsistency without actually making the respondents aware of the problem.

The results indicated that in at least one case we were incorrect in our assumptions about how best to structure the verification and resolution tasks.

To maximize the efficiency of our laboratory testing, we felt we could not rely on respondents to give inconsistent answers in the laboratory testing. Even with an extremely large sample, the number of respondents who will provide inconsistent answers is quite small. Because only 40 respondents were recruited for this task, it seemed entirely possible that none of the respondents who came into the laboratory would provide an inconsistent response. For this reason, we developed a laboratory task that incorporated the use of vignettes. The vignettes we used were essentially brief descriptions of a person and his or her drinking behavior. The laboratory subject was instructed to answer a series of questions about drinking alcohol as though

he or she was the person in the vignette. The subject read the vignette, then began answering the questions as they appeared on the computer screen. At a specific point in the questioning process, the subject was instructed to obtain additional information from the interviewer to be able to continue answering the questions. This additional information resulted in the subject providing inconsistent answers, which he or she was then required to verify or resolve.

The vignette methodology was somewhat artificial. That is, subjects did not answer the questions based on their own experiences. Thus, we did not learn anything about why people may give inconsistent answers (e.g., poorly worded questions, difficulty in recalling the information, desire to conceal information). However, our primary objective in this testing was to determine whether respondents could easily navigate through the verification and resolution process without becoming either confused or annoyed.

Our initial round of testing indicated one very pronounced problem. Respondents found it very confusing to be asked which of their answers was **incorrect** when two answers were identified as inconsistent. Respondents reported that it was much more logical to be asked which answer was **correct**. Even when the researcher pointed out that the reason for the question was to determine which question to re-ask, the subjects were nearly unanimous in their preference for identifying the correct answer.

Our initial testing also pointed to some problems with the vignette task itself. A number of respondents had difficulties figuring out which information in the vignette was applicable to which survey question. In our effort to make the vignettes seem more "realistic," we added information not specifically needed to complete the set of questions asked. To reduce the confusion caused by the vignettes, we scaled back the amount of information provided in each vignette to only that which was needed to complete the task.

Using the revised vignettes and the resolution process that asked the respondent to indicate which of his or her answers was correct, we began testing a second round of subjects. The second round of subjects seemed to find the resolution task much easier than subjects in the first round had found it. Subjects were able to easily select which of the answers was correct during the resolution process and understood that they were being routed back to the incorrect item to make the necessary correction. In general, respondents were not put off by the verification and resolution process. Some respondents went so far as to note that they would appreciate the computer pointing out inconsistencies in their data. Although not unanimous, most respondents preferred the less direct wording of "the computer recorded..." to the more direct wording that would say, "you reported that..." Respondents with little computer experience seemed to prefer this wording because they believed that the computer could make mistakes in how entries were stored. Respondents with greater computer literacy recognized that recording errors are made by the respondent and not the computer. However, the majority of these respondents still felt the less direct wording would be less confrontational and less likely to embarrass the respondent.

Verification screens were reported to be easier to complete than the resolution questions. Comments made by the subjects indicated that the verification screens were short and to the point. Resolution screens were reported to be "too wordy." Respondents indicated that there

was so much text to read that it was easy to get confused. Therefore, for our second round of testing, we reduced the amount of text in the resolution questions. For example, in round two the resolution screen that identified inconsistencies between 12-month frequency and 30-day frequency was worded as follows:

> The computer compared the answers for the last question and an earlier question. According to the answers it recorded, you drank one or more alcoholic beverages on more days in the past 30 days than in the past 12 months. This is not possible. Which of the following is correct?
>
> > I drank alcohol on [XX] days in the past 12 months
> > I drank alcohol on [YY] days in the past 30 days
> > Neither answer is correct

Subjects noted that the scripting in the body of the question was repeated again in the categories. This redundancy was viewed as unnecessary and sometimes confusing.

Prior to conducting the third (and last) round of laboratory interviews, we made some additional changes to the resolution screens to reduce the amount of text. Scripting was reduced again, and the revised version of the question shown above was worded as follows:

> The answers for the last question and an earlier question disagree. Which answer is correct?
>
> > I drank alcohol on [XX] days in the past 12 months
> > I drank alcohol on [YY] days in the past 30 days
> > Neither answer is correct

The revised wording seemed to work well in the laboratory. Subjects were still able to complete the resolution task with little trouble, and we did not have as many complaints about the wordiness of the screens.

Based on these three rounds of testing, we developed a method for resolving inconsistent responses in the NHSDA that was tested on respondents in the 1997 field experiment.

4.2.4 12-Month Frequency of Use

Two problems with the 12-month frequency of use item have been noted over the years. First, the question is difficult to answer because it requires the respondent to recall information over a long period of time. Second, the answer categories combine total number or days with a periodicity estimate, which confuses respondents who have "episodic" use patterns. The response categories were as follows:

> More than 300 days (every day or almost every day),
> At least 201 but not more than 300 days (5 to 6 days a week),
> At least 101 but not more than 200 days (3 to 4 days a week),

At least 51 but not more than 100 days (1 to 2 days a week),
At least 25 but not more than 50 days (3 to 4 days a month),
At least 12 but not more than 24 days (1 to 2 days a month),
At least 6 but not more than 11 days (less than 1 day a month),
At least 3 but not more than 5 days in the past 12 months, and
At least 1 but not more than 2 days in the past 12 months.

For example, a respondent who only drinks alcohol every day during a 2-week vacation each year may have difficulty choosing a category because the number of days she drank alcohol (14) falls in a category that is also identified as "1 to 2 days a month."

To begin to understand how people work with these categories, we developed a laboratory protocol that split the two parts of the response categories. Respondents were first asked to report the number of days they drank alcohol during the past 12 months using a showcard that displayed only the text shown in parentheses (i.e., "every day or almost every day," "5 to 6 days a week," etc.). Based on the category they selected, the interviewer provided them with a numerical estimate of their 12-month frequency. In each case, the estimate given was the numeric range that corresponds to the text in parentheses. Subjects were asked whether the numeric estimate seemed right for them and if not, whether the actual frequency was higher or lower than the estimate. We also developed a parallel set of items that asked about eating candy. These items were developed for use with respondents who had not drunk alcohol during the past 12 months. Initially, we used these items only with nondrinkers, but about halfway through the testing period, we began taking every subject through both sets of materials in order to maximize the amount of information we could collect about how people work with the set of categories.

Not surprisingly, respondents had difficulty recalling their alcohol use over the past 12 months. Although having categories to choose from made the task easier, a number of subjects still reported estimating their use based on their use over just the past few months. Respondents with more sporadic use patterns had more difficulties reporting their answer than did respondents with regular use patterns. Generally, the difficulty was caused by the implicit regularity of the response alternatives. Subjects who used the substance infrequently noted that the categories were difficult to work with because they implied that the use occurred on a regular basis. One subject volunteered that although he could answer the questions for alcohol, his sporadic use of other drugs (marijuana and inhalants) would make it difficult for him to report his use accurately.

However, among respondents who were able to select a category (about 70% of the respondents), the estimate provided by the interviewer was reported to be correct. Respondents noted that because there was such a broad range to the categories, it was possible to allow for less regular use without needing to switch categories. In only four cases did the subjects indicate that their actual use was lower or higher than the estimate provided. In these cases, respondents were not changing categories based on an exact count, but rather based on "gut reactions" that led them to believe that their use was outside the range provided. In some cases, respondents seemed to agree with the estimate provided by the interviewer simply because the interviewer was considered to be the authority. Several respondents made such comments as, "if that's how it multiplies out, then it must be right" or "if you say so." Some respondents checked the

interviewer's math themselves by multiplying the rate of use by the number of months or weeks in a year.

Based on this laboratory testing, it seemed that the biggest problem respondents had with the 12-month frequency of use question was dealing with the inherent periodicity contained in the categories. The question was particularly difficult for respondents with sporadic or infrequent use. In an attempt to facilitate the response task, we developed a revised 12-month frequency of use question series. This new series includes a question that allows the respondent to select the unit for reporting his or her use. Respondents can choose to report the number of days per week, the days per month, or the total number of days they used during the past 12 months. The follow-up question that the respondent receives is based on his or her choice of units.

We interviewed 10 subjects using this protocol and asked them to report any confusion or difficulties they experienced with any of the questions. With only 10 subjects, we must use caution in interpreting our results, but the general response was favorable. All but one of the subjects readily understood why the units question was being asked. Most of the subjects noted that the inclusion of the units question made the 12-month frequency question significantly easier to answer. Two respondents chose to report their use in total number of days, four chose to report the average number of days per month, and four chose to report the average number of days per week. When asked why they selected the category they did, respondents uniformly reported they had chosen their category based on the frequency of their use. The two respondents who chose to report total number of days both indicated that they drank only on special occasions (New Year's Eve, weddings, anniversaries, etc.) and thus it was easiest to count up the special occasions that had occurred during the past 12 months. Respondents who chose to report the average number of days per month did so because there were too many days for them to count up total days, but they did not drink frequently enough to report weekly. Respondents who chose to report the average number of days per week did so because they drink more often and more regularly than monthly.

Particularly for the respondents who chose to report the average number of days per month, we were concerned that their answer to the 12-month frequency question would affect how they answered the past 30-day frequency question. This proved not to be a problem, however. Respondents did not feel that their past 30-day frequency of use needed to parallel their answer for the 12-month frequency of use. The respondents commented that the 12-month frequency question was asking for an **average** number of days per month during the past 12 months, while the 30-day frequency question was asking about a specific 30-day period. Thus, respondents did not feel obligated to provide the same answer to the 30-day question that they had just provided to the 12-month question. Their responses were similar, however—usually within 1 to 3 days of their 12-month answer.

In short, the new 12-month series of questions seemed to work well. Respondents had a few comments for ways to improve the wording of the response categories to make them more clear, but overall this methodology seemed feasible and was used for the 1997 field experiment.

4.2.5 Selecting a Voice for ACASI

Originally, we had planned to examine respondent reactions to male and female voices in the 1997 field experiment by using a within-subjects design in which respondents heard some of the questions read in a female voice and some read in a male voice. We decided to drop this treatment from the experiment because there was not time at the end of the development period to record and program the entire interview in two voices. However, a cognitive laboratory study examined respondents' reactions to the different voices. Respondents listened to four pairs of voices (male and female in each pair) and were asked to indicate which pair they preferred. Following this, they listened to each voice and rated it on several voice characteristics. Respondents were able to reliably choose a preferred voice. There was some indication that respondents rated the preferred female voice as being deeper and slower and as having fewer changes in loudness than the other female voices. Among the male voices, there was little difference in the respondents' ratings of voice characteristics for the preferred voice relative to the others, although the two most preferred male voices were ranked as having higher pitch (see Caspar & Edwards, 1997, for detailed results of this study).

5. Summary of the Design and Conduct of the 1997 Field Experiment

The design of the 1997 field experiment was based on the results of the 1996 feasibility experiment, laboratory testing, power calculations, and discussions as to the operational feasibility of various designs for the field experiment. The experiment compared alternative versions of the computer-assisted interviewing (CAI) interview in a factorial design. These alternatives were compared to each other and to the results from the paper-and-pencil interview (PAPI) and self-administered answer sheets. To reduce the overall cost of the field experiment, we conducted the experiment in the fourth quarter of 1997 and used the Quarter 4 1997 NHSDA survey results as a comparison group.

In this chapter, we review the overall design of the CAI experimental treatments, sampling design, development of the electronic screener, screening and interview response rates, and analysis weights.

5.1 Experimental Design for Computer-Assisted Interviewing

We considered examining experimentally a number of factors, including the use of alternative skip patterns (i.e., contingent questioning structures), respondent-completed consistency and range checks, variations in the procedures of administering the 12-month frequency questions, audio computer-assisted self-interviewing (ACASI) versus computer-assisted personal interviewing (CAPI) administration of mental health questions, question wordings that were tailored or not tailored to the CAI environment, extra questions for nonusers of particular substances, and the effects of alternative voices on respondent preferences and responses. Other factors considered were examination of differences in responses under a "mark all that apply" questioning strategy versus the current "multiple question" strategy, asking income questions using ACASI, translating the alternative CAI instruments into Spanish, and specifically designing multiple questions on use to permit response variance modeling. Based on the cost and complexity of conducting such a study and the need to increase the power of the experimental comparisons, the number of experimental factors was reduced to three.

5.1.1 Structure of Alternative CAI Instruments

Several types of changes were made in the CAI instruments for all respondents participating in the 1997 field experiment:

1. The mental health items were administered via ACASI in all versions of the CAI questionnaire. In the 1997 NHSDA, these questions were administered by the interviewers using a PAPI interview. These items were moved to the ACASI section because other research has shown that self-administration of these items is likely to be easy and to improve response. Also, it was felt that it was not necessary to examine this change experimentally because these questions are not part of the core interview.
2. Question and response category wordings were tailored for CAI across all versions of the CAI questionnaire. A workable CAI version requires that the wording be tailored to the CAI environment; not doing so would create an awkward CAI interview.

In addition, the following decisions were made:

1. No additional questions would be added for nonusers. We considered including additional questions for nonusers in the core sections to prevent respondents from giving false negative answers to initial questions on use of a particular substance in order to either (a) avoid having to answer additional questions or (b) because they felt that answering a number of questions about a substance compromised their privacy. The latter could occur because the respondent reasoned as follows: "If the interviewer or an observer notices that I am taking a long time to answer questions, the interviewer or observer will conclude that I have used an illegal substance." We decided not to include additional questions for nonusers because (a) the CAI environment presents questions one by one, making it hard for respondents to know in advance that they will have to answer additional questions if they respond "yes" to the overall use question; (b) there was already considerable variability across answer sheets in the number of questions that follow a "yes" response; and (c) one of the contingent questioning structures that was to be tested required nonusers to answer several questions for each substance before they were routed to the next substance.
2. There was no experimental treatment for the 12-month frequency of use question. An alternative way of asking the 12-month frequency question was needed for the ACASI administration. An unfolding strategy was used in the fall 1996 feasibility experiment. Based on laboratory testing, we decided to have all respondents report 12-month frequency using the two-stage process in which they first indicated the metric that would be easiest for them (days per year, days per month, or days per week) and then reported the number of days for that period.

3. There was no Spanish translation. No Spanish translation was used because only about 150 Spanish-speaking respondents were likely to be included, and we felt that it was not cost effective to translate eight different versions for so few respondents.
4. The income questions were retained in the interviewer-administered portion of the CAI instrument across all versions. Income is a sensitive issue and could possibly benefit from being asked in ACASI; however, because the sample person is not always the person who knows the income and because of the complexity of the question, we decided to leave these questions as interviewer-administered.

Experimental features. We chose to examine questioning strategies that had the potential for reducing respondent burden and improving accuracy using a 2x2x2 factorial design. Random halves of the sample were assigned one of two levels within three experimental factors. These factors are described in the following paragraphs.

Factor 1: Structure of the contingent questioning in the CAI interview. Under a contingent questioning strategy, respondents are routed past detailed questions if they indicate they have not used the substance in earlier questions. The 1996 feasibility experiment indicated that the contingent questioning was likely to yield prevalence estimates that were the same or higher than those obtained using answer sheets and the "answer every question" strategy. Thus, two versions were tested in 1997: use of a single gate question and use of multiple gate questions. In the single gate question version, respondents were first asked if they had ever used a substance and were immediately routed to the next section if they had not. Respondents who reported some use were routed through the appropriate follow-up questions. Under the multiple gate question version, every respondent answered three gate questions for each substance: use in the past 30 days, use in the past 12 months, and lifetime use. Only those respondents who answered "no" to each of the three questions were routed to the next section, and those respondents answering "yes" to any of the three gate questions were routed to the appropriate follow-up questions.

Rationale: We examined whether having the respondent answer three gate questions rather than one reduced threats to complete reporting for the following reasons: The single gate question was tested in the fall of 1996 and resulted in a significantly shorter interview and increased reporting. However, allowing respondents only one opportunity to report use might result in decreased prevalence estimates because (a) the respondents mistakenly answer "no" to the sole gate question, (b) they can reduce the number of questions they answer by misreporting actual use to a single question, or (c) they may feel their privacy is compromised by taking a long time to answer questions.

Factor 2: Data quality checks within the ACASI interview. We examined the potential for improving data quality by having a random half of the respondents resolve inconsistent and questionable data during the interview. For a random half of the respondents, the ACASI program included additional questions that followed up on inconsistent answers and questionable reports, such as a suspiciously low age of first use for a substance. Consistency checks were included for the following inconsistent reports:

1. The 30-day frequency of use was greater than the 12-month frequency of use for cigarettes, alcohol, marijuana, cocaine, crack, heroin, hallucinogens, or inhalants.
2. A response of zero days used in the past 30 days was reported by persons previously reporting some use within the past 30 days for cigarettes, alcohol, marijuana, cocaine, crack, heroin, hallucinogens, or inhalants.
3. Age at first use was suspiciously low for cigarettes, alcohol, marijuana, cocaine, crack, heroin, hallucinogens, or inhalants.
4. Age at first use was greater than or equal to current age: cigarettes, alcohol, marijuana, cocaine, crack, heroin, hallucinogens, or inhalants.
5. The 12-month frequency of being very high or drunk was greater than the 12-month frequency of use for alcohol.
6. The number of days consumed five or more drinks on the same occasion was greater than the 30-day frequency of use for alcohol.
7. The last use of LSD was more recent than the last use of any hallucinogen.
8. The last use of phencyclidine (PCP) was more recent than the last use of any hallucinogen.

Rationale: Clearly, it is preferable to have respondents correct any inconsistencies in their data rather than having an analyst determine how to edit the data after the fact. In addition, although considerable effort must be expended to program these data quality checks, they have the potential to reduce the post-survey processing by reducing the number of edits. However, we were uncertain as to whether respondents would be able or willing to provide this type of information and speculated that it could increase either the number of breakoffs and or the overall length of the interview.

Factor 3: Number of chances to report 30-day and 12-month use. This factor was included at two levels: a single opportunity to report use and multiple opportunities to report use. Under the single opportunity to report use, regardless of the version of contingent questioning used, respondents were only asked once to indicate use during the past 30 days or during the past 12

months.[9] With the multiple opportunities, respondents who indicated at least lifetime use of a substance were routed through the additional follow-up questions even though they had not indicated use in the particular time period. For example, respondents who reported that their last use was more than 30 days ago were asked to report the number of days they had used a substance in the past 30 days in spite of this report. Similarly, respondents who reported that their most recent use was more than 12 months ago but within the past 3 years were routed to the question on frequency of use in the past 12 months. In addition, respondents who reported no cocaine use were asked about crack in spite of their denial of using any form of cocaine.

Rationale: With the paper answer sheets, respondents had many "opportunities" to indicate use beyond the basic lifetime, 12-month, and 30-day questions because there was no contingent questioning that routes them around questions that did not apply to them. When answering these other questions, respondents sometimes were inconsistent and indicated that they may in fact be a user of the substance. A significant number of respondents were classified as more frequent users based on editing rules. With only one opportunity for respondents to report 12-month or 30-day use, we thought that we might see a decline in prevalence rates. By adding a second question for these items, we could determine how the prevalence rates were likely to be affected by eliminating multiple questions on use during the 30-day and 12-month reporting periods.

While deciding whether it was necessary to include this factor, we examined the results from the 1996 NHSDA to determine how often respondents gave inconsistent answers on use within the core answer sheets. In the 1996 NHSDA, only 84% of the edited past month alcohol users indicated that they had used within the past month when they were responding to the recency question. The corresponding percentages for marijuana and cocaine are 77% and 58.7%, respectively.

In addition to the experimental factors described above, we included respondent and interviewer debriefing questions. The respondent debriefing questions gathered information on respondents' computer knowledge, attitudes and preferences, and perceptions of privacy and confidentiality. The interviewer debriefing questions consisted of a short set of questions for the interviewer on his or her impression of the interview. These focused on questions raised by the respondent, problems encountered, possible reasons for consistency checks being tripped, appraisal of the respondent's interest in and understanding of the interview, and so on.

5.1.2 Assignment to Treatments

The sample was designed to yield a total of 2,256 respondents. Exhibit 5.1.1 presents the expected number of respondents in each treatment combination.

The goal of the design was to yield 1,128 respondents for each of the major factors (main effects) in the experiment. In addition, because of the pressing need to understand substance abuse among youths, the sample was designed so that half of the respondents were expected to be 12 to 17 year olds. The CAI application included a case management system (CMS) that randomly assigned each person who agreed to be interviewed to one of the eight versions of the questionnaire.

5.1.3 Comparison Group

To reduce costs, we decided to use the 1997 NHSDA Quarter 4 respondents as the control group. This comparison group was restricted to those 1997 NHSDA respondents who were in the same primary sampling units (PSUs) that contained the 1997 field experiment sample. There were both positive and negative aspects of doing so. On the positive side, it provided a large sample size for the comparisons from data that were already being collected. A large sample size had the potential for both increasing the power of the comparisons and reducing the overall cost of the field experiment. Having interviewers conduct both the CAI and the PAPI versions would have required training the field staff to use two data collection methods. Thus, by selecting a subsample of 1997 NHSDA Quarter 4 respondents as the comparison group, we avoided the costs of this extra training. On the negative side, we felt that comparisons of overall response rates at the sample person level would be confounded by the fact that there would be two different interviewing teams collecting the data. We would be limited in our ability to disentangle any observed differences in response rates between the Quarter 4 NHSDA and the 1997 field experiment and to determine if these were due to the CAI interviewing, interviewer experience, or interviewing teams. On balance, it was felt that it was more important to have a large sample size for comparing the alternative questioning strategies rather than to focus on the response rate. The survey staff agreed that if CAI were adopted for the survey, they would be able to find procedures to achieve equivalent response rates.

It turned out that we achieved lower response rates in the 1997 field experiment than were achieved in the comparison group, but we are not able to clearly determine that this was not due to the use of electronic instruments.

To parallel the debriefing questions included in the field experiment, we selected a subsample of 1997 Quarter 4 respondents and administered an ACASI respondent debriefing questionnaire to them. This subsample was designed to yield 750 respondents for Quarter 4.

Exhibit 5.1.1 Desired Distribution of 1997 Field Experiment Sample									
		ACASI Treatment Groups							
		Contingent Questioning Structure							
		Single Gate Questions				Multiple Gate Questions			
Respondent Characteristics		Consistency Checks				Consistency Checks			
		Absent		Present		Absent		Present	
		Multiple Use Questions		Multiple Use Questions		Multiple Use Questions		Multiple Use Questions	
		Absent	Present	Absent	Present	Absent	Present	Absent	Present
	Treatment Version	1	2	3	4	5	6	7	8
		Number of Respondents							
Total		282	282	282	282	282	282	282	282
Age Group									
	12-17	141	141	141	141	141	141	141	141
	18+	141	141	141	141	141	141	141	141
Race/Ethnicity									
	Hispanic	71	71	71	71	71	71	71	71
	Non-Hisp., Black	71	71	71	71	71	71	71	71
Non-Hisp., All Other Races		141	141	141	141	141	141	141	141

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

5.2 Sample Design

In the following paragraphs, we describe the sampling design for the 1997 field experiment and the sampling design for the 1997 NHSDA debriefing subsample. A summary of the sampling plans for the 1997 Quarter 4 NHSDA, the field experiment, and the Quarter 4 debriefing sample is given in Exhibit 5.2.1.

5.2.1 1997 Field Experiment Sample Design

The sample for the field experiment was confined to 99 purposely selected, geographic PSUs that had been previously selected for the national 1997 NHSDA.[10] Hence, the respondent universe for this component of the field experiment is defined as the civilian, noninstitutionalized population aged 12 years old or older residing in the 99 geographic PSUs. Consistent with the national NHSDA, the respondent universe included civilians residing in military installations and residents of noninstitutional group quarters (e.g., college dormitories, homeless shelters, and rooming houses). The 99 PSUs were selected to be a representative subset of the type of geographic areas that comprise the national NHSDA sample frame and consisted of the 43 PSUs that were selected with certainty for the 1997 national survey and 56 of the 72 noncertainty PSUs. The exclusion of 16 of the 72 noncertainty PSUs was done to reduce field costs.

The 115 PSUs in the national sample consisted of single counties or groups of adjacent counties. There were 43 certainty PSUs, which were so designated because of their high concentration of the Hispanic population, and 72 noncertainty PSUs that were randomly selected with probability proportionate to a composite size with minimum replacement. The composite size measure facilitated oversampling of blacks and Hispanics while simultaneously equalizing the interviewers' workload between PSUs. Controlled ordering of the noncertainty PSUs during the selection process produced a deep implicit stratification based on

geography, metropolitan statistical area (MSA) size, and percentage minority. This implicit stratification helped ensure that the sample of PSUs was as representative of the Nation as possible.

At the second stage of sampling for the 1997 national NHSDA, each of the 115 sample PSUs were partitioned into noncompact clusters of dwelling units, called segments, that were constructed to yield an average of at least 18 responding individuals. A sample "dwelling unit" in the NHSDA refers to either a housing unit or a group quarters listing unit, such as a dormitory room or a shelter bed. The segments were constructed using 1990 Decennial Census data supplemented with revised population counts obtained from several outside sources. Again, segments within PSUs were selected with probability proportionate to a composite size measure with minimum replacement. Altogether, for the national 1997 NHSDA, 1,940 sample segments were selected from the 115 PSUs. After the segment sample was identified, staff visited each area in order to count and list all existing dwelling units that were contained within the segment's geographic boundaries. This list of dwelling units (within the 99 PSUs selected for the field experiment) became the sampling frame for the third stage of selection.

It was possible to use the segments that were listed for the 1997 NHSDA because of two features of the national design. First, the segments in the national 1997 NHSDA contained enough dwelling units for at least two national NHSDAs. This was done to reduce the costs of counting and listing geographic segments because it permitted counting and listing to occur just every other year for the NHSDA series of studies. Because the 1997 NHSDA segments had not been used for a prior NHSDA, some of the listed dwelling units had not been used and were available for use in the 1997 field experiment. Second, the 1997 NHSDA sample segments were assigned to a particular quarter of data collection. Because segments were randomly assigned to the four quarter panels of the national NHSDA, any set of quarters is also a national sample. Thus, for the field experiment, which was to be conducted in Quarter 4, we selected 282 segments from the sample segments that had been used in Quarters 1 through 3 of the 1997 NHSDA. Confining the sample to these segments ensured that the field experiment's data collection effort did not interfere with the national Quarter 4 NHSDA. Thus, for the 1997 field experiment sample, any dwelling unit that was counted and listed within these 282 segments and not selected for the 1997 national NHSDA was eligible for selection. More than enough dwelling units were available because of the practice of counting and listing a number sufficient to conduct two rounds of the national NHSDA.

Sample persons were selected using procedures similar to that used for the national NHSDA. After dwellings units were selected within each segment, an interviewer visited each selected dwelling unit and attempted to collect demographic information on all survey-eligible people residing there. This information was used to determine which racial/ethnic and age groups were represented in the dwelling unit, which in turn was used to classify the dwelling unit into one of 96 household types. As described below, sample selection at the dwelling unit level was completed electronically using the Newton. These 96 selection types were loaded into Apple Newton computers by the sampling staff, and a table lookup procedure that exactly mimicked the 1997 NHSDA paper procedure was used to determine the number of people selected in the dwelling unit. The number selected was either 0, 1, or 2.

After individuals were selected for the field experiment and after they agreed to participate, the computer randomly assigned the respondents to a treatment combination.

Subsampling segments. Initially, some 16,000 dwelling units were selected for the 1997 field experiment. At the end of October 1997, we determined that it was unlikely that the field staff could screen all of these dwelling units. Thus, we eliminated 40 segments that had not yet been released to the field interviewers (FIs). This action resulted in a sample of 14,327 dwelling units for the 1997 field experiment, which, at that time, we thought would still yield more than 2,296 interviews. This projection was based on the following expected yields for the sample: (a) eligibility rate: 80.6%, (b) dwelling unit response rate: 92.9%, (c) selected persons per dwelling unit: 30.5%, and (d) interview response rate: 80.0%.

5.2.2 1997 NHSDA Debriefing Subsample Design

The respondents in the 1997 field experiment received a debriefing questionnaire that asked them about the CAI interviewing experience. To have comparable information from those who had responded using the paper-and-pencil interviewing (PAPI) version of the 1997 NHSDA, a subsample of the 1997 Quarter 4 national NHSDA sample was selected. The debriefing subsample was confined to 66 PSUs and included the 43 PSUs selected with certainty for the 1997 national survey, as well as 23 PSUs that were randomly selected from the 72 noncertainty PSUs selected for the national 1997 survey. Within these PSUs, sampling rates were set to yield a total of 750 individuals. Again, the number of PSUs was limited to minimize interviewer training and data collection costs. Because it was a random sample of the 1997 NHSDA, the target population was the civilian, noninstitutionalized population aged 12 years old or older within the United States. Within these 66 PSUs, 150 segments were selected, with probability proportionate to a composite size, for the debriefing subsample.

At the time we selected this subsample, we anticipated that it would yield 1,068 eligible sample persons and 750 respondents. The actual yield was smaller than expected, and only 713 eligibles were selected.

As noted in Section 5.1.3, the part of this NHSDA Quarter 4 sample that was in the 99 PSUs selected for the field experiment became the comparison group for the field experiment.

Exhibit 5.2.1 Summary of 1997 NHSDA Field Experiment Sample Design: Expected Sample Yields

Sample Stage	National Quarter 4, NHSDA* Sample			Field Experiment Sample (Selection of 2,256 Interviews)			1997 Q4 Subsample (Subselecting 750 Interviews from National Q4 for Debriefing)		
	Certainty PSUs	Noncertainty PSUs	Total	Certainty PSUs	Noncertainty PSUs	Total	Certainty PSUs	Noncertainty PSUs	Total
First Stage - Select PSUs PSUs are counties or groups of counties. Field experiment sample and Q4 subsample are subsamples of the national sample. Total PSUs	43	72	115	43	56	99	43	23	66
Second Stage - Select Segments Field experiment sample PSUs selected from Q1, Q2, and Q4 national samples; 1997 subsample Q4 national sample. Total Segments	269	216	485	153	129	282	82	69	150
Third Stage - Select Dwellings	9,802	8,096	17,898	9,071	7,109	16,179	3,016	2,363	5,379

Field experiment sample selected from dwellings not previously selected for national NHSDA. Total Dwelling Units									
Estimated Eligibility Rate	84.00	84.00	84.00	84.00	84.00	84.00	84.00	84.00	84.00
Estimated Response Rate	94.00	94.00	94.00	94.00	94.00	94.00	94.00	94.00	94.00
Total Completed Screenings	7,740	6,393	14,132	7,162	5,613	12,775	2,381	1,866	4,247
Fourth/Fifth Stage - Select People Total People Selected	3,486	2,926	6,412	1,731	1,483	3,214	575	493	1,068
Estimated Selection Error Rate	90.00	90.00	90.00	90.00	990.00	90.00	90.00	90.00	90.00
Estimated Response Rate	78.00	78.00	78.00	78.00	78.00	78.00	78.00	78.00	78.00
Total Completed Interviews	2,447	2,054	4,501	1,215	1,041	2,256	404	346	750

*The part of this NHSDA Quarter 4 sample that was in the 99 PSUs selected for the field experiment became the comparison group for the field experiment.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

5.3 Electronic Screener

An electronic screener was also used in the 1997 field experiment. In early 1997, we reviewed hardware options and selected the Apple Newton 2000 handheld computer for the application. We used FormLogic from Wright Strategies, Inc., San Diego, California, to develop the screening program. This package offers an MS Windows development platform, suitable for the screener design requirements. Furthermore, Wright Strategies was willing to work with RTI programming staff to ensure timely and robust software support throughout the entire development process.

Development of the Newton NHSDA screening application began in early 1997. The development was an iterative process with extensive usability testing conducted after each cycle. In May 1997, we conducted a small feasibility test in Durham, North Carolina, to examine the application in realistic field conditions. Following this feasibility test, additional development was undertaken to add a number of case management features to the Newton application. Subsequent testing in August 1997 indicated that some of these additional case management features were unstable in the field interviewing environment, and they were dropped from the application. However, the cumulative results of our testing convinced us that the basic Newton screener application was ready to be tested. We developed a contact work sheet for handling the dropped case management activities.

Each interviewer received a Newton that contained the case ID numbers and addresses for all dwelling units in his or her field assignment. When visiting the dwelling unit, the interviewer accessed the address by tapping the specific line containing the address. The screening application guided the interviewers through a series of questions to determine and record in the Newton the number of persons aged 12 or older in the household along with their age, gender, marital status, race/ethnicity, and military status.[11] Using this information, the screening application consulted sample selection tables and indicated to the interviewer whether none, one, or two respondents had been selected for the interview. It displayed the characteristics of the selected persons to the interviewer. Because no names are collected, the sample persons were identified by their age, gender, marital status, and race/ethnicity.

5.4 Screening and Interview Response Rates

The screening and interview response rates in the 1997 field experiment were lower than those we achieved in the main study. Exhibit 5.4.1 displays the screening rates for the 1997 field experiment and for two comparison groups: 1) the NHSDA Quarter 4 sample that was in the same PSUs as the field experiment and 2) those test Quarters 1 through 3 of the NHSDA for the same sample segments as were used for the field experiment. If we consider group quarters, which were not included in the Newton application, as ineligible for the 1997 field experiment, the overall screening response rate was 86%, which is about 7% lower than that achieved in similar areas in the national NHSDA. About 2.5% of this shortfall was due to the failure to obtain access to restricted housing; 3.5% was due to increased refusals. It is unlikely that the electronic screener contributed to the failure to obtain access to restricted housing. However, we are not able, from this study alone, to verify that using the Newton did not contribute to increased refusals to the screening.

In Exhibit 5.4.2, we present the results of attempting to interview the selected persons. To obtain a consistent picture of the coverage of the population for the 1997 field experiment and the main NHSDA, we counted people who had completed the main NHSDA in Spanish as nonrespondents because a Spanish interview was not available for the field experiment. Overall, only about 62.7% of the selected persons in the 1997 field experiment were interviewed, whereas the equivalent percentage for the NHSDA main study is 75.6%, which is a 12.9% difference. Again, it is not possible to determine if this difference is due to using electronic instruments.

Exhibit 5.4.3 presents the distribution of the 1997 field experiment's respondents by treatment. The overall sample size was 1,982; 56% of the respondents were youths aged 12 to 17, which was higher than we planned for and is due to youths responding at a higher rate than adults.

We compared the demographics of the comparison group to those of the field experiment and show the results in Exhibits 5.4.4 through 5.4.6. Except for the planned-for larger numbers of youths in the field experiment, there are few differences. Thus, the comparisons are likely to be valid.

Exhibit 5.4.1 Screening Response Rates: Comparison of 1997 Field Experiment to Selected 1997 NHSDA Experience

	Percent of Dwelling Units			Number of Dwelling Units		
Response Rates	1997 Field Experiment	NHSDA Q4 Comparison Group	Quarters 1-3 Field Experiment Segments	1997 Field Experiment	NHSDA Q4 Comparison Group	Quarters 1-3 Field Experiment Segments
Ineligible Dwelling Unit	**16.28**	**15.30**	**16.75**	**2,333**	**8,782**	**1,415**
Vacant	11.8	11.32	11.35	1,690	6,500	959
Not a Primary Residence	1.63	1.79	2.66	234	1,028	225
Not a Dwelling Unit	2.3	2.07	2.62	329	1,187	221
Other	0.56	0.12	0.12	80	67	10

Eligible Dwelling Unit, Not Screened for Eligible Persons	**14.6**	**6.37**	**5.35**	**1,751**	**3,099**	**376**
No One at Home	2.08	2.09	1.98	250	1,014	139
Refusal	5.84	2.46	2.34	701	1,196	163
Denied Access	2.91	0.59	0.26	349	287	18
Newton Screener Problem	0.57	0.0	0.0	63	0	0
Other Nonresponse (Group Quarters)[1]	1.39	N/A	N/A	167	N/A	N/A
Other Nonresponse	1.80	1.24	0.8	216	602	56
Eligible Dwelling Units, Screened for Eligible Persons	**85.4**	**93.63**	**94.65**	**10,243**	**45,529**	**6,658**
Total Lines Selected	**100.0**	**100.0**	**100.0**	**14,327**	**57,410**	**8,449**

[1]The Newton application used for the 1997 field experiment did not handle group quarters.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 5.4.2 Distribution of Final Response Codes for Selected Persons

Selected Persons	**1997 Field Experiment**			**NHSDA Q 4 Comparison Group**			**Qtrs 1-3 NHSDA Field Experiment Segments**		
	Sample Size	% of Total Selected	% of Total Non-complete	Sample Size	% of Total Selected	% of Total Non-complete	Sample Size	% of Total Selected	% of Total Non-complete
Total	**3,163**	**100.00%**		**4,110**	**100.00%**		**2,891**	**100.00%**	
Respondents	**1,982**	**62.66%**		**3,105**	**75.55%**		**2,186**	**75.61%**	
Nonrespondents	**1,181**	**37.34%**	**100.00%**	**1,005**	**24.45%**	**100.00%**	**705**	**24.39%**	**100.00%**
No One Home, R Unavailable	180	5.69%	15.24%	293	7.13%	29.15%	179	6.19%	25.39%
Physical / Mentally Incompetent	57	1.80%	4.83%	37	0.90%	3.68%	22	0.76%	3.12%
Language Barrier	188	5.94%	15.92%	268	6.52%	26.67%	215	7.44%	30.50%
Refusal	625	19.76%	52.92%	358	8.71%	35.62%	256	8.86%	36.31%
Other	131	4.14%	11.09%	49	1.19%	4.88%	33	1.14%	4.68%

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse.

Exhibit 5.4.3 Distribution of 1997 Field Experiment Respondents

Respondent	**ACASI Treatment Groups**

Characteristics	Contingent Questioning Structure								Total ACASI
	Single Gate Questions				Multiple Gate Questions				
Treatment Version	Consistency Checks				Consistency Checks				
	Absent		Present		Absent		Present		
	Multiple Use Questions		Multiple Use Questions		Multiple Use Questions		Multiple Use Questions		
	Absent	Present	Absent	Present	Absent	Present	Absent	Present	
	1	2	3	4	5	6	7	8	
Total	**208**	**314**	**285**	**264**	**245**	**240**	**219**	**207**	**1,982**
Age Group									
12-17	118	179	157	148	142	142	118	113	1,117
18 +	90	135	128	116	103	98	101	94	865
Gender									
Males	112	139	138	123	119	110	98	88	927
Females	96	175	147	141	126	130	121	119	1,055
Race/Ethnicity									
Hispanic	45	73	66	62	63	61	49	51	470
Non-Hisp., Black	55	76	79	70	58	67	63	63	531
Non-Hisp., All Other Races	108	165	140	132	124	112	107	93	981
Education[1]									
< High School	26	28	20	32	25	27	22	19	199
High School	35	52	49	41	31	34	38	41	321
> High School	29	55	59	43	47	37	41	34	345

[1] Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 5.4.4 Demographic Distributions of Field Experiment and 1997 NHSDA Comparison Group

Respondent Characteristics		**1997 Field Experiment**		**Comparison Group**	
		CAPI / ACASI		**1997 Quarter 4 PAPI / SAQ**	
		Number of Respondents	*Percentage*	*Number of Respondents*	*Percentage*
Total		1,982	100.0%	3,105	100.0%

Age Group					
	12-17	1,117	56.36%	979	31.53%
	18+	865	43.64%	2,126	68.47%
Gender					
	Males	927	46.77%	1,265	40.74%
	Females	1,055	53.23%	1,840	59.26%
Race / Ethnicity					
	Hispanic	470	23.71%	572	18.42%
	Non-Hisp., Black	531	26.79%	1,023	32.95%
	Non-Hisp., All Other Races	981	49.50%	1,510	48.63%
Education[1]					
	< High School	199	23.01%	406	19.10%
	High School	321	37.11%	790	37.16%
	> High School	345	39.88%	930	43.74%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 5.4.5 Distribution of Field Experiment and Comparison Group Respondents, by Work Status

Work Status		**1997 Field Experiment**		**Comparison Group**	
		CAPI / ACASI		**1997 Quarter 4 PAPI / SAQ**	
		Number of Respondents	*Percentage*	*Number of Respondents*	*Percentage*
Total		1,979	100.0%	3,091	100.0%
Employed		755	38.15%	1,552	50.21%
	Full-Time	497	25.11%	1,161	37.56%
	Part-Time	228	11.52%	368	11.91%
	Extended Leave	30	1.52%	23	0.74%
Not Employed		1,224	61.85%	1,539	49.79%
	Looking	56	2.83%	113	3.66%
	Not Looking	25	1.26%	24	0.78%
	Student	979	49.47%	958	30.99%

	Disabled	19	0.96%	70	2.26%
	Other[1]	145	7.33%	374	12.10%

[1]Other includes full-time homemaker, retired, and not specified.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 5.4.6 Distribution of 1997 Field Experiment and Comparison Group Respondents, by Income

Income ($ US)		1997 Field Experiment		Comparison Group	
		CAPI / ACASI		1997 Quarter 4 PAPI / SAQ	
		Number of Respondents	*Percentage*	*Number of Respondents*	*Percentage*
Total		1,957	100.0%	3,080	100.0%
	< 20,000	1,497	76.49%	2,220	72.08%
	20,000 to 50,000	307	15.69%	617	20.03%
	> 50,000	153	7.82%	243	7.89%

Sources: National Household Survey on Drug Abuse: Development of Computer Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

5.5 Analysis Weights

The analysis weights for the 1997 field experiment sample were constructed as follows. Control totals were created by summing the fully adjusted national sample weights for the 99 PSUs that contained the field experiment sample. All four quarters of data were used to create the control totals. An initial weight was calculated as the ratio of the total sum of weights divided by the sum of weights for the respondents. These weights were adjusted to the control totals within age, race/ethnicity, and gender subgroups using a raking procedure at both the household and person at both the household and personal level.

[9] Because the structure of the questionnaire required respondents to first indicate the time period of their most recent use and to then indicate the number of days used in that period, there are some implicit multiple use questions in every interview and these were analyzed as well.

[10] The PSUs and segments (segments are the geographic second stage sampling units in the NHSDA) selected for the 1997 NHSDA California/Arizona supplemental sample were not considered during the design of the 1997 field experiment.

[11] People who are on active military duty are not eligible for the NHSDA.

6. Comparisons of CAPI/ACASI to PAPI for Selected Outcomes: 1997 Field Experiment

In this chapter, we compare, for selected outcomes, the computer-assisted interviewing (CAI) procedures with the paper-and-pencil interview (PAPI) procedures. We examine the effects for the different modes of interview on the following topics:

1. the reported prevalence of alcohol, marijuana, cocaine, and any illicit drug use;

2. the reported prevalence of nonmedical use of psychotherapeutic drugs;

3. answers to questions on mental health, which were interviewer administered in the 1997 NHSDA and self-administered via audio computer-assisted self-interviewing (ACASI) in the 1997 field experiment;

4. answers to questions on methods of use, risk and availability of drugs, and drug dependence; and

5. the reporting of 12-month frequency of use.

We also discuss the respondent and interviewer assessments of the interviewing environment.

6.1 Alcohol, Marijuana, Cocaine/Crack, and Any Illicit Drug Use in Lifetime, Past Year, and Past Month

Exhibit 6.1.1 contains the weighted prevalence of tobacco, alcohol, marijuana, cocaine, and any illicit drug use averaged across the ACASI treatment groups and for the 1997 NHSDA Quarter 4 comparison group. [12] Exhibit 6.1.2 presents the prevalence estimates by treatment level of the three experimental factors.

To compare ACASI treatments to PAPI, we fit a model with the main effects of the treatment group and the covariates. Levels 1-8 represented the eight ACASI experimental combinations used in the study, and Level 9 represented PAPI (the covariates in the model were gender, age, race/ethnicity, and education):

1 = Single Gate, No Consistency Checks, No Multiple Use;
2 = Single Gate, No Consistency Checks, Multiple Use;
3 = Single Gate, Consistency Checks, No Multiple Use;
4 = Single Gate, Consistency Checks, Multiple Use;
5 = Multiple Gate, No Consistency Checks, No Multiple Use;
6 = Multiple Gate, No Consistency Checks, Multiple Use;
7 = Multiple Gate, Consistency Checks, No Multiple Use;
8 = Multiple Gate, Consistency Checks, Multiple Use; and
9 = 1997 Quarter 4 PAPI.

The following single degree-of-freedom contrasts were constructed from the nine-level treatment effects and evaluated as follows:

1. ACASI as a whole (Levels 1-8) versus PAPI (Level 9),
2. Single Gate (Levels 1-4) versus PAPI (Level 9),
3. Multiple Gate (Levels 5-8) versus PAPI (Level 9),
4. Multiple Use Absent (Levels 1, 3, 5, and 7) versus PAPI (Level 9),
5. Multiple Use Present (Levels 2, 4, 6, and 8) versus PAPI (Level 9),

6. Consistency Checks Absent (Levels 1, 2, 5, and 6) versus PAPI (Level 9), and
7. Consistency Checks Present (Levels 3, 4, 7, and 8) versus PAPI (Level 9).

Because there are no interaction effects in this model, all contrasts were automatically adjusted for other covariates present in the model. Wald chi-square tests were used to evaluate each of these effects.

The logistic regression was used because response variable *Y* takes on the values 0 (nonuser) or 1 (drug user). The set of treatments and covariates is denoted by *X* and includes the treatment groups and the covariates. In the logistic regression model, $p = prob\,(Y = 1 \mid x)$ is the response probability that a respondent has used the drug, given his or her set of covariates *X*.

The logit form of the model is formally stated as follows:

$$\log\left(\frac{p}{1-p}\right) = \boldsymbol{X\beta} = \alpha + \beta_1 X_1 + \ldots \beta_k X_k ,$$

or, equivalently $p = prob(Y=1 \mid \boldsymbol{x}) = [1 + \exp(-\boldsymbol{X\beta})]^{-1}$.

The expression $X\beta$ represents the linear combination of covariate effects for each respondent, where β is the set of regression coefficients to be estimated and *q* is the number of covariates in the model.

The expression $\left(\frac{p}{1-p}\right)$ represents the *odds* that a particular respondent has used the drug. The model response, log $\left(\frac{p}{1-p}\right)$, is referred to as the log odds or the logit of p. Exponentiating the regression coefficients provides estimates of the *odds ratio* for different values of *X*, such as for respondents in one of the ACASI treatment cells versus the control (PAPI) group. These statistics quantify the strength of the treatment effect.

Exhibit 6.1.3 summarizes the modeling results for the overall comparison of ACASI to PAPI for the total sample and 12 to 17 year olds. In conducting these analyses, we classified the 12 to 17 year olds as less than high school graduates in order for all respondents to be included in the overall analysis. A total of 5,087 respondents were included in the analysis of the total sample, and 2,096 respondents were included in the analysis of 12 to 17 year olds. There were no missing values in this analysis because the privacy covariates were not included in the model (privacy information was unavailable for the PAPI group). Covariates in the model for the total sample analyses included gender, age (12 to 17 vs. 18+), race (Hispanic; Non-Hispanic, All Other Races; Non-Hisp., Black), and education (less than high school, high school, more than high school). Covariates in the model for the 12 to 17 year olds included gender and race only.

For the overall sample, there were only two significant differences: lifetime cocaine use and past year use of any illicit drug. In each case, ACASI yielded higher prevalence estimates, and the odds ratio was greater than 1.5. Among the youths (12 to 17 year olds), there were seven significant differences observed. Significantly higher reports of prevalence under ACASI were observed for lifetime cigarette use (OR=1.43), lifetime alcohol use (OR=1.35), lifetime marijuana use (OR=1.36), and past year marijuana use (OR=1.35). Also, for the summary variable, any illicit drug, ACASI yielded significantly higher reports for all three reporting periods: lifetime (OR=1.76), past year (OR=1.76), and past month (OR=1.57) use.

In addition to these analyses, for the total sample we also compared multiple gate versus PAPI, single gate versus PAPI, multiple chance present versus PAPI, multiple chance absent versus PAPI,

consistency checks present versus PAPI, and consistency checks absent versus PAPI. Exhibit 6.1.4 presents the results, and Exhibit 6.1.5 summarizes the comparisons that were either significant or marginally significant. Significantly higher reports under PAPI were observed in only one case. The reported prevalence of past month cigarette use by the total sample was significantly reduced in the absence of multiple chances to report use compared to PAPI (29.8% PAPI vs. 21.2% multiple chance absent).

The reported prevalence of past year alcohol use by the total sample was marginally increased ($0.05 < p < 0.10$) in the presence of multiple chances to report use and in the presence of consistency checks compared to PAPI. Prevalence rates ranged from 66% for PAPI to 72.8% and 73.3% for the presence of multiple chance and consistency checks, respectively.

Total sample analyses of cocaine use showed that reported prevalence of lifetime use was significantly increased for overall ACASI versus PAPI (13.6% vs. 9.0% for ACASI vs. PAPI, respectively). Reported prevalence of lifetime use was also significantly increased in the single gate version, presence of multiple chances, and presence of consistency checks versus PAPI. Prevalence rates ranged from 9.0% for PAPI versus 17.8% for the single gate version, 16.4% for the presence of multiple chances, and 17.2% for the presence of consistency checks. Reported prevalence of past year use was significantly increased in the single gate questions versus PAPI group, with prevalences of 3.5% for single gate and 1.5% for PAPI.

Total sample analyses of any illicit drug use revealed that reported prevalence of lifetime use was marginally increased ($0.05 < p < 0.10$) for overall ACASI compared to PAPI (43.5% for ACASI vs. 37.2% for PAPI). Also for lifetime use, prevalence was marginally increased in the single gate treatment compared to PAPI and in the presence of multiple chances compared to PAPI (prevalence rates were just over 46% for the two ACASI classifications and 37.2% for PAPI). Reported prevalence of past year use was significantly increased for overall ACASI compared to PAPI (14.8% for ACASI vs. 10.0% for PAPI). Also for past year use, prevalence was significantly increased in the single gate versus PAPI in the absence of multiple chances versus PAPI and in the presence of consistency checks versus PAPI. Prevalence was marginally increased in the absence of consistency checks versus PAPI (16.6% for single gate, 14.3% absence of multiple chances, 14.1% for absence and presence of consistency checks, 10.0% PAPI).

Exhibit 6.1.1 Prevalence by Mode of Interview and Age of Respondent--Weighted and Edited Estimates

Reporting Period	Overall		12 to 17 Year Olds		18+ Year Olds	
	ACASI	**PAPI**	**ACASI**	**PAPI**	**ACASI**	**PAPI**
Cigarettes						
Lifetime Use	72.0	71.0	44.5	36.1	75.2	74.9
Past Year Use	31.2	33.5	22.9	25.2	32.2	34.5
Past Month Use	26.2	30.5	16.1	18.6	27.4	31.8
Alcohol						
Lifetime Use	83.2	82.2	44.9	38.1	87.6	87.3

Past Year Use	68.3	65.0	36.5	32.7	72.0	68.7
Past Month Use	48.7	52.2	17.0	18.8	52.4	56.1
Marijuana						
Lifetime Use	35.2	35.0	20.6	16.1	36.9	37.2
Past Year Use	10.2	9.3	16.7	13.0	9.5	8.9
Past Month Use	4.7	5.3	8.4	7.3	4.3	5.0
Cocaine						
Lifetime Use	14.9	10.0	2.9	3.1	16.3	10.8
Past Year Use	2.6	1.6	1.8	2.5	2.7	1.5
Past Month Use	0.6	0.6	0.7	1.1	0.6	0.5
Any Illicit Drug[1]						
Lifetime Use	43.7	18.8	31.2	20.5	45.2	39.5
Past Year Use	15.2	10.3	23.0	14.5	14.2	9.8
Past Month Use	5.9	5.4	12.0	8.0	5.2	5.1

[1]Any illicit drug includes marijuana, cocaine, crack, heroin, inhalants, hallucinogens, and nonmedical use of analgesics, sedatives, stimulants, and tranquilizers.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.1.2 Prevalence Estimates for ACASI Experimental Factors: Overall and by Age									
	Overall			**18+ Years Old**			**12 to 17 Years Old**		
	Lifetime	12 Month	30 Day	Lifetime	12 Month	30 Day	Lifetime	12 Month	30 Day

Cigarettes									
Single gate questions	74.2	29.2	26.6	77.9	30.0	27.8	43.2	22.7	16.6
Multiple gate questions	69.4	33.5	25.7	72.0	34.6	26.9	46.0	23.0	15.5
Multiple use questions absent	73.9	29.2	21.7	76.8	29.4	22.2	47.4	27.4	17.3
Multiple use questions present	70.0	33.3	31.0	73.4	35.2	33.0	41.7	18.5	15.0
Consistency checks absent	69.7	28.6	22.9	72.6	29.1	23.8	45.5	23.8	15.5
Consistency checks/present	74.5	34.1	29.9	78.0	35.5	31.3	43.3	21.8	16.8
Alcohol									
Single gate questions	82.0	65.1	47.3	86.4	68.5	50.9	44.9	36.4	17.4
Multiple gate questions	84.5	72.0	50.3	89.0	76.0	54.1	44.9	36.5	16.6
Multiple use questions absent	83.5	66.2	50.2	87.5	69.3	53.7	46.9	37.7	18.0
Multiple use questions present	82.7	70.6	47.0	87.7	75.0	50.9	43.0	35.3	16.1
Consistency checks absent	82.1	65.1	44.8	86.5	68.5	48.3	45.6	37.0	16.2
Consistency checks/present	84.3	71.8	53.0	88.8	75.8	56.9	44.1	35.8	18.1
Marijuana									
Single gate questions	35.4	12.5	5.5	36.9	11.9	5.1	22.1	17.6	9.4
Multiple gate questions	35.0	7.6	3.8	36.8	6.7	3.4	18.7	15.6	7.1
Multiple use questions absent	33.4	9.1	5.3	34.6	8.3	4.9	21.8	17.0	9.8
Multiple use questions	37.1	11.4	4.0	39.3	10.8	3.7	19.4	16.4	7.0

present									
Consistency checks absent	35.0	10.0	3.7	36.8	9.2	3.2	20.7	16.2	7.7
Consistency checks/present	35.3	10.6	5.8	37.0	9.8	5.5	20.4	17.3	9.1
Cocaine									
Single gate questions	19.2	3.7	0.6	21.0	3.9	0.6	3.6	2.2	1.0
Multiple gate questions	9.9	1.4	0.5	10.8	1.4	0.6	1.9	1.2	0.3
Multiple use questions absent	12.1	3.1	0.6	13.1	3.3	0.6	2.7	1.7	0.2
Multiple use questions present	17.9	2.1	0.6	19.8	2.2	0.5	3.0	1.8	1.1
Consistency checks absent	11.6	2.3	0.6	12.7	2.3	0.5	2.5	1.7	1.0
Consistency checks/present	18.5	3.0	0.5	20.3	3.2	0.6	3.3	1.8	0.2
Any Illicit Drug[1]									
Single gate questions	46.3	16.9	6.9	47.8	16.1	6.2	33.5	23.7	13.0
Multiple gate questions	40.7	13.1	4.7	42.2	12.1	4.1	28.3	22.2	10.6
Multiple use questions absent	41.6	14.5	6.6	42.6	13.5	5.9	32.8	24.3	13.1
Multiple use questions present	46.0	15.8	5.1	48.0	15.0	4.4	29.6	21.9	10.8
Consistency checks absent	43.4	14.6	4.7	45.0	13.8	4.0	30.8	21.2	10.4
Consistency checks/present	44.0	15.8	7.3	45.4	14.8	6.6	31.5	25.2	13.7

[1]Any illicit drug includes marijuana, cocaine, crack, heroin, inhalants, hallucinogens, and nonmedical use of analgesics, sedatives, stimulants, and tranquilizers.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.1.3 Results of Logistic Regression Modeling of ACASI Versus PAPI

	Adjusted Prevalence			
	ACASI	**PAPI**	***p* value**	**Odds Ratio**
Overall: Adjusted for gender, age group, education, and race/ethnicity				
Cigarettes: Lifetime Use	73.0	71.8	0.72	1.06
Past Year Use	30.8	33.1	0.51	0.90
Past Month Use	25.5	29.8	0.16	0.81
Alcohol: Lifetime Use	86.5	85.5	0.66	1.09
Past Year Use	69.5	66.0	0.32	1.17
Past Month Use	47.8	51.9	0.32	0.85
Marijuana: Lifetime Use	34.6	34.4	0.95	1.01
Past Year Use	9.5	8.6	0.63	1.11
Past Month Use	4.1	4.6	0.72	0.90
Cocaine: Lifetime Use	13.6	9.0	0.05	1.59
Past Year Use	2.5	1.5	0.15	1.67
Past Month Use	0.4	0.4	0.97	0.98
Any Illicit Drug:[1]	46.5	37.2	0.08	1.30

Lifetime Use				
Past Year Use	14.8	10.0	0.00	1.57
Past Month Use	5.2	4.8	0.73	1.10
12 to 17 Year Olds: Adjusted for gender and race/ethnicity				
Cigarettes: Lifetime Use	44.5	35.7	0.002	1.43
Past Year Use	22.7	24.9	0.37	0.89
Past Month Use	15.9	18.3	0.22	0.84
Alcohol: Lifetime Use	45.1	37.9	0.007	1.35
Past Year Use	36.6	32.6	0.12	1.19
Past Month Use	16.7	18.5	0.45	0.88
Marijuana: Lifetime Use	20.6	16.1	0.013	1.36
Past Year Use	16.7	12.9	0.041	1.35
Past Month Use	8.3	7.3	0.47	1.16
Cocaine: Lifetime Use	2.6	2.9	0.83	0.91
Past Year Use	1.5	2.1	0.51	0.69
Past Month Use	0.6	0.9	0.47	0.60
Any Illicit Drug:[1]	31.1	20.4	0.000	1.76

Lifetime Use				
Past Year Use	23.0	14.5	0.000	1.76
Past Month Use	11.9	7.9	0.023	1.57

[1]Any illicit drug includes marijuana, cocaine, crack, heroin, inhalants, and hallucinogens, as well as nonmedical use of analgesics, sedatives, stimulants, and tranquilizers.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.1.4 Logistic Regression Modeling Results: ACASI Treatment Groups Versus PAPI

Effect	**Effect Categories**	***p* value**	**Odds Ratio**	**Adjusted Prevalence**
Lifetime Cigarette Use				
ACASI vs. PAPI	ACASI	0.72	1.06	73.0%
Multiple Gate vs. PAPI	Multiple Gate	0.81	0.94	70.5%
Single Gate vs. PAPI	Single Gate	0.32	1.18	75.1%
Multiple Chance: Present vs. PAPI	Present	0.85	0.96	71.5%
Multiple Chance: Absent vs. PAPI	Absent	0.45	1.16	74.8%
Consistency Checks: Present vs. PAPI	Present	0.30	1.24	75.9%
Consistency Checks: Absent vs. PAPI	Absent	0.68	0.93	70.2%
Past Year Cigarette Use				

ACASI vs. PAPI	ACASI	0.51	0.90	30.8%
Multiple Gate vs. PAPI	Multiple Gate	0.93	1.02	33.6%
Single Gate vs. PAPI	Single Gate	0.29	0.80	28.5%
Multiple Chance: Present vs. PAPI	Present	0.96	0.99	32.9%
Multiple Chance: Absent vs. PAPI	Absent	0.38	0.82	28.9%
Consistency Checks: Present vs. PAPI	Present	0.74	1.05	34.2%
Consistency Checks: Absent vs. PAPI	Absent	0.28	0.78	27.8%
Past Month Cigarette Use				
ACASI vs. PAPI	ACASI	0.16	0.81	25.5%
Multiple Gate vs. PAPI	Multiple Gate	0.34	0.81	25.6%
Single Gate vs. PAPI	Single Gate	0.34	0.81	25.5%
Multiple Chance: Present vs. PAPI	Present	0.92	1.02	30.2%
Multiple Chance: Absent vs. PAPI	Absent	0.04	0.64	21.2%
Consistency Checks: Present vs. PAPI	Present	0.99	0.99	29.7%
Consistency	Absent	0.05	0.66	21.8%

Checks: Absent vs. PAPI				
Lifetime Alcohol Use				
ACASI vs. PAPI	ACASI	0.66	1.09	86.5%
Multiple Gate vs. PAPI	Multiple Gate	0.42	1.21	87.7%
Single Gate vs. PAPI	Single Gate	0.98	1.00	85.4%
Multiple Chance: Present vs. PAPI	Present	0.62	1.10	86.6%
Multiple Chance: Absent vs. PAPI	Absent	0.78	1.07	86.3%
Consistency Checks: Present vs. PAPI	Present	0.34	1.20	87.6%
Consistency Checks: Absent vs. PAPI	Absent	0.97	0.99	85.4%
Past Year Alcohol Use				
ACASI vs. PAPI	ACASI	0.32	1.17	69.5%
Multiple Gate vs. PAPI	Multiple Gate	0.13	1.40	73.1%
Single Gate vs. PAPI	Single Gate	0.92	1.02	66.5%
Multiple Chance: Present vs. PAPI	Present	0.06	1.38	72.8%
Multiple Chance: Absent vs. PAPI	Absent	0.95	1.01	66.3%

Consistency Checks: Present vs. PAPI	Present	0.08	1.42	73.3%
Consistency Checks: Absent vs. PAPI	Absent	0.99	0.99	66.0%
Past Month Alcohol Use				
ACASI vs. PAPI	ACASI	0.33	0.85	47.8%
Multiple Gate vs. PAPI	Multiple Gate	0.61	0.89	48.9%
Single Gate vs. PAPI	Single Gate	0.37	0.82	46.9%
Multiple Chance: Present vs. PAPI	Present	0.38	0.82	47.0%
Multiple Chance: Absent vs. PAPI	Absent	0.49	0.88	48.6%
Consistency Checks: Present vs. PAPI	Present	0.91	1.03	52.6%
Consistency Checks: Absent vs. PAPI	Absent	0.11	0.72	43.5%
Lifetime Marijuana Use				
ACASI vs. PAPI	ACASI	0.95	1.01	34.6%
Multiple Gate vs. PAPI	Multiple Gate	0.94	0.98	34.1%
Single Gate vs. PAPI	Single Gate	0.87	1.03	35.1%
Multiple Chance: Present vs. PAPI	Present	0.53	1.12	37.1%

Multiple Chance: Absent vs. PAPI	Absent	0.61	0.91	32.4%
Consistency Checks: Present vs. PAPI	Present	0.92	1.02	34.9%
Consistency Checks: Absent vs. PAPI	Absent	1.00	1.00	34.4%
Past Year Marijuana Use				
ACASI vs. PAPI	ACASI	0.60	1.11	9.5%
Multiple Gate vs. PAPI	Multiple Gate	0.32	0.80	7.0%
Single Gate vs. PAPI	Single Gate	0.15	1.40	11.6%
Multiple Chance: Present vs. PAPI	Present	0.39	1.25	10.6%
Multiple Chance: Absent vs. PAPI	Absent	0.94	0.98	8.5%
Consistency Checks: Present vs. PAPI	Present	0.58	1.18	10.0%
Consistency Checks: Absent vs. PAPI	Absent	0.80	1.05	9.0%
Past Month Marijuana Use				
ACASI vs. PAPI	ACASI	0.72	0.90	4.1%
Multiple Gate vs. PAPI	Multiple Gate	0.31	0.72	3.4%
Single Gate vs. PAPI	Single Gate	0.91	1.04	4.8%

Multiple Chance: Present vs. PAPI	Present	0.43	0.75	3.5%
Multiple Chance: Absent vs. PAPI	Absent	0.92	1.04	4.7%
Consistency Checks: Present vs. PAPI	Present	0.73	1.15	5.2%
Consistency Checks: Absent vs. PAPI	Absent	0.16	0.68	3.2%
Lifetime Cocaine Use				
ACASI vs. PAPI	ACASI	0.05	1.59	13.6%
Multiple Gate vs. PAPI	Multiple Gate	0.93	0.98	8.8%
Single Gate vs. PAPI	Single Gate	0.01	2.19	17.8%
Multiple Chance: Present vs. PAPI	Present	0.01	1.99	16.4%
Multiple Chance: Absent vs. PAPI	Absent	0.50	1.24	11.0%
Consistency Checks: Present vs. PAPI	Present	0.04	2.11	17.2%
Consistency Checks: Absent vs. PAPI	Absent	0.57	1.17	10.3%
Past Year Cocaine Use				
ACASI vs. PAPI	ACASI	0.15	1.67	2.5%
Multiple	Multiple Gate	0.68	0.87	1.3%

Gate vs. PAPI				
Single Gate vs. PAPI	Single Gate	0.04	2.40	3.5%
Multiple Chance: Present vs. PAPI	Present	0.47	1.32	1.9%
Multiple Chance: Absent vs. PAPI	Absent	0.14	2.02	2.9%
Consistency Checks: Present vs. PAPI	Present	0.19	1.94	2.8%
Consistency Checks: Absent vs. PAPI	Absent	0.27	1.44	2.1%
Past Month Cocaine Use				
ACASI vs. PAPI	ACASI	0.97	0.98	0.36%
Multiple Gate vs. PAPI	Multiple Gate	0.87	0.93	0.34%
Single Gate vs. PAPI	Single Gate	0.95	1.03	0.38%
Multiple Chance: Present vs. PAPI	Present	0.91	0.96	0.35%
Multiple Chance: Absent vs. PAPI	Absent	0.98	1.01	0.37%
Consistency Checks: Present vs. PAPI	Present	0.89	0.94	0.35%
Consistency Checks: Absent vs. PAPI	Absent	0.96	1.03	0.38%

Any Lifetime Illicit Drug Use[1]				
ACASI vs. PAPI	ACASI	0.08	1.30	43.5%
Multiple Gate vs. PAPI	Multiple Gate	0.50	1.14	40.4%
Single Gate vs. PAPI	Single Gate	0.06	1.45	46.2%
Multiple Chance: Present vs. PAPI	Present	0.06	1.44	46.0%
Multiple Chance: Absent vs. PAPI	Absent	0.41	1.18	41.2%
Consistency Checks: Present vs. PAPI	Present	0.22	1.33	44.0%
Consistency Checks: Absent vs. PAPI	Absent	0.22	1.28	43.1%
Any Past Year Illicit Drug Use[1]				
ACASI vs. PAPI	ACASI	0.004	1.57	14.8%
Multiple Gate vs. PAPI	Multiple Gate	0.21	1.32	12.8%
Single Gate vs. PAPI	Single Gate	0.01	1.79	16.6%
Multiple Chance: Present vs. PAPI	Present	0.02	1.64	15.4%
Multiple Chance: Absent vs. PAPI	Absent	0.03	1.50	14.3%
Consistency Checks: Present vs. PAPI	Present	0.04	1.67	15.7%

Consistency Checks: Absent vs. PAPI	Absent	0.05	1.48	14.1%
Any Past Month Illicit Drug Use[1]				
ACASI vs. PAPI	ACASI	0.73	1.10	5.2%
Multiple Gate vs. PAPI	Multiple Gate	0.63	0.88	4.2%
Single Gate vs. PAPI	Single Gate	0.44	1.29	6.1%
Multiple Chance: Present vs. PAPI	Present	0.81	0.93	4.4%
Multiple Chance: Absent vs. PAPI	Absent	0.45	1.27	6.0%
Consistency Checks: Present vs. PAPI	Present	0.31	1.41	6.6%
Consistency Checks: Absent vs. PAPI	Absent	0.50	0.83	4.0%

[1]Any illicit drug includes marijuana, cocaine, crack, heroin, inhalants, hallucinogens, and nonmedical use of analgesics, sedatives, stimulants, and tranquilizers.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.1.5 Statistically Significant Modeling Results: Prevalence of Drug Use in the ACASI Treatment Groups Versus PAPI (Weighted Results)

Outcome	Effects	*p* value	Odds Ratio	Adjusted Prevalence
Total Sample				
Cigarettes: Past Month	Multiple Chance Absent vs. PAPI	0.04	0.64	21.2%

Alcohol: Past Year	Multiple Chance Present vs. PAPI	0.06	1.38	72.8%
	Consistency Checks Present vs. PAPI	0.08	1.42	73.3%
Cocaine: Lifetime	ACASI vs. PAPI	0.05	1.59	13.6%
	Single Gate vs. PAPI	0.01	2.19	17.8%
	Multiple Chance Present vs. PAPI	0.01	1.99	16.4%
	Consistency Checks Present vs. PAPI	0.04	2.11	17.2%
Cocaine: Past Year	Single Gate vs. PAPI	0.04	2.40	3.5%
Any Illicit Drug Use:[1] Lifetime	ACASI vs. PAPI	0.08	1.3	43.5%
	Single Gate vs. PAPI	0.06	1.45	46.2%
	Multiple Chance Present vs. PAPI	0.06	1.44	46.0%
Any Illicit Drug Use:[1] Past Year	ACASI vs. PAPI	0.004	1.57	14.8%
	Single Gate vs. PAPI	0.01	1.79	16.6%
	Multiple Chance Absent vs. PAPI	0.03	1.50	14.3%
	Consistency Checks Absent vs. PAPI	0.05	1.48	14.1%
	Consistency Checks Present vs. PAPI	0.04	1.67	14.1%

[1]Any illicit drug includes marijuana, cocaine, crack, heroin, inhalants, hallucinogens, and nonmedical use of analgesics, sedatives, stimulants, and tranquilizers.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

6.2 Reporting of Nonmedical Use of Psychotherapeutic Drugs

There was a single version of the ACASI program for questions on the nonmedical use of psychotherapeutic drugs, which include analgesics, tranquilizers, stimulants, and sedatives. This version differed from the PAPI answer sheets in several respects. As for the other sections of the core interview, the ACASI provided a more private mode of data collection. The "answer all questions" privacy-enhancing procedure was not used in these answer sheets, and the protocol allowed respondents to be routed past the detailed questions if they had never used one of the specific drugs. Thus, the ACASI privacy enhancement was even greater for these drugs than for other sections of the core. In the PAPI answer sheet, respondents received the list of drugs in a matrix format listing the drug on the left and two columns of check boxes on the right. One column was labeled "yes" and the other "no," and the respondent was asked to check a box for each drug. In contrast, in the ACASI

version, respondents were presented with an individual question on each of the drugs of interest and asked to indicate whether he or she had ever used it. This served to focus their attention on each individual drug.

Taken together, these differences made for both a much more private interviewing environment and one that required increased attention to the response task. One would expect that this would have some impact on reporting, which it did. Exhibit 6.2.1 presents the results.

Overall, lifetime reporting of the nonmedical use of analgesics increased by 300%, from an estimated 4.9% of the population to 14.8%. Both youths and adults showed dramatic increases, with the youths' lifetime prevalence rate being 3.7 times higher under ACASI and the adults' 3 times greater. Use of analgesics during the past 12 months showed a similar dramatic increase. For the other psychotherapeutic drugs, similar but less dramatic results were obtained. For example, overall the reported lifetime prevalence of nonmedical use of tranquilizers was 2.7 times higher under ACASI, 1.8 times higher for stimulants, and 2.5 times higher for sedatives.

Exhibit 6.2.1 Prevalence of Psychotherapeutics, by Interview Mode and Age

	Overall			18+ Years Old			12 to 17 Years	
Mode	Lifetime	Past 12 Month	Past 30 Day	Lifetime	Past 12 Month	Past 30 Day	Lifetime	Past 12 Month
Analgesics								
ACASI	14.8	4.8	1.4	15.1	4.5	1.1	12.5	7.7
PAPI	4.9	1.6	0.4	5.1	1.6	0.4	3.4	2.1
Tranquilizers								
ACASI	8.1	1.3	0.1	8.7	1.4	0.0	2.8	1.2
PAPI	3.0	1.1	0.3	3.2	1.1	0.3	1.8	0.9
Stimulants								
ACASI	8.1	1.2	0.2	8.6	1.1	0.1	3.6	2.0
PAPI	4.5	0.7	0.2	4.8	0.7	0.1	2.1	0.6
Sedatives								
ACASI	4.6	1.1	0.0	4.9	1.2	0.0	1.6	0.7
PAPI	1.9	0.5	0.0	2.0	0.5	0.0	0.7	0.2

6.3 Effect of ACASI on Mental Health Questions

Questions about adult mental health syndromes have been included in the NHSDA instrument since 1994. These questions were interviewer-administered. Consequently, respondents who had experienced symptoms suggestive of mental health disorders in the past 12 months were required to report these problems aloud to the interviewers. Because of stigma or shame associated with mental health problems, some respondents who experienced such problems may have been reluctant to report them to an interviewer, particularly if another household member was within hearing distance. We included the adult mental health questions in the ACASI section in the 1997 field experiment because we felt that increased privacy might facilitate the reporting of potentially sensitive mental health problems.

The mental health questions also included numerous routing patterns, some of which required checking of multiple items to determine if a condition had been met (e.g., existence of at least one affirmative answer to a series of gate agoraphobia questions). Other contingent questioning structures in the mental health section involved follow-up questions to route respondents back into a "branch" if

they initially answered "no" to a gate question but then answered affirmatively to the follow-up question. The field interviewers sometimes made routing errors. Errors were particularly problematic if interviewers incorrectly bypassed mental health questions that respondents should have answered, because some of these respondents might have reported significant mental health problems had they been asked the questions. In contrast, routing errors are virtually nonexistent in a well-tested CAI instrument because the computer program routinely checks the routing conditions and then executes the instructions according to the contingent questioning structure that was set up by the researchers.

In this section, we first compare ACASI mental health data across the eight experimental treatments. We then compare estimates of different mental health syndromes based on the PAPI or ACASI modes of administration. Finally, we examine selected data quality issues.

6.3.1 Comparisons of Rates Within ACASI Treatments

In the 1997 field experiment, adult respondents were asked the same set of mental health questions regardless of the experimental treatment to which they were assigned. All other things being equal, respondents in the different treatments should not vary significantly in how they answered the mental health questions. We did not know, however, whether the number or types of questions posed in a given treatment might affect respondents' subsequent answers in nonexperimental sections. Respondents in a given treatment might be more likely than respondents in other treatments to answer gate mental health questions negatively and to thus bypass the branch questions (i.e., if a given treatment induces more "nay-saying").

To test the assumption that the experimental treatment did not appreciably affect respondents' answers to the mental health questions, we examined the unweighted percentages of adults in each treatment who answered affirmatively to lead questions for the different mental health syndromes covered in the ACASI instrument: major depressive episode, generalized anxiety disorder, agoraphobia, and panic attack. As expected, no clear patterns emerged in terms of the tendency of respondents to answer affirmatively to the gate mental health questions according to their experimental treatment. For example, about 23% to 30% of respondents who were asked the gate depression question answered affirmatively. Similarly, 8% to 14% of respondents in the different treatments who were asked the agoraphobia questions gave an affirmative answer, and 7% to 16% of respondents who were asked the panic attack questions gave an affirmative answer.

One exception was that 28% of respondents in Treatment 4 (single gate question, consistency check present, multiple use questions present) who were asked the gate anxiety question answered affirmatively compared to about 13% to 20% of the respondents in the other treatments who were asked this question. However, these respondents were not consistently more likely than respondents in other treatments to answer affirmatively to the gate questions. For example, respondents in four other treatments were somewhat more likely than respondents in Treatment 4 to answer affirmatively to the gate depression question.

Despite the higher rate of respondents in Treatment 4 who answered affirmatively to the lead anxiety question, these findings generally suggest that the experimental treatment did not affect adults' willingness to answer affirmatively to the lead questions in the mental health section. Therefore, in the remainder of this section, we present ACASI mental health data pooled across all eight treatments.

6.3.2 Comparison of Rates of Mental Health Syndromes

We compared mental health short-form scores for the adult field experiment respondents (n=865) and the comparison group of adults, that is, the Quarter 4 adult PAPI respondents subsetted to English-language interviews in the primary sampling units (PSUs) that corresponded to the field experiment (n=2,126). The PAPI scoring routines that have been used since the 1994 NHSDA create two sets of short-form score variables: one that does not take into account logical editing of the data and one that does take edits into account. For the ACASI/PAPI comparisons, we used the PAPI short-form score variables that do not take editing into account because we did no editing of the ACASI mental health data. To create short-form scores for the ACASI data, we adapted the PAPI scoring routines.

In both the ACASI and PAPI data, a short-form score of "3 or more" for major depressive episode, generalized anxiety disorder, or panic attack indicates that a person is a "probable case" for that mental syndrome. For agoraphobia, a short-form score of "1 or more" indicates a probable case.

Exhibit 6.3.1 shows the results of the mental health syndrome short-form score comparisons between the ACASI and the PAPI comparison group. Exhibit 6.3.2 compares results for ACASI and PAPI by gender and educational level. Exhibit 6.3.3 compares results for the two modes of interviewing by race/ethnicity. All estimates are weighted. We did not test for the statistical significance of any differences that we observed; therefore, we discuss differences that we observed as trends.

We observed a definite trend for ACASI to yield higher estimates of these mental syndromes compared with PAPI. In particular, the estimate of "probable caseness" for major depressive episode based on ACASI (14.6%) was nearly double the rate for PAPI (7.4%). In addition, the ACASI estimate for generalized anxiety disorder (5.8%) was nearly four times the PAPI rate (1.6%). Rates for panic attack and agoraphobia also tended to be higher for the ACASI instrument, although the rates of these two syndromes were generally low. Overall, the PAPI rates for this subset of the 1997 NHSDA were comparable to rates that have been observed in prior NHSDAs (e.g., SAMHSA, 1997).

Although rates of these different mental health syndromes tended to be higher in ACASI for both males and females compared with corresponding PAPI rates, the effects were particularly pronounced for males (see Exhibit 6.3.2). For example, the ACASI estimate of probable caseness for generalized anxiety disorder among males (5.1%) was more than 4.5 times the PAPI rate for males (1.1%). Similarly, ACASI estimated that 12.7% of adult males were probable cases for major depressive episode, or more than twice the PAPI rate of 5.5% for males.

In both ACASI and PAPI, women were more likely than men to score as probable cases for major depressive episode. This is consistent with prior NHSDA findings and with the mental health literature in general (Kessler et al., 1994; SAMHSA, 1996b, 1997). However, the gender differential in the rates for major depressive episode was not as pronounced for ACASI as it was for PAPI. The PAPI estimate of the prevalence of major depressive episode among women was almost twice that for men (9.1% vs. 5.5%, respectively). In comparison, the ACASI estimate for women was only 30% higher than the rate for men (16.4% vs. 12.7%). These findings suggest that the privacy of ACASI may provide respondents with more freedom to report these problems than might be the case if they had to report these problems to an interviewer.

Among the educational groups shown in Exhibit 6.3.2, there was a clear trend for ACASI to yield higher estimates for the different mental health syndromes for adults with some education beyond high school compared with PAPI data for this group. For the other educational groups, there was not a consistent pattern of ACASI yielding higher estimates compared with PAPI. These findings suggest that increased reporting of mental health problems by adults with higher levels of education may be driving the higher estimates in ACASI. Interestingly, PAPI yielded somewhat higher estimates for generalized anxiety disorder among adults who had not finished high school (4.4%) compared with ACASI estimates for this group (2.4%). Similarly, PAPI estimates for agoraphobia were somewhat higher for adults with a high school education or less compared with the corresponding ACASI estimates.

Among the different racial/ethnic groups shown in Exhibit 6.3.3, ACASI tended to yield higher estimates of the different mental health syndromes for whites and blacks relative to the corresponding PAPI estimates for these groups. With the smaller number of adult Hispanic respondents in the field experiment (n=145), estimates for Hispanics showed more variability across the different modes of administration. One notable finding was that blacks (0.8%) had a lower PAPI estimate for generalized anxiety disorder compared with whites (1.7%), but the ACASI rates for both of these groups were considerably higher and were comparable (6.7% vs. 5.9%, respectively). Similarly, the PAPI estimate for panic attack among blacks was lower than the estimate for whites (1.3% vs. 1.9% respectively), but the ACASI rate was higher for blacks than it was for whites (4.4% vs. 3.8%, respectively).

6.3.3 Comparison of Data Quality in the ACASI and PAPI Mental Health Items

We investigated the extent to which the ACASI administration of the mental health items improved data quality relative to PAPI, and whether improvements in data quality for the ACASI instrument might explain the higher rates of mental health problems based on ACASI. In particular, if interviewers administering the PAPI instrument made errors that caused them to skip over key questions used in scoring the mental health variables, respondents who might otherwise have reported significant problems would have missing data for these items. In contrast, we would expect careful testing of the ACASI program to lead to the elimination of these types of errors in ACASI. In addition, the programming of contingent routing instructions in ACASI can reduce or eliminate the occurrence of inconsistencies, such as respondents who report zero occurrences of a problem in the past 12 months (e.g., panic attacks) after having previously reported that they had the problem in that same period.

Exhibit 6.3.4 presents data on missing data patterns in the ACASI and Quarter 4 PAPI comparison group where (a) all items pertaining to a given mental health problem were blank (e.g., due to a breakoff prior to that point in the interview), or (b) the gate question was refused and data within the branch were missing, indicating that these questions within the branch were bypassed. In both the ACASI and PAPI groups, these rates of missing data were low. In particular, these missing percentages based on the PAPI comparison group were lower than the ACASI percentages. These findings suggest that the higher estimates of mental health problems based on ACASI cannot be explained by a tendency for entire sets of PAPI questions to have missing data.

Exhibit 6.3.5 compares selected patterns of inconsistent or ambiguous data in the ACASI or Quarter 4 PAPI comparison samples. These patterns include the following:

1. the reported year when the respondent last had a 2-week period of depression being inconsistent with prior data indicating that the respondent had a 2-week period of depression in the past 12 months;
2. the respondent not knowing or refusing to answer questions about how long the period of anxiety lasted or has been going on;
3. the reported number of months that the respondent had agoraphobia-related fears being inconsistent with the reported length of these fears;
4. zero panic attacks being reported in the past 12 months in question after the respondent had answered affirmatively to one of the gate panic attack questions; and
5. the reported year when the respondent last had a panic attack being inconsistent with prior data indicating that the respondent had a panic attack in the past 12 months.

If respondents reported a 2-week period of feeling depressed or a panic attack in the past 12 months, the ACASI programming did not allow them to report a year other than 1996 or 1997 for the most recent occurrence of these problems. Furthermore, the ACASI program asked the question about the year when these problems last occurred only if the reported month when these problems last occurred matched the interview month (e.g., if a respondent reported last having a panic attack in November and the interview was conducted in November 1997).

Consequently, out-of-range years for the most recent symptoms of depression or panic attack did not occur in the ACASI data. Although these out-of-range years sometimes occurred in the PAPI data, they occurred relatively rarely in the Quarter 4 comparison sample. However, the occurrence of out-of-range years in the PAPI data raises the question of whether the ACASI program should "force" this consistency. Given the ability of CAI to resolve inconsistencies in the course of an interview, an alternative might be to permit a wider range of allowable years and then to prompt ACASI respondents to resolve any inconsistencies. That could involve resolving inconsistencies in the direction of less recent occurrence of depression symptoms or panic attacks.

Information on responses of "don't know" or "refused" to the questions about the duration of anxiety symptoms are presented in Exhibit 6.3.5 because ACASI respondents who answer "don't know" or "refused" to the duration questions are routed out of subsequent questions about symptoms associated with anxiety. All four of the ACASI respondents who had this particular pattern did not know how long they had experienced periods of anxiety or worry. If ACASI respondents are unable to

recall the length of time that they experienced periods of anxiety, we would recommend adding a follow-up probe to determine if the period lasted fewer than 6 months, or 6 months or more, which is the key piece of information sought. Adding this follow-up probe would allow more information to be collected on symptoms associated with anxiety, as opposed to letting these respondents be routed out.

In addition, the ACASI programming specifications did not completely eliminate the potential for inconsistent reporting in the mental health items. Specifically, in the agoraphobia questions, it was possible for ACASI respondents to report a number of months that they had these fears that was inconsistent with the prior question on the length of time that they had these fears. Six ACASI respondents showed this pattern of inconsistent reporting. Therefore, this indicated that ACASI programming needs to be revised either by restricting the allowable range or by prompting respondents to resolve inconsistent answers.

Exhibit 6.3.6 shows selected patterns of skip errors for the Quarter 4 PAPI data only. These errors concerned the following:

1. respondents being incorrectly asked entire sets of related questions twice, when they should have been asked only one set or the other (i.e., both sets of depression questions or both sets of anxiety questions); and
2. respondents not being asked questions that are used in computing mental health scores, when other data suggest that they should have been asked (e.g., if a lead question was answered as "yes" but questions within the branch were blank).

The first type of contingent questioning error is problematic from a burden perspective in that an entire set of questions is essentially repeated. This unnecessarily increases the interview time, can add to respondent fatigue, and can result in breakoffs if respondents become angry about being re-asked questions that they already answered. The second type of routing error is problematic from a measurement perspective in that some respondents who might otherwise have reported significant problems would not be asked the relevant questions.

As shown in Exhibit 6.3.6, about 1% of the PAPI respondents (i.e., unweighted) in the Quarter 4 comparison group were incorrectly asked both sets of depression questions. Instead of being routed to the start of the anxiety section, these respondents were asked a second set of depression questions. Although relatively few respondents in the Quarter 4 PAPI comparison group were asked both sets of depression questions (n=26), a rate of 1.2% translates to more than 900 respondents in a sample of more than 75,000 interviews, which was the number planned for the 1999 NHSDA.

No respondents in the Quarter 4 PAPI comparison group had an affirmative answer to a lead depression question, with the questions being blank within the branch pattern. Similarly, rates of contingent questioning errors that affected the short-form score variables were low for the anxiety and agoraphobia sections of the interview. The rate of contingent questioning errors increased for the panic attack section, approaching 1% of respondents in the Quarter 4 PAPI comparison group (n=17). Compared with the other mental health sections, the panic attack section had more routing patterns that would provide opportunities for interviewers to exit the section prematurely.

However, these findings in Exhibit 6.3.6 suggest that contingent questioning errors in the PAPI instrument that could affect scoring of the mental health syndromes were relatively uncommon, at least in the fourth quarter among respondents completing an English-language interview. However, one caution is that we might expect error rates in the PAPI mental health data to be low by the fourth quarter due to interviewer experience. Most of the productive, high-quality interviewers working on the NHSDA by the start of the fourth quarter had three or more quarters of experience with the NHSDA interview by that time.

Nevertheless, these findings suggest that contingent questioning errors on the part of the PAPI interviewers are not likely to explain the considerably lower rates of some mental health syndromes in the PAPI data. Rather, the findings from these analyses of data quality issues suggest that the higher

estimates of mental health syndromes based on ACASI reflect a greater tendency of ACASI respondents to report mental health problems, particularly among males and adults with higher levels of education.

Exhibit 6.3.1 Adult Mental Health Short-Form Scores		
	1997 Field Experiment	**Comparison Group**
Score	CAPI/ACASI	1997 Q4 PAPI
Total 18+	**(*n*=865)**	**(*n*=2,126)**
Major Depressive Episode Short-Form Score		
0	85.2	91.9
1 or 2	0.1	0.7
3 or more	14.6	7.4
Generalized Anxiety Disorder Short-Form Score		
0	94.2	98.3
1 or 2	0.0	0.1
3 or more	5.8	1.6
Agoraphobia Short-Form Score		
0	98.1	98.4
1 or more	1.9	1.6
Panic Attack Short-Form Score		
0	92.5	96.4
1 or 2	3.6	1.5
3 or more	3.9	2.0

Note: Percentages are weighted. PAPI short-form scores do not take into account logically imputed data. A score of "3 or more" indicates a probable case for major depressive episode, generalized anxiety disorder, or panic attack. A score of "1 or more" indicates a "probable case" for agoraphobia. Some table entries may not sum to 100% because of rounding.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.3.2 Mental Health Syndromes, by Gender and Education

Syndrome	Gender	1997 Field Experiment CAPI/ACASI	Comparison Group 1997 Q4 PAPI	Education	1997 Field Experiment CAPI/ACASI	C 19
Unweighted Numbers	**Total 18+**	**(*n*=865)**	**(*n*=2,126)**	**Total 18+**	**(*n*=865)**	
of Respondents	Male	(*n*=355)	(*n*=809)	< High School	(*n*=199)	
	Female	(*n*=510)	(*n*=1,317)	High School	(*n*=321)	
				> High School	(*n*=345)	
Probable Case for Major	**Total 18+**	**14.6**	**7.4**	**Total 18+**	**14.6**	
Depressive Episode	Male	12.7	5.5	< High School	16.7	
	Female	16.4	9.1	High School	6.9	
				> High School	19.2	
Probable Case for Generalized	**Total 18+**	**5.8**	**1.6**	**Total 18+**	**5.8**	
Anxiety Disorder	Male	5.1	1.1	< High School	2.4	
	Female	6.5	2.1	High School	3.1	
				> High School	8.6	
Probable Case for	**Total 18+**	**1.9**	**1.6**	**Total 18+**	**1.9**	
Agoraphobia	Male	1.3	0.4	< High School	2.5	
	Female	2.4	2.7	High School	1.4	
				> High School	2.0	
Probable Case for	**Total 18+**	**3.9**	**2.0**	**Total 18+**	**3.9**	
Panic Attack	Male	1.8	1.0	< High School	4.3	
	Female	5.9	3.0	High School	4.6	
				> High School	3.4	

Note: PAPI short-form scores do not take into account logically imputed data.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Exper
National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.3.3 Mental Health Syndromes, by Race/Ethnicity

		1997 Field Experiment	Comparison Group
Syndrome	**Race/Ethnicity**	**CAPI/ACASI**	**1997 Q4 PAPI**
Unweighted Numbers of Respondents	Total 18+	(*n*=865)	(*n*=2,126)
	Hispanic	(*n*=145)	(*n*=347)
	Non-Hisp., Black	(*n*=283)	(*n*=748)
	Non-Hisp., All Other Races	(*n*=437)	(*n*=1,031)
Probable Case for Major Depressive	Total 18+	14.6	7.4

Episode			
	Hispanic	12.7	3.2
	Non-Hisp., Black	7.5	3.9
	Non-Hisp., All Other Races	16.0	8.3
Probable Case for Generalized Anxiety Disorder	Total 18+	5.8	1.6
	Hispanic	2.7	2.3
	Non-Hisp., Black	6.7	0.8
	Non-Hisp., All Other Races	5.9	1.7
Probable Case for Agoraphobia	Total 18+	1.9	1.6
	Hispanic	4.0	1.3
	Non-Hisp., Black	2.4	1.9
	Non-Hisp., All Other Races	1.6	1.6
Probable Case for Panic Attack	Total 18+	3.9	2.0
	Hispanic	4.1	4.5
	Non-Hisp., Black	4.4	1.3
	Non-Hisp., All Other Races	3.8	1.9

Note: PAPI short-form scores do not take into account logically imputed data.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.3.4 Selected Missing Data Patterns for Adult Mental Health Questions

	1997 Field Experiment			Comparison Gr	
	CAPI/ACASI			1997 Q4 PAP	
Syndrome	Unweighted	Unweighted	Weighted %	Unweighted	Unweighted
Total 18+		(*n*=865)			(*n*=2,126)
Major Depressive Episode					
All blank	10	1.2%	0.8%	7	0.3%
Lead question refused; no data within branch	2	0.2%	0.1%	0	0.0%
Generalized Anxiety Disorder					
All blank	11	1.3%	0.9%	10	0.5%
Lead question refused; no data within branch	2	0.2%	0.1%	0	0.0%
Agoraphobia					
All blank	11	1.3%	0.9%	10	0.5%
Lead questions all refused; no data within branch	2	0.2%	0.1%	0	0.0%
Panic Attack					
All blank	11	1.3%	0.9%	7	0.3%
Either lead question refused; no data within branch	3	0.3%	0.4%	0	0.0%

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experim
National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.3.5 Inconsistent or Ambiguous Data in the Adult Mental Health Questions

	1997 Field Experiment			Comparison Group	
	CAPI/ACASI			1997 Q4 PAPI (Edite	
Syndrome	Unweighted *n*	Unweighted %	Weighted %	Unweighted *n*	Unweighted %
Total 18+	(*n*=865)			(*n*=2,126)	
Major Depressive Episode					
Out-of-range year	NA	NA	NA	7	0.3%
Generalized Anxiety Disorder					
Duration of anxiety unknown/refused	4	0.5%	0.3%	1	0.0%
Agoraphobia					
Inconsistent length of fears	6	0.7%	0.3%	0	0.0%
Panic Attack					
Zero attacks reported in past year (and prior affirmative answer)	NA	NA	NA	0	0.0%
Out-of-range year	NA	NA	NA	1	0.0%

NA = Not applicable; ACASI programming specifications do not allow this type of response to occur.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 F
Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.3.6 Skip Error Patterns in the PAPI Adult Mental Health Questions

	1997 Q4 PAPI (Edited)		
Syndrome	Unweighted *n*	Unweighted %	Weighte
Total 18+	(*n*=2,126)		
Depression			
Short-form score items incorrectly skipped	0	0.0%	0.0%
Both sets of depression questions incorrectly administered	26	1.2%	1.4%
Generalized Anxiety Disorder			
Short-form score items incorrectly skipped	6	0.3%	0.3%
Both sets of anxiety questions incorrectly administered	2	0.1%	0.0%
Agoraphobia			
Short-form score items incorrectly skipped	3	0.1%	0.0%
Panic Attack			
Short-form score items incorrectly skipped	17	0.8%	0.7%

Source: 1997 National Household Survey on Drug Abuse; Quarter 4.

6.4 Effect of Audio Computer-Assisted Self-Interviewing on Responses to Questions on Risk and Availability

The perceived risk of using drugs was assessed by asking how much people risk harming themselves when they engage in each of 11 different drug-using behaviors. The response choices offered were "no risk," "slight risk," "moderate risk," and "great risk." Respondents were also asked about how difficult or easy it would be to get each of five drugs. Response options were "probably impossible," "very difficult," "fairly difficult," "fairly easy," and "very easy."

6.4.1 Perceived Risk of Different Drug-Using Behaviors

Questions about levels of risk perceived to be associated with using different drugs were included in every NHSDA since 1985 except for the 1995 NHSDA. The major purpose for asking these questions has been to quantify the national population's knowledge about risks of harm, physical and otherwise, that might result from using drugs. Such information has been useful for guiding substance abuse prevention and education efforts. The questions used in the 1997 NHSDA constitute a set that have been consistently asked, although there have been some improvements in the frequency of use wordings. The risk questions were presented on their own answer sheet using the NHSDA self-administered questionnaire (SAQ).

Exhibit 6.4.1 shows the percentages (weighted) of the ACASI and PAPI/SAQ total populations and the populations within specific age groups, as well as current cigarette smokers and those who used any illicit drug in the past year who reported perceptions of great risk or other harm associated with different drug-using behaviors. For the total population and the two age groups, the level of risk was perceived as great by more respondents to the PAPI/SAQ instrument than by ACASI respondents. The only noteworthy exceptions were youths aged 12 to 17 who answered with ACASI. Fewer of them than PAPI/SAQ youths perceived that smoking one or more packs of cigarettes and smoking marijuana once a month entailed great risk. Differences between the methodologies in percentages of the total population and the two age groups who reported great risk ranged from 0% to about 6% for most of

the drug-using behaviors. For the two drinking behaviors, however, differences between the two methodologies were greater, with at least 7% to 10% more PAPI/SAQ respondents reporting perceptions of great risk.

Perceptions that great risk is associated with different drug-using behaviors varied with drug use. In Exhibit 6.4.1, perceptions about risk associated with the different drug-using behaviors are reported for current cigarette smokers and for persons who used any illicit drug in the past year. Almost without exception, fewer smokers and fewer illicit drug users reported that the different drug-using behaviors entailed great risk than did the total populations of the two age groups and of all ages. In a few instances, drug users rated an activity as more risky: (a) current cigarette smokers of all ages when they answered for trying heroin once or twice and (b) smokers aged 12 to 17 when they answered questions on using heroin and cocaine. In addition, more 12- to 17-year-old past year illicit drug users reported great risk for heroin and cocaine use. This pattern of more (usually younger) smokers and past year illicit drug users who reported great risk of using heroin and cocaine appears in both the ACASI and the PAPI/SAQ data and is more pronounced under ACASI.

In general, past year illicit drug users were less likely than the overall population to perceive that smoking marijuana entailed great risk. The ACASI respondents who were classified as past year drug users were somewhat more likely than the PAPI/SAQ respondents to rate monthly marijuana use as harmful.

Differences between the ACASI and PAPI/SAQ instruments may be contributing to the general pattern of ACASI percentages reporting great risk being smaller than the PAPI/SAQ percentages. On the PAPI/SAQ answer sheet, the boxes for indicating one's response are all printed on the same line as the text of the drug-using behavior, with the actual text of the response options printed only once on the page as headings above columns of boxes and ordered left to right from "no risk" to "great risk." In ACASI, each response option appears, along with its numbers to be keyed to indicate one's choice, in a list below the text of the general question's introductory phrase and the phrase specifying the drug-using behavior. The response options are identical to those on the PAPI/SAQ answer sheet and are ordered top to bottom from "no risk" to "great risk." Respondents to the PAPI/SAQ risk questions, most of whom were likely to have been right-handed, may have been somewhat more likely to mark the boxes physically located on the right; if so, reports about the risk associated with the drug-using behaviors will be biased toward greater risk. On the other hand, respondents to the ACASI risk questions may have been somewhat more likely to key in the number for a response option printed higher in the list and therefore encountered more quickly than they were to read the entire list or even further down the list before keying an answer. There is considerable evidence that, for response options similar in content to those in the risk questions (e.g., rankings of satisfaction, importance), those printed closer to the top of a list of options will be chosen more often than items printed closer to the end of the list, even when the ordering of the listing is reversed. If this tendency is occurring, ACASI reports about risk of drug-using behaviors will be biased toward lower levels of risk.

Exhibit 6.4.2 presents the (unweighted) percentages of the ACASI and PAPI/SAQ total populations and the populations of specific age groups who failed to answer each question about how much risk they thought each drug-using behavior exposed one to. The missing data types "refused" and "no answer (blank)" were combined in Exhibit 6.4.2 because PAPI/SAQ respondents were not required to indicate their refusal to answer any particular question or even the entire set. So some of the "no answer (blank)" missing data in the PAPI/SAQ data may actually reflect any such refusals.

The amount of missing data was small for both data collection methodologies, never exceeding 4.5% for any risk question and occurring much less often for most of the risk questions. Most ACASI missing data resulted from "don't know" responses; most of the PAPI/SAQ missing data resulted from "no answer (blank)."

More data are "missing" in the ACASI mostly because more ACASI respondents answered "don't know," an option not explicitly offered in the PAPI/SAQ. PAPI/SAQ respondents apparently either tried harder to respond with a risk answer or simply left the item blank. In PAPI/SAQ, the generally higher rates of "no answer (blank)" missing data were still very small (never exceeding 1.7% and usually no

greater than 0.7%), suggesting that most uncertain respondents tried to give a risk answer. In ACASI, however, the explicit availability of "don't know" and "refused" response options may have made it easier for respondents not to provide risk answers, resulting in higher rates of missing data of both kinds, especially the "don't know" type. If the goal was to reduce the absolute amount of missing data, PAPI/SAQ did a better job; however, enabling respondents who were very uncertain about the risk associated with a drug-using behavior to register that uncertainty may actually result in more valid data.

6.4.2 Availability of Drugs

Exhibit 6.4.3 presents percentages (weighted) of the ACASI and PAPI/SAQ total populations, current cigarette smokers, and past year illicit drug users who reported that five different drugs were fairly or very easy to get. The overall pattern of responses was for more PAPI/SAQ than ACASI respondents to report that getting each drug would be easy, although the differences between the two modes were usually smaller than in the risk data. Indeed, for all ages and the two age-group total populations, differences between the ACASI and PAPI/SAQ percentages reporting that getting the drugs would be fairly or very easy were unremarkable, except that about 15% more PAPI/SAQ than ACASI respondents reported that getting marijuana was easy. For both modes, subgroups who used cigarettes or drugs were more likely to report that getting each drug would be fairly or very easy. This last finding is consistent with our expectation that persons who use drugs, even if only cigarettes, are more likely to feel as though they can get drugs without too much difficulty, while persons who use no drugs are more likely to feel that getting drugs would be difficult.

This constancy of the pattern of differences between the two modes points toward a systematic mode effect. Being able to see all the items and availability response options on the PAPI/SAQ answer sheet may encourage more variation in responses with a net effect of shifting the average response toward the "fairly easy" or "very easy" options. This factor does not seem sufficient by itself, however, to cause the regular and consistent pattern in the differences between the ACASI and the PAPI/SAQ availability responses.

Exhibit 6.4.4 presents (unweighted) percentages of the ACASI and PAPI/SAQ total populations and the populations in specific age groups who did not provide a response to availability questions. The missing data types "refused" and "no answer (blank)"were combined because of the impossibility of distinguishing between the two types of missing data for PAPI/SAQ respondents.

The amounts of missing data in both the ACASI and PAPI/SAQ data for the total samples and for the subgroups defined by age ranged from 2.1% to 6.8% of the ACASI subsamples, with most of the rates exceeding 4% and from 0.5% to 5.8% of the PAPI/SAQ subsamples with most of the rates falling below 2.2%. As in the risk data, most ACASI missing data consisted of "don't know" responses, while most of the PAPI/SAQ missing data resulted from "no answer (blank)."

Again, rates of missing data in ACASI were higher than in PAPI/SAQ because of the "don't know" responses. The ready availability of this response for those who used ACASI may mean that they are truly uncertain of the answer.

Exhibit 6.4.1 Weighted Percentages Reporting Perceptions of Great Risk of Physical or Other H Associated with Different Drug-Using Behaviors, by Drug Use Status and Age

Respondent Characteristics		1997 Field Experiment	Comparison Group	1997 Field Experiment	Comparison Group	1997 Field Experiment	C
Age Group	Drug Use	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ	CAPI/ACASI	
		a. Smoking 1 or More Packs of Cigarettes a Day		**b. Smoking Marijuana Once a Month**		**c. Smoking Marijuana a Week**	
All Ages	**Total**	**63.3%**	**67.9%**	**40.5%**	**42.0%**	**53.1%**	
	Current Cigarette Smokers	43.7%	51.2%	27.3%	30.4%	36.9%	
	Used Any Illicit Drug Past Year	62.8%	64.5%	14.2%	11.3%	18.5%	
12-17	**Total**	**56.4%**	**53.1%**	**36.5%**	**32.7%**	**50.8%**	
	Current Cigarette Smokers	42.0%	35.4%	15.8%	16.6%	16.4%	
	Used Any Illicit Drug Past Year	47.1%	39.5%	13.3%	8.5%	22.0%	
18+	**Total**	**64.1%**	**69.7%**	**41.0%**	**43.1%**	**53.4%**	
	Current Cigarette Smokers	43.8%	52.3%	28.1%	31.3%	38.3%	
	Used Any Illicit Drug Past Year	65.8%	68.5%	14.4%	11.7%	17.8%	
		d. Trying LSD Once or Twice		**e. Using LSD Once or Twice a Week**		**f. Trying Heroin On**	
All Ages	**Total**	**70.1%**	**73.6%**	**88.5%**	**89.0%**	**79.3%**	
	Current Cigarette Smokers	71.9%	67.4%	86.5%	84.3%	84.8%	
	Used Any Illicit Drug Past Year	45.1%	48.1%	76.9%	79.3%	73.8%	
12-17	**Total**	**50.3%**	**49.0%**	**71.3%**	**76.4%**	**54.4%**	
	Current Cigarette Smokers	45.0%	43.8%	66.4%	70.9%	61.9%	
	Used Any Illicit Drug Past Year	50.4%	37.6%	71.3%	77.2%	61.5%	
18+	**Total**	**72.4%**	**76.5%**	**90.5%**	**90.5%**	**82.2%**	
	Current Cigarette Smokers	73.8%	69.0%	87.9%	85.2%	86.4%	
	Used Any Illicit Drug Past Year	44.1%	49.8%	78.0%	79.6%	76.1%	

Exhibit 6.4.1 (continued)

Respondent Characteristics		1997 Field Experiment	Comparison Group	1997 Field Experiment	Comparison Group	1997 Field Experiment	C
Age Group	Drug Use	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ	CAPI/ACASI	
		g. Using Heroin Once or Twice a Week		**h. Using Cocaine Once a Month**		**i. Using Cocaine On Week**	
All Ages	**Total**	**92.7%**	**93.2%**	**73.2%**	**77.4%**	**90.4%**	
	Current Cigarette Smokers	91.5%	92.5%	68.9%	74.8%	84.1%	
	Used Any Illicit Drug Past Year	90.9%	91.5%	53.8%	61.3%	80.5%	
12-17	**Total**	**78.2%**	**82.1%**	**49.4%**	**55.3%**	**76.4%**	
	Current Cigarette Smokers	79.9%	82.6%	53.7%	57.1%	78.1%	
	Used Any Illicit Drug Past Year	83.4%	87.3%	51.6%	52.7%	76.0%	
18+	**Total**	**94.4%**	**94.4%**	**76.0%**	**79.9%**	**92.0%**	
	Current Cigarette Smokers	92.3%	93.1%	70.0%	76.0%	84.5%	
	Used Any Illicit Drug Past Year	92.3%	92.2%	54.2%	62.6%	81.4%	
		j. Having 4 or 5 Drinks Nearly Every Day		**k. Having 5 or More Drinks Once or Twice a Week**			
All Ages	**Total**	**70.3%**	**77.1%**	**45.7%**	**55.6%**		
	Current Cigarette Smokers	57.0%	67.8%	32.8%	47.7%		
	Used Any Illicit Drug Past Year	58.2%	64.2%	29.0%	33.7%		
12-17	**Total**	**58.9%**	**67.0%**	**39.0%**	**47.9%**		
	Current Cigarette Smokers	46.2%	52.7%	30.0%	28.6%		
	Used Any Illicit Drug Past Year	44.1%	52.8%	24.3%	24.9%		
18+	**Total**	**71.7%**	**78.3%**	**46.5%**	**56.5%**		
	Current Cigarette Smokers	57.7%	68.9%	33.0%	49.0%		
	Used Any Illicit Drug Past Year	60.9%	66.1%	29.9%	35.1%		

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures, 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.4.2 Unweighted Missing Data in Questions About Perceptions of Risk of Physical or Other Harm Associated with Different Drug-Using Behaviors for the Total Population, by Age Group

Respondent Characteristics		1997 Field Experiment	Comparison Group	1997 Field Experiment	Comparison Group	1997 Field Experiment	Comparison Group
Age Group	Type of Missing Data	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ
		a. Smoking 1 or More Packs of Cigarettes a Day		b. Smoking Marijuana Once a Month		c. Smoking Marijuana Once or Twice a Week	
All Ages	**Total**	**1.6%**	**0.5%**	**2.0%**	**0.7%**	**2.2%**	**1.9%**
	Don't Know	1.2%	0.1%	1.5%	0.1%	1.7%	0.2%
	Multiple Response	0.0%	0.0%	0.0%	0.1%	0.0%	0.1%
	Refusal, No Answer (Blank)	0.5%	0.3%	0.5%	0.5%	0.5%	1.6%
12-17	**Total**	**1.6%**	**0.1%**	**1.3%**	**0.2%**	**1.7%**	**1.9%**
	Don't Know	1.3%	0.0%	0.9%	0.0%	1.3%	0.0%
	Multiple Response	0.0%	0.0%	0.0%	0.1%	0.0%	0.2%
	Refusal, No Answer (Blank)	0.4%	0.1%	0.4%	0.1%	0.4%	1.7%
18+	**Total**	**1.6%**	**0.6%**	**2.9%**	**1.0%**	**2.9%**	**1.9%**
	Don't Know	1.0%	0.1%	2.2%	0.2%	2.2%	0.3%
	Multiple Response	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.6%	0.4%	0.7%	0.7%	0.7%	1.6%
		d. Trying LSD Once or Twice		e. Using LSD Once or Twice a Week		f. Trying Heroin Once or Twice	
All Ages	**Total**	**3.7%**	**0.9%**	**3.5%**	**1.4%**	**2.7%**	**0.8%**
	Don't Know	3.1%	0.3%	2.9%	0.3%	2.1%	0.2%
	Multiple Response	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.6%	0.6%	0.6%	1.1%	0.6%	0.6%
12-17	**Total**	**4.5%**	**0.9%**	**4.0%**	**1.4%**	**2.5%**	**0.5%**
	Don't Know	3.8%	0.3%	3.5%	0.2%	2.0%	0.0%
	Multiple Response	0.0%	0.2%	0.0%	0.1%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.6%	0.4%	0.5%	1.1%	0.5%	0.5%
18+	**Total**	**2.8%**	**0.9%**	**2.8%**	**1.4%**	**2.9%**	**1.0%**
	Don't Know	2.2%	0.3%	2.2%	0.3%	2.2%	0.3%
	Multiple Response	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.6%	0.7%	0.6%	1.1%	0.7%	0.7%

(continued)

Exhibit 6.4.2 (continued)

Respondent Characteristics		1997 Field Experiment	Comparison Group	1997 Field Experiment	Comparison Group	1997 Field Experiment	Comparison Group
Age Group	Type of Missing Data	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ
		g. Using Heroin Once or Twice a Week		**H. Using Cocaine Once a Month**		**i. Using Cocaine Once or Twice a Week**	
All Ages	**Total**	**2.5%**	**1.3%**	**2.1%**	**0.7%**	**2.1%**	**1.1%**
	Don't Know	1.9%	0.2%	1.6%	0.2%	1.7%	0.2%
	Multiple Response	0.0%	0.1%	0.0%	0.1%	0.0%	0.1%
	Refusal, No Answer (Blank)	0.6%	1.0%	0.5%	0.5%	0.5%	0.9%
12-17	**Total**	**2.3%**	**0.9%**	**1.7%**	**0.3%**	**1.8%**	**0.9%**
	Don't Know	1.8%	0.0%	1.3%	0.0%	1.4%	0.0%
	Multiple Response	0.0%	0.2%	0.0%	0.2%	0.0%	0.2%
	Refusal, No Answer (Blank)	0.5%	0.7%	0.4%	0.1%	0.4%	0.7%
18+	**Total**	**2.8%**	**1.4%**	**2.5%**	**0.9%**	**2.5%**	**1.2%**
	Don't Know	2.1%	0.3%	1.8%	0.3%	2.0%	0.3%
	Multiple Response	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.7%	1.1%	1.3%	0.6%	0.6%	0.9%
		j. Having 4 or 5 Drinks Nearly Every Day		**k. Having 5 or More Drinks Once or Twice a Week**		***Unweighted n's (Denominators)***	***Unweighted n's (Denominators)***
All Ages	**Total**	**1.6%**	**0.9%**	**1.7%**	**0.8%**	***1,982***	***3,105***
	Don't Know	1.2%	0.2%	1.2%	0.2%	*1,982*	*3,105*
	Multiple Response	0.0%	0.1%	0.0%	0.1%	*1,982*	*3,105*
	Refusal, No Answer (Blank)	0.5%	0.6%	0.5%	0.5%	*1,982*	*3,105*
12-17	**Total**	**1.5%**	**0.5%**	**1.4%**	**0.3%**	***1,117***	***979***
	Don't Know	1.2%	0.0%	1.0%	0.0%	*1,117*	*979*
	Multiple Response	0.0%	0.1%	0.0%	0.0%	*1,117*	*979*
	Refusal, No Answer (Blank)	0.4%	0.4%	0.4%	0.3%	*1,117*	*979*
18+	**Total**	**1.7%**	**1.1%**	**2.0%**	**1.0%**	***865***	***2,126***
	Don't Know	1.2%	0.2%	1.4%	0.3%	*865*	*2,126*
	Multiple Response	0.0%	0.1%	0.0%	0.1%	*865*	*2,126*
	Refusal, No Answer (Blank)	0.6%	0.7%	0.6%	0.7%	*865*	*2,126*

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures, 1997 Field Experiment.
1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.4.3 Weighted Percentages Reporting Perceptions That Getting Different Drugs Is Fairly or Very Easy, by Drug Use Status and Age

Respondent Characteristics		1997 Field Experiment	Comparison Group	1997 Field Experiment	Comparison Group
Age Group	Drug Use	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ	CAPI/ACASI	1997 Q4 NHSDA: PAPI/SAQ
		a. Fairly or Very Easy to Get Some Marijuana		d. Fairly or Very Easy to Get Some "Crack"	
All Ages	**Total**	**31.3%**	**36.5%**	**20.3%**	**21.1%**
	Current Cigarette Smokers	39.8%	48.3%	21.1%	28.3%
	Used Any Illicit Drug Past Year	48.6%	54.1%	20.9%	30.4%
12-17	**Total**	**34.5%**	**36.3%**	**16.5%**	**13.5%**
	Current Cigarette Smokers	62.5%	58.2%	17.6%	18.0%
	Used Any Illicit Drug Past Year	59.9%	66.8%	22.8%	19.0%
18+	**Total**	**30.9%**	**36.7%**	**20.8%**	**21.9%**
	Current Cigarette Smokers	38.2%	47.6%	21.3%	29.1%
	Used Any Illicit Drug Past Year	46.4%	52.1%	20.6%	32.2%
		b. Fairly or Very Easy to Get Some LSD		e. Fairly or Very Easy to Get Some Heroin	
All Ages	**Total**	**13.4%**	**16.3%**	**13.3%**	**16.7%**
	Current Cigarette Smokers	12.0%	20.3%	14.7%	21.4%
	Used Any Illicit Drug Past Year	12.3%	20.2%	9.7%	16.3%
12-17	**Total**	**12.4%**	**12.9%**	**10.0%**	**10.8%**
	Current Cigarette Smokers	15.6%	19.9%	10.2%	14.2%
	Used Any Illicit Drug Past Year	16.8%	30.4%	10.0%	13.6%
18+	**Total**	**13.5%**	**16.7%**	**13.7%**	**17.4%**
	Current Cigarette Smokers	11.7%	20.3%	15.0%	21.9%
	Used Any Illicit Drug Past Year	11.5%	18.6%	9.7%	16.8%
		c. Fairly or Very Easy to Get Some Cocaine			
All Ages	**Total**	**18.3%**	**20.6%**		
	Current Cigarette Smokers	19.1%	24.4%		
	Used Any Illicit Drug Past Year	18.5%	27.1%		
12-17	**Total**	**14.2%**	**15.0%**		
	Current Cigarette Smokers	14.2%	20.7%		
	Used Any Illicit Drug Past Year	17.3%	23.9%		
18+	**Total**	**18.7%**	**21.3%**		
	Current Cigarette Smokers	19.4%	28.2%		
	Used Any Illicit Drug Past Year	18.8%	27.7%		

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures, 1997 Field Experiment.
1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.4.4 Unweighted Missing Data in Questions About How Difficult or Easy Getting Different Drugs Is for the Total Population, by Age Group

Respondent Characteristics		1997 Field Experiment	Comparison Group	1997 Field Experiment	Comparison Group
Age Group	**Type of Missing Data**	**CAPI/ACASI**	**1997 Q4 NHSDA: PAPI/SAQ**	**CAPI/ACASI**	**1997 Q4 NHSDA: PAPI/SAQ**
		a. Availability of Marijuana		**d. Availability of "Crack"**	
All Ages	Total	3.0%	1.4%	4.5%	1.8%

	Don't Know	2.4%	0.6%	3.8%	0.8%
	Multiple Response	0.0%	0.0%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.7%	0.8%	0.7%	1.0%
12-17	Total	2.1%	0.6%	3.8%	1.0%
	Don't Know	1.6%	0.1%	3.1%	0.2%
	Multiple Response	0.0%	0.0%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.5%	0.5%	0.6%	0.8%
18+	Total	4.2%	1.8%	5.5%	2.2%
	Don't Know	3.4%	0.8%	4.7%	1.1%
	Multiple Response	0.0%	0.0%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.8%	0.9%	0.8%	1.1%

		b. Availability of LSD		**e. Availability of Heroin**	
All Ages	Total	5.6%	2.2%	4.9%	2.8%
	Don't Know	4.9%	0.9%	4.2%	1.3%
	Multiple Response	0.0%	0.1%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.7%	1.2%	0.8%	1.5%
12-17	Total	4.7%	1.4%	3.6%	0.5%
	Don't Know	3.9%	0.3%	2.9%	0.2%
	Multiple Response	0.0%	0.2%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.7%	0.9%	0.7%	0.4%
18+	Total	6.8%	2.5%	6.7%	5.8%
	Don't Know	6.1%	1.2%	5.9%	2.8%
	Multiple Response	0.0%	0.1%	0.0%	0.0%
	Refusal, No Answer (Blank)	0.7%	1.3%	0.8%	3.0%

		c. Availability of Cocaine		***Unweighted n's (Denominators)***	***Unweighted n's (Denominators)***
All Ages	Total	4.4%	1.8%	*1,982*	*3,105*
	Don't Know	3.6%	0.8%	*1,982*	*3,105*
	Multiple Response	0.0%	0.1%	*1,982*	*3,105*
	Refusal, No Answer (Blank)	0.8%	0.9%	*1,982*	*3,105*
12-17	Total	3.5%	1.0%	*1,117*	*979*
	Don't Know	2.7%	0.2%	*1,117*	*979*
	Multiple Response	0.0%	0.2%	*1,117*	*979*
	Refusal, No Answer (Blank)	0.8%	0.6%	*1,117*	*979*

18+	Total	5.5%	2.2%	*865*	*2,126*
	Don't Know	4.7%	1.1%	*865*	*2,126*
	Multiple Response	0.0%	0.0%	*865*	*2,126*
	Refusal, No Answer (Blank)	0.8%	1.1%	*865*	*2,126*

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures, 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

6.5 Effect of New ACASI Strategy on Frequency of Use

Each of the treatments employed the new method of asking about the frequency of use in the past 12 months.[13] Our primary interest in analyzing the data from the revised 12-month frequency-of-use items was to determine whether the new format affected the quality of the data collected. To investigate this, we conducted two types of analyses. First, we investigated the rates of missing data for the gate question into each 12-month frequency-of-use series and compared the rates to the Quarter 4 PAPI data collection version. Second, we calculated the 12-month frequency of use for each respondent by multiplying the reporting units by the frequency of use to create a continuous variable. We then created categories similar to those used in the PAPI questionnaire and compared the distribution of responses to those collected in the Quarter 4 PAPI interview. Each of these analyses is presented below.

6.5.1 Missing Data Rates

By design, CAI respondents cannot mistakenly bypass an item or leave an item blank. The CAI program requires an answer to be entered for every question. However, respondents are still free to provide a "don't know" or "refused" response in the CAI instrument. In fact, they may be more likely to use these responses in the CAI interview because they have been specifically trained in the use of the "don't know" key and both the "don't know" and "refused" keys are explicitly labeled on the keyboard. This differs from the PAPI environment where these responses do not appear on the answer sheet and the respondent is not explicitly told how to record one of these answers. A respondent who wished to provide one of these responses in the self-administered portion of the PAPI interview had to ask the interviewer how to record the answer,[14] come up with a method on his or her own, or simply leave the item blank.

Because the process for providing the 12-month frequency-of-use data differs between CAI and PAPI, it is somewhat difficult to compare the missing data rates. For our purposes here, we chose to examine the data in two ways. We first compared the "don't know" and "refused" from the CAI interview to the explicit "don't know" and "refused" responses from the PAPI data. We found that there were no explicit "don't know" or "refused" responses for any of the 12-month frequency-of-use items in the PAPI data. In Exhibit 6.5.1, we present data that adds all the "bad data" into the PAPI data. These include multiple responses, inappropriately answered questions, and blanks because it seems likely that questions left blank could reflect an inability or unwillingness to respond on the part of the respondent. In an effort to maintain comparability between the two modes, we include only the data provided during the initial pass through the 12-month frequency-of-use item for respondents assigned to the consistency check treatment.

The data in Exhibit 6.5.1 show that for the alcohol and marijuana sections, the CAI data were more complete than the PAPI data. In fact, for the alcohol section, the bad data rate was more than twice that for CAI (7.9% vs. 3.5%). However, for all other drugs types except hallucinogens, the "don't know" and "refused" rates for CAI were still higher than for all the bad data associated with the PAPI item. For hallucinogens, the rates were comparable. Comparing the CAI and PAPI data in this manner, however, is somewhat artificial. For the CAI data, the "don't know" and "refused" rates were calculated based on only those respondents who were routed to the question (users of the drug). By

contrast, all respondents were included in the calculation for the PAPI data because every respondent answered every question regardless of use status. The result is that the CAI rates may appear artificially high when compared to the PAPI data. Nevertheless, the comparison may provide some insight into the quality of the data collected for the 12-month frequency item.

6.5.2 Distribution of the Responses

To compare the distribution of responses collected in each mode, we collapsed the data into two categories: (a) use for 50 days or less and (b) use for 51 days or more. Then we compared the distribution of responses. These analyses excluded those respondents who indicated no use during the 12-month period.

Exhibit 6.5.2 shows the weighted estimates for youths and adults separately. We see that the younger respondents consistently reported higher frequencies of use in CAI than in PAPI. This trend was not as strong for adult respondents, however, suggesting that youths were more influenced by question wording.

These findings, although intriguing, are somewhat difficult to explain. It is possible that the increased frequency of use reported in CAI is due to the revised question wording. Because respondents were generating their own responses without being able to see the other response choices (as is the case on the answer sheets used in PAPI), they may have been less influenced by the desire not to appear a "serious user" or trying to appear "like everyone else." The ability to influence frequency reporting by altering the endpoints of the scale has been well documented in the literature (see, e.g., Schwarz & Hippler, 1991). A similar phenomenon may be taking place here, in that respondents' answers were more accurate because they were less influenced by outside factors. It is also possible that these results were just another indication of the benefits of audio computer-assisted self-interviewing (ACASI) for collecting sensitive data. Having improved the privacy of the interview, we may now be improving respondents' honesty in reporting. Because the higher reporting is most obvious for younger respondents, this second hypothesis seems especially likely as youths are more likely to have the privacy of their interview compromised by the presence of other household members. Younger respondents, however, could be more likely to be influenced by the presence of scale categories, so the first hypothesis should not be ruled out. Most likely, both hypotheses are contributing to the results seen here.

Further research should be conducted to fully disentangle the effect of ACASI from the effect of the revised question text. As a first step in this direction, we compared the original 12- month frequency distribution to the 12-month distribution that was obtained from multiplying respondents' 30-day frequency by 12 to see whether the CAI or the PAPI method was more similar. In an attempt to understand the differences, we examined the consistency of the 30-day and 12-month frequency-of-use responses by multiplying the reports of use in the past 30 days by 12. As noted by the respondents, these will not necessarily agree because the 12-month frequency is likely to reflect the respondent's assessment of his or her average rate of use over a 12-month period.

Exhibit 6.5.3 presents the results of this comparison. We show the estimated number of days used and the ratio of 12 times the 30-day reports to the reported frequency of use in the past 12 months. It appears that the ACASI gave more consistent results in that all of the ratios were greater than one. This is what we would expect because respondents are likely to forget instances of use over a 12-month recall period. In addition, we note that, in general, the ratios tended to be similar; however, when they were different, in most cases the ACASI ratios were closer to one.

Exhibit 6.5.3 also shows the average number of days used in the past 30 days that were reported by respondents under each interview mode. We note that for the more sensitive drugs, ACASI resulted in higher reports of the number of days used.

Finally, we examined the patterns of reporting to determine if there is evidence that respondents were completing the response task as we intended. Exhibit 6.5.4 presents data on the number and percentage of respondents who chose to provide their answers using the three available reporting

frequencies (days per week, month, or year) and the distribution of the responses among those who chose to report using a particular reporting period. Data for alcohol, marijuana, and cocaine are shown.

The data indicate that the majority of respondents appeared to be responding as intended. For example, we would expect that respondents who chose to report monthly would tend to use less frequently than several days a week and those who selected yearly reporting would tend to use less than once a month. Thus, we see that most of the respondents who selected the monthly or yearly reporting periods were infrequent users. For example, if we examine the reported frequency of alcohol use among the youths, 91.1% of those who chose to report their use on a monthly basis reported using 10 or less days per month. For those who chose to report days per year, 80.8% reported using less than 10 or less days per year. Corresponding percentages for adults were 85.8% and 72.9%. Similar results were observed for marijuana and cocaine.

Based upon these findings, this revised procedure for asking about frequency of use in the past 12 months was adopted for the 1999 NHSDA.

Exhibit 6.5.1 Comparison of CAI to PAPI for "Don't Know" and "Refused" Responses for the 12-Month Frequency-of-Use Item (Including All Bad Data from the PAPI Data)

	DK and REF Responses			
	CAI		PAPI	
Drug Type	**%**	***n***	**%**	***n***
Alcohol	3.5	36	7.9	246
Marijuana	3.3	11	4.6	143
Cocaine	4.3	3	2.4	75
Crack	7.1	2	1.9	58
Heroin	8.3	1	2.2	67
Hallucinogens	1.4	1	1.5	47
Inhalants	7.2	6	1.2	56

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.5.2 Weighted Percentages of Past 12-Month Users Using 51 Days or More, by Age and Mode of Interview

	ACASI	PAPI Comparison Group
Age/Days of Use	**Alcohol**	
12 to 17		

Year Olds		
51 Days or More	34.86	9.83
50 Days or Less	65.14	90.17
18+ Year Olds		
51 Days or More	47.59	41.09
50 Days or Less	52.41	58.91
	Marijuana	
12 to 17 Year Olds		
51 Days or More	41.06	30.82
50 Days or Less	58.94	69.18
18+ Year Olds		
51 Days or More	41.73	33.70
50 Days or Less	58.27	66.30
	Cocaine	
12 to 17 Year Olds		

51 Days or More	28.15	3.62
50 Days or Less	71.85	96.38
18+ Year Olds		
51 Days or More	36.85	6.75
50 Days or Less	63.15	93.25
	Hallucinogens	
12 to 17 Year Olds		
51 Days or More	22.36	13.86
50 Days or Less	77.64	86.14
18+ Year Olds		
51 Days or More	36.94	13.93
50 Days or Less	63.06	86.07
	Inhalants	
12 to 17 Year Olds		
51 Days	31.33	14.37

or More		
50 Days or Less	68.67	85.63
18+ Year Olds		
51 Days or More	16.07	3.49
50 Days or Less	83.93	96.51

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.5.3 Consistency Between 30-Day Frequency and 12-Month Frequency-of-Use Respon
Rating of Most Consistent and Average Number of Days of Use Reported in Past 3

Drug	Type of Estimate	Total ACASI	Comparison Group 1997 Quarter 4 PAPI/SAQ	Ratio of 12*30 to 12-Month Frequency		Most Consistent	A
				ACASI	PAPI		
Alcohol	12*30	86.28	102.77	1.16	1.36	ACASI	
	12-Month Freq	74.10	75.31				
Marijuana	12*30	117.11	112.71	1.41	1.35	Equal	
	12-Month Freq	82.88	83.24				
Cocaine	12*30	82.41	47.53	1.60	2.43	ACASI	
	12-Month Freq	51.63	19.53				
Crack	12*30	85.85	34.14	1.68	0.63	Equal	
	12-Month Freq	51.16	54.05				
Hallucinogens	12*30	40.30	21.79	1.37	0.52	ACASI	
	12-Month Freq	29.41	41.90				
Inhalants	12*30	53.22	44.84	1.45	1.58	Equal	
	12-Month Freq	36.77	28.34				
Heroin	12*30	34.93	32.94	1.48	0.98	PAPI	
	12-Month Freq	23.61	33.67				

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 F

Exhibit 6.5.4 Number and Percentage of Respondents Selecting Weekly, Monthly, and Yearly Reporting Periods and Distributions of Responses Within Reporting Periods

	Alcohol					
	12 to 17 Year Olds			18+ Year Olds		
Period	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent
Per Week	95	23.8	23.8	164	28.9	28.9
Per Month	112	29.0	52.8	197	34.7	63.6

Per Year	182	47.2	100.0	207	36.4	100.0
Days Per Week for Weekly Reporters						
Days	**Frequency**	**Percent**	**Cumulative Percent**	**Frequency**	**Percent**	**Cumulative Percent**
1	48	52.2	52.2	60	36.6	36.6
2	17	18.5	70.7	39	23.8	60.4
3	15	16.3	87.0	21	12.8	73.2
4	4	4.3	91.3	16	9.8	82.9
5	4	4.3	95.7	16	9.8	92.7
6	1	1.1	96.7	8	4.9	97.6
7	3	3.3	100.0	4	2.4	100.0
Days Per Month for Monthly Reporters						
1	25	22.3	22.3	27	13.7	13.7
2	16	14.3	36.6	31	15.7	29.4
3	16	14.3	50.9	38	19.3	48.7
4	8	7.1	58.0	25	12.7	61.4
5	16	14.3	72.3	23	11.7	73.1
6	5	4.5	76.8	6	3.0	76.1
7	4	3.6	80.4	5	2.5	78.7
8	6	5.4	85.7	1	0.5	79.2
9	1	0.9	86.6	1	0.5	79.7
10	5	4.5	91.1	12	6.1	85.8
11-15	4	3.6	94.6	11	5.6	91.4
16-20	4	3.6	98.2	4	2.0	93.4
21-25	0	0.0	98.2	3	1.5	94.9
26-30	1	0.9	99.1	8	4.1	99.0
31	1	0.9	100.0	2	1.0	100.0
Days Per Year for Yearly Reporters						
1	36	19.8	19.8	17	8.2	8.2
2	33	18.1	37.9	24	11.6	19.8

3	35	19.2	57.1	26	12.6	32.4
4	9	4.9	62.1	15	7.2	39.6
5	12	6.6	68.7	26	12.6	52.2
6	4	2.2	70.9	16	7.7	59.9
7	2	1.1	72.0	6	2.9	62.8
8	3	1.6	73.6	4	1.9	64.7
9	1	0.5	74.2	2	1.0	65.7
10	12	6.6	80.8	15	7.2	72.9
11-15	6	3.3	84.1	13	6.3	79.2
16-20	13	7.1	91.2	8	3.9	83.1
21-30	8	4.4	95.6	14	6.8	89.9
31-50	2	1.1	96.7	7	3.4	93.2
50+	6	3.3	100.0	14	6.8	100.0
	Marijuana					
	12 to 17 Year Olds			**18+ Year Olds**		
Period	**Frequency**	**Percent**	**Cumulative Percent**	**Frequency**	**Percent**	**Cumulative Percent**
Per Week	54	30.2	30.2	40	34.2	34.2
Per Month	54	30.2	60.3	31	26.5	60.7
Per Year	71	39.7	100.0	46	39.3	100.0
Days Per Week for Weekly Reporters						
Days	**Frequency**	**Percent**	**Cumulative Percent**	**Frequency**	**Percent**	**Cumulative Percent**
1	16	29.6	29.6	12	30.0	30.0
2	6	11.1	40.7	3	7.5	37.5
3	6	11.1	51.9	5	12.5	50.0
4	5	9.3	61.1	2	5.0	55.0
5	4	7.4	68.5	3	7.5	62.5

6	5	9.3	77.8	5	12.5	75.0
7	12	22.2	100.0	10	25.0	100.0
Days Per Month for Monthly Reporters						
1	15	27.8	27.8	5	16.1	16.1
2	10	18.5	46.3	8	25.8	41.9
3	11	20.4	66.7	1	3.2	45.2
4	2	3.7	70.4	2	6.5	51.6
5	2	3.7	74.1	4	12.9	64.5
6	1	1.9	75.9	0	0.0	64.5
7	0	0.0	75.9	1	3.2	67.7
8	0	0.0	75.9	0	0.0	67.7
9	1	1.9	77.8	0	0.0	67.7
10	1	1.9	79.6	2	6.5	74.2
11-15	5	9.3	88.9	4	12.9	87.1
16-20	1	1.9	90.7	2	6.5	93.5
21-25	1	1.9	92.6	0	0.0	93.5
26-30	2	3.7	96.3	2	6.5	100.0
31	2	3.7	100.0	0	0.0	100.0
Days Per Year for Yearly Reporters						
1	19	26.8	26.8	9	19.6	19.6
2	6	8.5	35.2	6	13.0	32.6
3	3	4.2	39.4	8	17.4	50.0
4	10	14.1	53.5	4	8.7	58.7
5	6	8.5	62.0	5	10.9	69.6
6	2	2.8	64.8	0	0.0	69.6
7	2	2.8	67.6	0	0.0	69.6
8	2	2.8	70.4	0	0.0	69.6
9	2	2.8	73.2	0	0.0	69.6
10	1	1.4	74.6	6	13.0	82.6
11-15	4	5.6	80.3	2	4.3	87.0

16-20	4	5.6	85.9	0	0.0	87.0
21-30	2	2.8	88.7	2	4.3	91.3
31-50	4	5.6	94.4	1	2.2	93.5
50+	4	5.6	100.0	3	6.5	100.0
	Cocaine					
	12 to 17 Year Olds			**18+ Year Olds**		
Period	**Frequency**	**Percent**	**Cumulative Percent**	**Frequency**	**Percent**	**Cumulative Percent**
Per Week	6	30	30	6	17.1	17.1
Per Month	5	25.0	55	11	31.4	48.6
Per Year	9	45	100.0	18	51.6	100.0
Days Per Week for Weekly Reporters						
Days	**Frequency**	**Percent**	**Cumulative Percent**	**Frequency**	**Percent**	**Cumulative Percent**
1	3	5.0	50.0	2	33.3	33.3
2	2	33.3	83.3	0	0.0	33.3
3	0	0.0	83.3	1	16.7	50.0
4	0	0.0	83.3	1	16.7	66.7
5	0	0.0	83.3	2	33.3	100.0
6	0	0.0	83.3	0	0.0	100.0
7	1	16.7	100.0	0	0.0	100.0
Days Per Month for Monthly Reporters						
1	1	20.0	20.0	0	0.0	0.0
2	1	20.0	40.0	2	18.2	18.2
3	0	0.0	40.0	3	27.3	45.5
4	1	20.0	60.0	1	9.1	54.5
5	0	0.0	60.0	1	9.1	63.6
6	0	0.0	60.0	0	0.0	63.6

7	0	0.0	60.0	1	9.1	72.7
8	0	0.0	60.0	0	0.0	72.7
9	0	0.0	60.0	0	0.0	72.7
10	0	0.0	60.0	0	0.0	72.7
11-15	1	20.0	80.0	1	9.1	81.8
16-20	0	0.0	80.0	1	9.1	90.9
21-25	1	20.0	100.0	1	9.1	100.0
26-30	0	0.0	100.0	0	0.0	100.0
31	0	0.0	100.0	0	0.0	100.0
Days Per Year for Yearly Reporters						
1	5	55.6	55.6	3	16.7	16.7
2	3	33.3	88.9	4	22.2	38.9
3	0	0.0	88.9	2	11.1	50.0
4	0	0.0	88.9	4	22.2	72.2
5	0	0.0	88.9	1	5.6	77.8
6	0	0.0	88.9	0	0.0	77.8
7	0	0.0	88.9	0	0.0	77.8
8	0	0.0	88.9	0	0.0	77.8
9	0	0.0	88.9	0	0.0	77.8
10	0	0.0	88.9	0	0.0	77.8
11-15	0	0.0	88.9	0	0.0	77.8
16-20	0	0.0	88.9	1	5.6	83.3
21-30	0	0.0	88.9	0	0.0	83.3
31-50	0	0.0	88.9	1	5.6	88.9
50+	1	11.1	100.0	2	11.1	100.0

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

6.6 Comparison of ACASI Respondent Attitudes to Attitudes of PAPI Respondents: Results from Debriefing Interviews

All ACASI respondents in the 1997 field experiment were asked to complete a debriefing questionnaire at the end of the interview. ACASI was used so that respondents could answer the debriefing questions privately. A subsample of the respondents in Quarter 4 of the 1997 NHSDA also completed a debriefing interview using ACASI methodology. Using ACASI to answer debriefing questions gave the comparison group some minimal experience with answering questions using the computer and allowed us to query them as to their preferences for using a computer versus paper-and-pencil methods and about the relative privacy of the two modes of interview.

Confirming the results from the 1996 feasibility experiment, the debriefings clearly indicate that ACASI provides a more private setting for the interview and that most respondents prefer to use the computer. This is particularly true for 12 to 17 year olds.

6.6.1 Participation in the Debriefing Interviews

A total of 1,953 (99%) of the respondents in the 1997 field experiment completed the debriefing interview. We obtained debriefing interviews from 584 of the respondents in Quarter 4 of the 1997 NHSDA. A total of 713 respondents were selected for these interviews; the interviewers neglected to request participation in the ACASI debriefing interview from 25 respondents (4%); among the 688 respondents who were asked to complete the debriefing interview, 584 (85%) agreed to complete the interview. Exhibit 6.6.1 summarizes the characteristics of the debriefing respondents.

The CAI respondents answered 25 debriefing questions, and the PAPI/SAQ respondents answered 21. Eighteen of the questions were nearly identical in the two debriefing interviews in that they addressed the same topic for the two different interview modes.

6.6.2 Comparison of Respondent Attitudes

In Exhibits 6.6.2 through 6.6.22, we present unweighted tabulations of the debriefing responses for the 18 questions that were answered by both the ACASI and PAPI/SAQ respondents. Some of these exhibits present comparisons of debriefing interview respondents and interviewer debriefing responses.

Ability to complete the response task. Respondents in the 1997 Quarter 4 NHSDA debriefing sample were asked to rate the difficulty of recording their answers on the answer sheets; those in the 1997 field experiment were asked to rate the difficulty of using the computer. In both cases, the large majority of the respondents indicated that the task was not at all difficult

(see Exhibit 6.6.2). An estimated 86% of the respondents who completed PAPI/SAQ answer sheets indicated that the task was not at all difficult, and there was a slight tendency for more of the ACASI respondents to indicate that the task was not at all difficult (91.4%). There was a larger difference between the two modes for the youths (aged 12 to 17); the corresponding percentages are 83.6% for those completing answer sheets and 91.4% for those using the computer. Among the adult respondents aged 18 or older, those with less than a high school education had a tendency to report more difficulty with the response task, and this was true for both interview modes.

Larger differences between the two interview modes were observed in response to the question concerning whether or not the respondent was able to complete the self-response task without the help of the interviewer. Data are shown in Exhibit 6.6.3. Among the ACASI respondents, 88.3% indicated that they did not need help, whereas only 73.5% of those who filled out the answer sheets indicated that they did not need help from the interviewer to do so. This difference was even larger for the youths, with 20% fewer of the ACASI respondents indicating that they needed help from the interviewer (33.2% for PAPI vs. 13.2% for ACASI). Again, among the adult respondents, those with less than a high school education were more likely to indicate that they needed help from the interviewer, and fewer of the ACASI respondents needed help.

One of the reasons for using ACASI is to ease the response task for poor readers. The results of the debriefing interviews indicated that ACASI did make it easier for poor readers to respond. To investigate this, we first asked respondents to rate their own reading ability. About the same proportion of people in the experimental and comparison group classified themselves as having fair to poor reading ability (15.2% and 16.8%, respectively). We also asked those who had received an ACASI interview how much the recorded voice helped them and how often they listened to it while completing the interview.[15]

Respondents who reported fair to poor reading ability were more likely to find the audio helpful and to listen to it during the interview (Exhibit 6.6.4). About 49% of the ACASI respondents reporting fair or poor reading abilities indicated that the recorded voice helped them a lot to understand the questions, whereas only about 15% of ACASI respondents reporting excellent reading abilities indicated that the recorded voice was of some help in understanding the questions. Among the ACASI respondents reporting fair or poor reading abilities, about 65% indicated that they listened to the recorded voice most of the time as compared to 37% of the ACASI respondents reporting excellent reading abilities (Exhibit 6.6.5).

We were concerned that respondents who did not routinely use computers would not like using them to complete the interview. To investigate this issue, we asked the comparison group who had used paper answer sheets to indicate whether they would rather use the computer or the paper answer sheets.

About 37% of the respondents who reported that they had not used the computer prior to the interview indicated that they would rather use the computer (Exhibit 6.6.6), and 32% of them indicated that it does not make any difference. ACASI respondents had no experience with the paper-and-pencil answer sheets (PAPI/SAQ); thus, we were not able to examine their preferences for an interview mode. We did have information, however, on their need for help and examined how this related to their computer experience. About 77% of ACASI respondents who reported they were using a computer for the first time indicated that they could easily enter their answers into the computer without asking the interviewer for help (Exhibit 6.6.7). An estimated 92% of the everyday users required no help.

Privacy of the interviewing environment. ACASI was designed to increase the privacy of the interviewing environment. Dramatic differences between the two modes of data collection were observed in response to the question on how many of the responses the interviewer saw (Exhibit 6.6.8). Only 41.3% of those completing answer sheets indicated that the interviewer saw none of their answers, whereas 82.6% of the ACASI respondents indicated such. In addition, under the paper-and-pencil mode, 42% of respondents indicated that the interviewer saw some of their answers, but only 13% of the ACASI respondents indicated such. This finding supports results from the 1996 feasibility experiment. In that experiment, we did not debrief respondents using a structured questionnaire but did ask interviewers to report, for each respondent, how often the respondent let them know his or her answer. In that study, interviewers indicated that they were aware of none of the respondent's answers in only 38% of the PAPI/SAQ cases, and for the two ACASI versions, the corresponding percentages were 63.7% and 73.8% for the MIRROR and the SKIP versions.

We also examined how the seating arrangement affected the privacy of the interview (Exhibit 6.6.9). About 73% of the ACASI respondents indicated that an interviewer seated next to them with no table space did not see any of their answers; about 84.2% indicated that the interviewer seated across with no table space did not see any of the answers. This indicates that the survey protocol still needs to stress obtaining a private setting, including making sure that the interviewer is not able to observe the respondent's answers.

Respondents were asked how many of their answers they thought that someone in their household other than the interviewer saw (see Exhibit 6.6.10). Under both modes, a large majority of the respondents indicated that no one other than the interviewer saw their answers (78.3% for those who used answer sheets vs. 83.6% for those using the computer). However, a more than desirable proportion of the respondents indicated that someone in the household saw some of their answers. For adult respondents, about 10% of those completing answer sheets indicated that someone else had seen their answers, and 7.5% of the ACASI respondents reported this as well. For the youths, the figures were even higher (21.2% for the PAPI; 18.9% for the ACASI).

Respondents also were asked how important it was to them that the interviewers and other household members not see their answers. Results for these two questions are presented in Exhibits 6.6.11 and 6.6.12. In both cases, the ACASI respondents were more likely to indicate that it was very important that their answers be concealed from either the interviewers or other household members. However, only about 26% of respondents indicated that this was very important. When asked about the importance of concealing their answers from the interviewer, nearly 59% of those completing the answer sheets reported that it was not at all important; the corresponding figure for the ACASI respondents was about 49%. Adults in both settings were more likely to say that it was not important to conceal their answers from the interviewer when compared to youths. Similar results were observed when respondents were asked about concealing their responses from other household members. Adults were less concerned than youths with concealing their answers from other household members, and ACASI respondents were somewhat more likely to report that it was very important to conceal the answers from other household members.

Exhibit 6.6.13 displays information on the interviewing environment and the respondents' reports as to whether or not someone in the household saw their answers. It is encouraging to note that among the 240 youths for whom a parent was present during the interview, 73% reported that none of their answers was seen by another household member. The fact that parents saw answers, however, in over 25% of the cases is a problem. The 240 youths from whom parents were present comprise 22% of the youth respondents. Overall, there was someone present during the interview for 497 (45%) of the youths. For 30% of these youths, at least someone in the household saw some of their answers. We note similar results for the adults. Exhibit 6.6.14 indicates that respondents for whom someone saw their answers find privacy just as important as those for whom no one saw their answers.

This series of exhibits has some worrisome findings. It indicates that respondents want a private setting for the interview and that this is not always achieved. The ACASI technology has the potential to achieve greater privacy; however, the interviewer also has some control over this. During interviewer training and supervision, the need for privacy needs to be continually emphasized.

Members of the 1997 Quarter 4 NHSDA debriefing sample were asked which method of completing the survey was best for protecting their privacy while they were completing the survey (see Exhibit 6.6.15). About 11% of these respondents indicated that the answer sheets were the best; 57.2% of youths and 42.5% of adults felt that the computer was best; and 23.3% and 28.4% of youths and adults, respectively, felt that the two methods were equal in protecting privacy. It is interesting to note that overall some 13% of the respondents felt that neither protected their privacy while they were completing the interview (9.2% of youths vs. 16.8% of adults).

Similar results (see Exhibit 6.6.16) were observed in response to the question on protection of privacy after the interviewer had left the household. Around 12% reported feeling that the answer sheets provided the best protection after the interviewer had left the household; 46.6% chose the computer; and 24.5% indicated that the methods were equal. The percentage reporting that neither protected their privacy after the interviewer left their home was around 16%, with adults slightly more likely than youths to report neither. These percentages are larger than those who reported that they were not confident that their results will not be linked with their names, indicating that respondents may be confident of the promises of confidentiality.

Respondents who completed the ACASI interview had no experience completing the interview using answer sheets. They were told, however, that in some cases people recorded their answers on answer sheets and sealed them in an envelope. They were asked which method they felt was most private. The responses are very similar (see Exhibit 6.6.17) to those of the 1997 NHSDA debriefing sample, with about 13% saying that the paper would be less private. About 47% chose the computer, and 39% said that there was no difference. Again, youths were more likely than adults to choose the computer.

Accuracy of responses. Two questions were asked about the need for accuracy and the overall accuracy of responses. In Exhibit 6.6.18, we compare the responses to the question on their opinions as to how accurate the sponsor wanted them to be. In both samples, about 80% of respondents thought that very accurate data were needed, and adults were more likely

than youths to think that the accuracy was needed. Exhibit 6.6.19 presents respondents' ratings of the accuracy of their responses. The vast majority said that they were very accurate (80.8% of the 1997 Quarter 4 respondents vs. 83.1% of the 1997 field experiment respondents). Very small minorities indicated that their responses were not very accurate.

Interest in the interview and willingness to respond again. The ACASI respondents were more likely to say that the interview was very or somewhat interesting (55.3% vs. 46.9%). These results are presented in Exhibit 6.6.20. In both samples, adults found it more interesting than youths; however, the difference between the ACASI and PAPI respondents was greater for youths than adults. Among youths, 51.6% of ACASI respondents reported that the interview was somewhat or very interesting, and only 39.7% of those who completed the answer sheets gave such ratings. The corresponding percentages for adults were 60.2% and 54.1%. Adults with less education were most likely to find the interview interesting (35.5% of the ACASI respondents with less than a high school education found the interview interesting vs. 22.5% of the adults with a high school education and 11.8% of those with more than a high school education).

Most respondents said that they were very or somewhat likely to complete a similar interview if asked, and those responding by computer were slightly more likely to report willingness to participate in the future (79.9% vs. 75.1%). Youths were about 4% to 5% more likely than adults to say that they would do it again (Exhibit 6.6.21).

Under both modes, the majority of respondents were either very or somewhat confident that their answers would never be linked with their names (Exhibit 6.6.22). There was a small difference between the two samples in the percentage who indicated that they were not at all confident, with 8.7% of the 1997 NHSDA Quarter 4 debriefing respondents being not at all confident and 5.6% of the ACASI respondents.

Exhibit 6.6.1 Characteristics of Debriefing Interview Respondents

	1997 Field Experiment		Comparison Group	
	CAPI/ACASI		1997 Quarter 4 PAPI/SAQ	
Respondent Characteristics	*n*	%	*n*	%
Total	1,953	100	584	100

Age Group				
12-17	1,102	56.4	292	50.0
18+	851	43.6	292	50.0
Gender				
Male	915	46.9	246	42.1
Female	1,038	53.1	338	57.9
Race/Ethnicity				
Hispanic	463	23.7	121	20.7
Non-Hisp., Black	521	26.7	193	33.0
Non-Hisp., All Other Races	969	49.6	270	46.2
Education[1]				
< High School	197	23.1	47	16.1
High School	316	37.1	100	34.2
> High School	338	39.7	145	49.7

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.2 Comparison of Debriefing Interview Respondents on Selected Debriefing Questio Difficulty Recording Answers (Unweighted Percentages)

	1997 Field Experiment				Comparison Grou		
	CAPI/ACASI				1997 Quarter 4 PAPI/SAQ		
	How difficult was it for you to use the computer to do this survey?				How difficult was it for you to answers on the answer sh		
Respondent Characteristics	Very	Somewhat	Not at All	DK/REF	Very	Somewhat	Not at A
Total	2.2	6.3	91.4	0.1	1.4	12.2	86
Age Group							
12-17	1.7	6.8	91.4	0.1	1.4	15.1	83
18+	2.7	5.8	91.5	0.0	1.4	9.2	88
Gender							
Male	2.5	6.6	90.8	0.1	1.6	12.2	86
Female	1.8	6.2	92.0	0.0	1.2	12.1	85
Race/Ethnicity							
Hispanic	2.8	8.6	88.3	0.2	3.3	12.4	84
Non Hisp., Black	3.5	7.7	88.9	0.0	2.1	13.5	83
Non Hisp., All Other Races	1.1	4.5	94.3	0.0	0.0	11.1	88
Education[1]							
< High School	4.1	10.2	85.8	0.0	2.1	14.9	83
High School	3.5	7.0	89.6	0.0	3.0	10.0	85
> High School	1.2	2.1	96.7	0.0	0.0	6.9	92

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experim
Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.3 Comparison of Debriefing Interview Respondents on Selected Debriefing Questio of Answering Questions (Unweighted Percentages)

	1997 Field Experiment			Comparison Group		
	CAPI/ACASI			1997 Quarter 4 PAPI/SAQ		
	Were you able to enter your answers easily into the computer without having to ask the interviewer for help?			Were you able to complete th sheets easily without having t interviewer for help?		
Respondent Characteristics	Yes	No	DK/REF	Yes	No	
Total	88.3	11.7	0.0	73.5	26.4	
Age Group						
12-17	86.8	13.2	0.0	66.8	33.2	
18+	90.2	9.8	0.0	80.1	19.5	
Gender						
Male	88.5	11.5	0.0	72.0	28.0	
Female	88.1	11.9	0.0	74.6	25.1	
Race/Ethnicity						
Hispanic	83.2	16.8	0.0	72.7	27.3	
Non Hisp., Black	87.7	12.3	0.0	69.9	30.1	
Non Hisp., All Other Races	91.9	9.0	0.0	76.3	23.3	
Education[1]						
< High School	83.2	16.8	0.0	68.1	31.9	
High School	90.8	9.2	0.0	84.0	16.0	
> High School	93.8	6.2	0.0	81.4	17.9	

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experim
National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.4 Comparison of Assistance Provided by the Recorded Voice, by Respondent's Rating of Reading Ability (Unweighted Percentages)

	1997 Field Experiment				
	CAPI/ACASI				
	Some people believe that having a recorded voice read the questions will help respondents understand the questions better. How much did the recorded voice help you to understand the questions?				
Reading Ability	**No Help**	**Some Help**	**A Lot of Help**	**Did Not Listen**	**DK/REF**
Total	35.3	31.6	24.6	8.4	0.1

Excellent	46.0	26.5	14.7	12.8	0.0
Good	30.1	37.1	26.6	5.7	0.1
Fair/Poor	16.8	31.9	48.7	2.7	0.0
DK/Ref	25.0	50.0	0.0	0.0	25.0

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 6.6.5 Use of Recorded Voice, by Respondent's Rating of Reading Ability (Unweighted Percentages)

	1997 Field Experiment			
	CAPI/ACASI			
	Did you listen to the recorded voice most or all of the time, some of the time, or little or none of the time?			
Reading Ability	**Most of the Time**	**Some of the Time**	**Little or None of the Time**	**DK/REF**
Excellent	36.5	19.5	44.0	0.1
Good	48.7	26.7	24.3	0.1
Fair/Poor	64.8	22.5	12.8	0.0
DK/Ref	25.0	25.0	25.0	25.0

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 6.6.6 Respondent's Preference to Use Computers or Answer Sheets, by Respondent's Computer Experience (Unweighted Percentages)

	Comparison Group			
	1997 Field Experiment			
	Would you rather use the computer, fill out the answer sheet or wouldn't it matter to you?			
Computer Experience	**Would Rather Use the**	**Would Rather Fill Out**	**Doesn't Make Any Difference**	**DK/REF**

	Computer	Answer Sheet		
First-Time User[1]	36.5	31.8	31.8	0.0
Prior User[2]				
Currently not using	46.2	3.9	50.0	0.0
Less than once a month	39.2	20.0	41.2	0.0
One to a few days a month	55.4	18.5	26.1	0.0
One to four days a week	47.9	9.4	42.7	0.0
Everyday	59.4	6.8	33.8	0.0
DK/REF	0.0	0.0	0.0	100

[1]Respondents who answered "yes" when asked, "Is this the first time you have ever used a computer?"

[2]Respondents who answered "no" when asked, "Is this the first time you have ever used a computer?"

Source: 1997 National Household Survey on Drug Abuse; Quarter 4.

Exhibit 6.6.7 Respondent's Ability to Enter Answers into the Computer Without the Interviewer's Help, by Respondent's Computer Experience (Unweighted Percentages)

	1997 Field Experiment	
	CAPI/ACASI	
	Were you able to enter answers into the computer easily, without having to ask the interviewer for help?	
Computer Experience	**Yes**	**No**
First-Time User[1]	76.7	23.2
Prior User[2]		
Currently not using	82.8	17.2

Less than once a month	90.4	9.6
One to a few days a month	89.2	10.8
One to four days a week	87.5	12.5
Everyday	91.7	8.3
DK/REF	71.4	28.6

[1]Respondents who answered "yes" when asked, "Is this the first time you have ever used a computer?"

[2]Respondents who answered "no" when asked, "Is this the first time you have ever used a computer?"

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures: 1997 Field Experiment.

Exhibit 6.6.8 Comparison of Debriefing Interview Respondents on Selected Debriefing Questio Interviewer See Answers? (Unweighted Percentages)

	1997 Field Experiment					Comparison Grou			
	CAPI/ACASI					1997 Quarter 4 PAPI/SAQ			
	How many of your answers that you entered into the computer do you think that the interviewer saw?					How many of your answers marked on the answer shee think that the interviewe			
Respondent Characteristics	None	Some	A Lot	All	DK/REF	None	Some	A Lot	All
Total	82.6	13.1	1.9	1.8	0.6	41.3	42.1	7.0	8.7
Age Group									
12-17	79.4	16.3	2.2	1.3	0.8	38.7	43.8	8.9	7.5
18+	86.7	8.9	1.6	2.6	0.1	43.8	40.4	5.1	9.9
Gender									
Male	81.1	14.1	2.2	2.0	0.6	39.8	41.5	7.3	10.6
Female	83.9	12.2	1.7	1.7	0.4	42.3	42.6	6.8	7.4
Race/Ethnicity									
Hispanic	77.1	17.1	2.8	2.4	0.6	39.7	43.8	5.8	10.7
Non Hisp., Black	80.6	13.2	2.5	3.1	0.6	41.5	35.2	9.3	14.0
Non Hisp., All Other Races	86.3	11.1	1.2	0.9	0.4	41.9	46.3	5.9	4.1
Education[1]									
< High School	78.2	13.2	3.6	5.1	0.0	40.4	36.2	6.4	17.0
High School	87.0	9.2	1.3	2.2	0.3	50.0	34.0	7.0	9.0
> High School	91.4	6.2	0.9	1.5	0.0	40.7	46.2	3.4	8.3

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experime Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.9 Comparison of Debriefing Interview Respondents to Interviewer Debriefing Responses: Seating Arrangement Versus Answers Interviewer Saw (Unweighted Percentages)

	1997 Field Experiment				
	CAPI/ACASI				
Place Where the Interviewer Was Seated During	How many of the answers that you entered into the computer do you think the interviewer saw?				
	None of the	**Some of the Answers**	**A Lot of the Answers**	**All of the Answers**	**DK/REF**

ACASI According to the Interviewer	**Answers**				
Seated Next to the Respondent with Table Space Available	79.3	14.7	2.8	2.8	0.5
Seated Next to the Respondent with No Table Space Available	72.9	20.8	2.1	3.5	0.7
Seated Across from Respondent with Table Space Available	86.3	10.8	1.3	1.1	0.5
Seated Across from Respondent with No Table Space Available	84.2	11.2	2.0	2.0	0.7
Some Other Arrangement	83.1	13.2	2.1	1.1	0.5

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 6.6.10 Comparison of Debriefing Interview Respondents on Selected Debriefing Ques
Anyone Else in Household See Answers? (Unweighted Percentages)

	1997 Field Experiment					Comparison Gr			
	CAPI/ACASI					1997 Quarter PAPI/SAQ			
	How many of the answers that you entered into the computer do you think someone in your household other than the interviewer saw?					How many of your answe you think someone oth interviewer sav			
Respondent Characteristics	None	Some	A Lot	All	DK/REF	None	Some	A Lot	A
Total	83.6	13.9	1.2	1.1	0.2	78.3	15.6	2.9	3
Age Group									
12-17	78.4	18.9	1.5	0.9	0.3	70.9	21.2	4.8	3
18+	90.4	7.5	0.8	1.3	0.0	85.6	9.9	1.0	3
Gender									
Male	81.6	16.0	1.2	0.9	0.3	77.2	15.4	3.7	3
Female	85.4	12.1	1.3	1.3	0.0	79.0	15.7	2.4	2
Race/Ethnicity									
Hispanic	79.0	18.1	1.7	0.9	0.2	70.3	19.5	2.5	6
Non Hisp., Black	84.8	11.5	1.3	2.1	0.2	81.9	13.0	2.6	2
Non Hisp., All Other Races	85.1	13.2	0.9	0.6	0.1	80.0	14.8	3.3	1
Education[1]									
< High School	85.8	9.1	1.5	3.6	0.0	87.2	6.4	2.1	4
High School	90.2	8.9	0.6	0.3	0.0	83.0	12.0	1.0	4
> High School	93.2	5.3	0.6	0.9	0.0	86.9	9.7	0.7	2

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experime
Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.11 Comparison of Debriefing Interview Respondents on Selected Debriefing Ques
Important That Interviewer Not See Answers? (Unweighted Percentages)

	1997 Field Experiment				Comparison Grou		
	CAPI/ACASI				1997 Quarter 4 PAPI/SAQ		
	How important is it to you that the interviewer not see the answers you entered into the computer?				How important is it to you that t not see how you answered the the answer sheets?		
Respondent Characteristics	Very	Somewhat	Not at All	DK/REF	Very	Somewhat	Not at
Total	26.0	25.0	48.6	0.5	15.2	25.7	58
Age Group							
12-17	26.7	27.7	45.0	0.6	18.5	28.8	52
18+	25.0	21.5	53.2	0.2	12.0	22.6	65
Gender							
Male	27.8	24.4	47.4	0.4	16.3	26.4	57
Female	24.4	25.5	49.6	0.5	14.5	25.1	60
Race/Ethnicity							
Hispanic	31.3	26.1	42.3	0.2	19.8	19.8	60
Non Hisp., Black	25.9	21.9	51.4	0.8	13.5	23.8	62
Non Hisp., All Other Races	23.5	26.1	50.0	0.4	14.4	29.6	55
Education[1]							
< High School	21.8	18.8	58.9	0.5	10.6	25.5	63
High School	27.5	20.3	51.9	0.3	12.0	22.0	66
> High School	24.6	24.3	51.2	0.0	12.4	22.1	64

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experim
Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.12 Comparison of Debriefing Interview Respondents on Selected Debriefing Ques
Important That No One in Your Household See Answers? (Unweighted Percent

	1997 Field Experiment				Comparison Group		
	CAPI/ACASI				1997 Quarter 4 PAPI/SAQ		
	How important is it that no one in your household see the answers you entered into the computer?				How important is it that no o household see how you ansv questions on the answer s		
Respondent Characteristics	Very	Somewhat	Not at All	DK/REF	Very	Somewhat	Not at A
Total	26.2	21.0	52.4	0.5	17.0	23.5	59.4
Age Group							
12-17	31.8	24.1	43.5	0.6	21.6	30.1	48.3
18+	18.9	17.0	63.8	0.2	12.3	16.8	70.5
Gender							
Male	26.6	20.9	51.0	0.5	16.7	26.8	56.5
Female	25.0	21.1	53.6	0.4	17.2	21.0	61.5
Race/Ethnicity							
Hispanic	29.4	24.4	46.0	0.2	19.0	22.3	58.7
Non Hisp., Black	25.1	16.3	56.6	1.0	17.1	18.1	64.8
Non Hisp., All Other Races	25.2	21.9	52.6	0.3	15.9	27.8	55.9
Education[1]							
< High School	18.8	15.7	65.0	0.5	19.1	19.1	61.7
High School	19.3	17.7	62.7	0.3	12.0	18.0	70.0
> High School	18.6	17.2	64.2	0.0	10.3	15.2	73.6

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experime
Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.13 Comparison of Debriefing Interview Respondents to Interviewer Debriefing Resp Other People Present During the Interview Versus How Many Answers Household Members Saw (Unweighted Percentages)

Other People Present or Listening to the Interview as Recorded by the Interviewer	1997 Field Experiment CAPI/ACASI									
	How many of the answers that you entered into the computer do you think someone in other than the interviewer saw?									
	Age Group: 12 to 17 Years						Age Group: 18 Years o			
	n	None of Answers	Some of the Answers	A Lot of Answers	All of Answers	DK/ REF	*n*	None of Answers	Some of the Answers	A Lot Answe
Parents Only	240	73.3	21.7	3.0	1.7	0.4	31	74.2	22.6	0
Spouse Only	1	100	0.0	0.0	0.0	0.0	41	75.6	19.5	2
Live-in Partner/Boyfriend/ Girlfriend Only	1	100	0.0	0.0	0.0	0.0	13	76.0	15.4	0
Adult Relatives	30	66.7	30.0	0.0	3.3	0.0	22	90.9	4.6	0
Other Adults	6	100.0	0.0	0.0	0.0	0.0	19	94.7	5.3	0
Child(ren) Less Than 15 Years of Age	92	70.7	25.0	2.2	1.1	1.1	104	84.6	13.5	1
Other	20	80.0	20.0	0.0	0.0	0.0	16	87.5	6.3	6
Family	80	56.3	42.5	0.0	1.3	0.0	32	68.8	25.0	3
Family and Others	26	77.0	15.4	7.7	0.0	0.0	1	0.0	100.0	0
DK/REF	1	100.0	0.0	0.0	0.0	0.0	14	85.7	7.1	0
No One Present	605	84.8	13.6	0.0	0.5	0.2	558	95.2	3.6	0

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experimen

Exhibit 6.6.14 Comparison of Debriefing Interview Respondents to Interviewer Debriefing Respo Other People Present During the Interview Versus Importance of Household Mem Seeing the Answers (Unweighted Percentages)

Other People Present or Listening to the Interview as Recorded by the Interviewer	1997 Field Experiment CAPI/ACASI								
	How important is it to you that no one in your household see how you answered that you entered into the computer yourself?								
	Age Group: 12 to 17 Years					Age Group: 18 Years or			
	n	Very Important	Somewhat Important	Not Very Important	DK/ REF	*n*	Very Important	Somewhat Important	No Im
Parents Only	240	30.0	20.8	47.9	1.3	31	25.8	19.4	
Spouse Only	1	100.0	0.0	0.0	0.0	41	9.8	22.0	
Live-in Partner/Boyfriend/ Girlfriend Only	1	0.0	0.0	100.0	0.0	13	7.7	15.4	
Adult Relatives	30	30.0	13.3	56.7	0.0	22	27.3	22.7	
Other Adults	6	33.3	33.3	33.3	0.0	19	10.5	10.5	
Child(ren) Less Than 15 Years of Age	92	34.8	23.9	40.2	1.1	104	14.4	18.3	
Other	20	30.0	35.0	35.0	0.0	16	25.0	18.8	
Family	80	28.8	17.5	53.8	0.0	32	18.8	15.6	
Family and Others	26	30.8	19.2	50.0	0.0	1	0.0	0.0	
DK/REF	1	0.0	100.0	0.0	0.0	14	28.6	14.3	
No One Present	605	32.6	26.5	40.4	0.5	558	19.9	16.5	

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures: 1997 Field Experimen

Exhibit 6.6.15 Comparison of Debriefing Interview Respondents on Selected Debriefing Questions: Privacy Protection (Unweighted Percentages)

	Comparison Group				
	1997 Quarter 4 NHSDA				
	Which method do you think is best for protecting your privacy while completing the survey?				
Respondent Characteristics	**Computer**	**Answer Sheets**	**Equally Well**	**Neither**	**DK/REF**
Total	49.8	11.0	25.9	13.0	0.4
Age Group					
12-17	57.2	10.3	23.3	9.2	0.0
18+	42.5	11.6	28.4	16.8	0.6

Gender					
Male	51.2	10.2	23.6	14.6	0.4<
Female	48.8	11.5	27.5	11.8	0.3
Race/Ethnicity					
Hispanic	52.9	10.7	25.6	10.7	0.0
Non-Hisp., Black	51.8	13.5	19.2	15.0	0.5
Non-Hisp., All Other Races	47.0	9.3	30.7	12.6	0.4
Education[1]					
< High School	46.8	12.8	23.4	14.9	2.1
High School	43.0	13.0	29.0	15.0	0.0
> High School	40.7	10.3	29.7	18.6	0.7

[1]Education includes only individuals aged 18 or older.

Source: 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.16 Comparison of Debriefing Interview Respondents on Selected Debriefing Questions: Privacy Protection After Interviewer Leaves (Unweighted Percentages)

Respondent Characteristics	Comparison Group: 1997 Field Experiment: Which method do you think is best for protecting the privacy of your answers once the interview is complete and the interviewer has left your house?				
	Computer	Answer Sheets	Equally Well	Neither	DK/RE
Total	46.6	12.2	24.5	16.4	0.
Age Group					
12-17	40.8	10.3	18.5	14.4	0.
18+	52.4	14.0	30.5	18.5	0.
Gender					
Male	48.4	12.2	22.0	7.7	0.
Female	45.3	12.1	26.3	16.0	0.
Race/Ethnicity					
Hispanic	50.4	9.1	28.1	12.4	0.
Non Hisp., Black	49.7	13.0	18.7	18.1	0.
Non Hisp., All Other Races	42.6	13.0	27.0	17.0	0.

Source: 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.17 Comparison of Debriefing Interview Respondents on Selected Debriefing Questions: Private Answers (Unweighted Percentages)

	1997 Field Experiment			
	CAPI/ACASI			
	Do you think your answers are more private, about as private, or less private because you entered them into the computer?			
Respondent Characteristics	More	About as	Less	DK/REF
Total	46.7	39.2	13.3	0.8
Age Group				
12-17	50.5	34.4	13.9	1.2
18+	41.8	45.4	12.6	0.2
Gender				
Male	44.6	38.3	16.7	0.3
Female	48.6	40.0	10.3	1.2
Race/Ethnicity				
Hispanic	49.9	35.6	13.2	1.3
Non Hisp., Black	43.0	38.4	17.9	0.8
Non Hisp., All Other Races	47.2	41.3	11.0	0.5
Education[1]				
< High School	55.2	34.4	10.4	0.0
High School	42.7	44.3	12.7	0.3
> High School	34.0	53.3	12.4	0.3

[1]Education includes only individuals aged 18 or older.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 6.6.18 Comparison of Debriefing Interview Respondents on Selected Debriefing Questi Importance of Accuracy (Unweighted Percentages)

	1997 Field Experiment				Comparison Grou		
	CAPI/ACASI				1997 Quarter PAPI/SAQ		
	How important do you think it is to the sponsor of this survey that you give accurate and complete answers to the questions on this survey?				How important do you think it is of this survey that you give a complete answers to the ques survey?		
Respondent Characteristics	Very	Somewhat	Not at All	DK/REF	Very	Somewhat	Not at
Total	81.4	15.6	2.5	0.5	80.1	16.6	2
Age Group							
12-17	76.6	19.6	3.2	0.7	75.3	19.9	4
18+	87.7	10.6	1.5	0.2	84.9	13.4	1
Gender							
Male	80.8	15.6	3.1	0.5	79.3	17.5	3
Female	82.0	15.7	1.9	0.4	80.8	16.0	2
Race/Ethnicity							
Hispanic	75.8	20.5	3.0	0.6	79.3	18.2	2
Non Hisp., Black	81.0	15.6	2.9	0.6	77.2	17.1	5
Non Hisp., All Other Races	84.4	13.3	2.0	0.3	82.6	15.6	1
Education[1]							
< High School	85.3	12.7	2.0	0.0	78.7	17.0	4
High School	86.7	11.4	1.3	0.6	84.0	15.0	1
> High School	89.9	8.6	1.5	0.0	86.6	11.0	0

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experim Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.19 Comparison of Debriefing Interview Respondents on Selected Debriefing Questions: Accuracy and Completeness of Answers (Unweighted Percentages)

	1997 Field Experiment				Comparison Group			
	CAPI/ACASI				1997/Quarter 4 PAPI/SAQ			
	How accurate and complete are the answers that you entered into the computer?				How accurate and complete are the answers that you gave?			
Respondent Characteristics	Very	Fairly	Not Very	DK/REF	Very	Fairly	Not Very	DK/REF
Total	83.1	15.6	1.1	0.2	80.8	17.0	1.9	0.4
Age Group								
12-17	82.3	15.8	1.6	0.3	79.1	18.5	2.1	0.3
18+	84.1	15.3	0.5	0.1	82.5	15.4	1.7	0.3
Gender								
Male	81.9	16.2	1.6	0.2	78.9	17.9	2.8	0.4
Female	84.1	15.0	0.7	0.2	82.2	16.3	1.2	0.3
Race/Ethnicity								
Hispanic	79.3	18.8	1.7	0.2	81.7	15.0	3.3	0.0
Non Hisp., Black	84.4	14.6	0.8	0.2	78.2	18.1	3.6	0.0
Non Hisp., All Other Races	84.2	14.6	1.0	0.2	82.6	16.7	0.0	0.8
Education[1]								
< High School	81.2	17.8	1.0	0.0	80.9	12.8	6.4	0.0
High School	83.5	15.8	0.3	0.3	81.0	18.0	1.0	0.0
> High School	86.4	13.3	0.3	0.0	84.1	14.5	0.7	0.7

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997

Exhibit 6.6.20 Comparison of Debriefing Interview Respondents on Selected Debriefing Questi Interest (Unweighted Percentages)

	1997 Field Experiment						Comparison Group			
	CAPI/ACASI						1997 Quarter 4 PAPI/SAQ			
	On a scale from 1 to 5, where 1 is very boring and 5 is very interesting, how would you rate this interview?						On a scale from 1 to 5, where 1 is 5 is very interesting, how would interview that you completed			
Respondent Characteristics	Very Boring	Somewhat Boring	Neither Boring nor Interesting	Somewhat Interesting	Very Interesting	DK/ REF	Very Boring	Somewhat Boring	Neither Boring nor Interesting	Somewhat Interesting
Total	6.6	11.6	26.0	34.3	21.0	0.5	9.9	13.7	29.3	33.
Age Group										
12-17	7.8	12.4	27.7	30.8	20.8	0.5	11.0	17.5	31.8	27.
18+	4.9	10.7	23.9	38.9	21.3	0.4	8.9	9.9	26.7	38.
Gender										
Male	8.1	12.3	28.1	31.9	19.2	0.4	11.4	15.4	28.9	29.
Female	5.2	11.1	24.2	36.4	22.6	0.5	8.9	12.4	29.6	36.
Race/Ethnicity										
Hispanic	4.3	9.1	24.0	32.6	29.8	0.2	4.1	11.6	24.8	39.
Non-Hisp., Black	7.9	10.6	20.6	32.9	27.7	0.4	11.9	10.9	26.4	33.
Non-Hisp., All Other Races	6.9	13.4	29.9	35.9	13.2	0.6	11.1	16.7	33.3	30.
Education[1]										
< High School	4.6	9.6	14.7	35.5	35.5	0.0	12.8	6.4	19.1	36.
High School	6.3	10.4	24.4	35.8	22.5	0.6	7.0	5.0	24.0	44.
> High School	3.8	11.5	28.7	43.8	11.8	0.3	9.0	14.5	31.0	35.

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experim Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.21 Comparison of Debriefing Interview Respondents on Selected Debriefing Questions: Likelihood to Participate in Future (Unweighted Percentages)

	1997 Field Experiment					Comparison Group				
	CAPI/ACASI					1997 Quarter 4 PAPI/SAQ				
	If you were asked to take part in an interview like the one you just completed on the computer at some time in the future, how likely do you think you would be to participate?					If you were asked to take part in an interview like the one you just completed on paper at some time in the future, how likely do you think you would be to participate?				
Respondent Characteristics	Very Likely	Somewhat Likely	Somewhat Unlikely	Very Unlikely	DK/REF	Very Likely	Somewhat Likely	Somewhat Unlikely	Very Unlikely	DK
Total	41.5	38.4	9.9	9.9	0.3	35.4	39.7	13.9	10.6	
Age Group										
12-17	44.8	37.3	8.7	8.7	0.5	36.3	40.8	14.7	7.9	
18+	37.3	39.7	11.5	11.4	0.1	34.6	38.7	13.0	13.4	
Gender										
Male	41.0	37.8	10.9	10.0	0.3	33.3	39.4	15.4	11.4	
Female	41.9	38.8	9.2	9.8	0.3	37.0	39.9	12.7	10.1	
Race/Ethnicity										
Hispanic	42.8	39.9	9.9	8.0	0.0	45.5	36.4	11.6	6.6	
Non Hisp., Black	40.4	35.2	10.4	13.7	0.4	37.4	35.8	12.6	14.2	
Non Hisp., All Other Races	41.5	39.6	9.7	8.8	0.4	30.0	44.4	14.8	10.0	
Education[1]										
< High School	41.6	37.1	9.6	11.7	0.0	36.2	38.3	10.6	14.9	
High School	36.1	38.0	10.8	14.9	0.3	39.0	38.0	10.0	13.0	
> High School	35.8	42.9	13.3	8.0	0.0	31.0	39.3	15.9	13.1	

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experime
1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 6.6.22 Comparison of Debriefing Interview Respondents on Selected Debriefing Questic Confidence in Confidentiality (Unweighted Percentages)

	1997 Field Experiment				Comparison Gr		
	CAPI/ACASI				1997 Quarter PAPI/SAQ		
	How confident are you that your answers will never be linked to your name?				How confident are you that will never be linked to y		
Respondent Characteristics	Very	Somewhat	Not at All	DK/REF	Very	Somewhat	Not a
Total	65.8	28.2	5.6	0.4	58.0	32.9	
Age Group							
12-17	68.2	25.9	5.4	0.5	61.0	31.5	
18+	62.7	31.3	5.9	0.1	55.1	34.2	
Gender							
Male	66.2	26.6	5.9	0.3	60.6	30.5	
Female	65.5	28.8	5.3	0.4	56.2	34.6	
Race/Ethnicity							
Hispanic	67.0	27.4	5.4	0.2	61.2	33.9	
Non-Hisp., Black	65.1	27.3	7.3	0.4	57.0	29.5	
Non-Hisp., All Other Races	65.7	29.1	4.8	0.4	57.4	34.8	
Education[1]							
< High School	70.6	23.9	5.6	0.0	59.6	27.7	
High School	62.3	30.1	7.3	0.3	56.0	32.0	
> High School	58.6	36.7	4.7	0.0	53.1	37.9	

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experir Household Survey on Drug Abuse: Quarter 4.

[12] More detailed tables by treatment, age group, gender, race/ethnicity, and education can be found in Chapter 7.

[13] See Chapter 4 for a discussion of the cognitive laboratory testing of the 12-month frequency of use question.

[14] If asked, interviewers are trained to tell the respondent to record "DK" or "REF" in the right margin next to the particular item.

[15] In the ACASI application used, respondents could turn off the voice if they wished.

7. Effect of ACASI Experimental Factors on Prevalence and Data Quality: 1997 Field Experiment

In this chapter, we examine the effect of audio computer-assisted self-interviewing (ACASI) experimental factors on reported prevalence, inconsistencies in reporting, time required to complete the interview, and breakoff rates. In addition, we examine the effect that the experimental factors had on respondents' ability to complete the interview and their attitudes about the survey task. Sections 7.1 through 7.3 contain the descriptive results; in Sections 7.4 and 7.5, we present the results of the statistical tests.

Because all respondents did not answer the same questions, it was necessary to create a set of edited recency of use variables for constructing the lifetime, 12-month, and 30-day prevalence of use estimates. This was done in three stages. In the ACASI application, all respondents had multiple chances to report use. For example, even respondents who did not receive multiple use questions could report that they had used alcohol in the past 30 days and subsequently report that they had used it on 0 days in the past 30 days. And some respondents who were in the multiple use treatment category were asked twice about their use in a particular time period. Thus, we first created a series of recodes for 30-day and 12-month use that documented whether the respondent was consistent or inconsistent in reporting use in these time periods. For example, we created a variable MU30ALC that had the following four levels:

1 = consistent 30-day drinker;
2 = inconsistent 30-day drinker: "yes" to period; 0 days reported;
3 = inconsistent 30-day drinker: "no" to period; more than 0 days reported;[16] and
4 = consistent nonuser of alcohol in the past 30 days.

For the random half of the respondents who received consistency checks during the interview and who had inconsistent data (e.g., MU30ALC = 2 or 3), we created a variable that documented the outcome of the consistency check. This variable also had four levels based on the path through the questionnaire and classified the respondent as either a user or nonuser for the period in question. The edited recency variables were created from these intermediate recodes.

This process resulted in our inability to determine the recency of use of some respondents; therefore, we had a residual category of indeterminate answers. In the NHSDA series of surveys, a similar logical editing procedure is used; however, before the final estimates are made, all indeterminate recency of use answers are removed by statistical imputation. We did not use any statistical imputations for the field experiment analysis. Therefore, when we compared 1997 field experiment data with data from the 1997 NHSDA Quarter 4, we used the recency variable thatwas edited but not imputed because it was the most analogous paper-and pencil interview (PAPI) comparison variable.

For cigarettes, there is one final important difference between the field experiment and the 1997 NHSDA that affects the prevalence estimates. In the field experiment, we routed respondents who had only had one or two puffs of a cigarette to the next section after they had reported their age at first use.

In addition to the edited variables, we also produced tables of weighted estimates using *raw variables* that were created using the most direct question on recency of use. For example, respondents in Treatments 5, 6, 7, and 8 who answered "yes" to the question *"During the past 30 days, have you _____?"* were counted as 30-day users even if they later indicated having used 0 days in the past 30 days. For the comparison group, we used answers to the recency question *"How long has it been since you last used _____?"* to create the raw variables. Estimates using raw variables are given in RTI (1998).

Exhibit 7.1 summarizes the findings for the main effects of the ACASI experimental factors by showing the ratios of the prevalence estimates by the level of the main effects. In the following section, each factor is discussed separately, followed by the results of the modeling.

7.1 Structure of Contingent Questioning

As described in Chapter 5, a random half of the respondents received a single gate question and a random half received three gate questions. We hypothesized that using multiple gate questions would result in fewer inconsistent reports and higher reports of prevalence of use, particularly for lifetime and 12-month use. We speculated that lifetime prevalence might be higher because (a) respondents could first deny more proximate use, which other research (e.g., Turner, Lessler, & Devore, 1992a) has shown to be a more sensitive behavior, and (b) it would prevent routing respondents who inadvertently answered "no" to a single gate question to the next section, thereby missing a chance to gather information about their use. In addition, we felt that having

multiple gate questions that permitted the respondent to focus on each reference period individually would reduce inconsistencies in reporting.

7.1.1 Prevalence Estimates

Exhibits 7.1.1 through 7.1.15 display the weighted estimates of prevalence by selected demographic variables. For cigarette use, a mixed pattern was observed in the overall estimates, with single gate questions sometimes yielding higher reports and sometimes yielding lower reports. Differences between the two approaches were not large in any case, and most ratios were close to one. A similar pattern was observed for alcohol use. For marijuana, cocaine, and any illicit drug use, the single gate questions tended to yield higher reports of use. These findings were not in the direction hypothesized prior to the field experiment.

7.1.2 Inconsistencies in Reporting

We also examined inconsistencies in reporting by identifying the number of people who had a chance to give an inconsistent report. We confined this detailed analysis to cigarettes, alcohol, and marijuana so that we would have enough people to make meaningful comparisons. Exhibits 7.1.16 to 7.1.20, contrary to our expectations, indicate that multiple gate questions resulted in a higher percentages of inconsistent reports in most of the cases compared to single gate questions. The denominators (*N*) of the percentages included people who were able to get to the questions (i.e., they were not routed to the next section at a previous question), where they recorded the number of days of use within the particular time period. For past 12-month use, respondents answering the treatment versions with no multiple use questions had no chance of recording an inconsistent response.

7.1.3 Operational Aspects of Reporting Under Alternative Contingent Questioning Strategies

Time required to complete the interview. The average length of time required to complete each drug section with multiple gate questions was slightly higher than the time required to complete each drug section with a single gate question. However, the differences were not large (see Exhibit 7.1.21).

Exhibit 7.1 Ratio of Prevalence Estimates for ACASI Experimental Factors Overall and by Age Group

	Overall			18+ Years Old			12 to 17 Years Old		
Factor	Lifetime	12 Month	30 Day	Lifetime	12 Month	30 Day	Lifetime	12 Month	30 Day
Cigarettes									
Multiple Gate Qs: Single/Multiple	1.07	0.87	1.03	1.08	0.87	1.03	0.94	0.99	1.07
Multiple 30-Day and 12-Month Qs:	1.06	0.88	0.70	1.05	0.84	0.67	1.14	1.48	1.16
Absent/Present Consistency Checks: Absent/Present	0.94	0.84	0.77	0.93	0.82	0.76	1.05	1.09	0.92
Alcohol									
Multiple Gate Qs: Single/Multiple	0.97	0.90	0.94	0.97	0.90	0.94	1.00	1.00	1.05
Multiple 30-Day and 12-Month Qs:	1.01	0.94	1.07	1.00	0.92	1.06	1.09	1.07	1.12
Absent/Present Consistency Checks: Absent/Present	0.97	0.91	0.85	1.03	1.03	0.89	0.97	0.90	0.85
Marijuana									
Multiple Gate Qs: Single/Multiple	1.01	1.63	1.45	1.00	1.76	1.48	1.18	1.12	1.32
Multiple 30-Day and 12-Month									

Qs:	0.90	0.80	1.31	0.88	1.76	1.31	1.14	1.48	1.16
Absent/Present Consistency Checks: Absent/Present	0.99	0.94	0.64	0.99	0.94	0.59	1.02	0.94	0.85
Cocaine									
Multiple Gate Qs: Single/Multiple	1.93	2.64	1.13	1.94	2.76	1.00	1.92	1.86	3.77
Multiple 30-Day and 12-Month Qs:	0.68	1.47	0.98	0.66	1.51	1.18	0.93	0.98	0.19
Absent/Present Consistency Checks: Absent/Present	0.62	0.75	1.09	0.63	0.74	0.95	0.74	0.96	4.68
Any Illicit Drug[1]									
Multiple Gate Qs: Single/Multiple	1.14	1.29	1.45	1.13	1.34	1.51	1.18	1.06	1.23
Multiple 30-Day and 12-Month Qs:	0.90	0.92	1.30	0.89	0.90	1.35	1.11	1.10	1.21
Absent/Present Consistency Checks: Absent/Present	0.99	0.92	0.64	0.99	0.93	0.60	0.98	0.84	0.76

[1]Any illicit drug includes marijuana, cocaine, crack, heroin, inhalants, hallucinogens, and nonmedical use of analgesics, sedatives, stimulants, and tranquilizers.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures: 1997 Field Experiment.

Breakoff rates. In Exhibit 7.1.22, we compared the percentage of cases that were reported as a breakoff (even if the interviewer went on to finalize the case as a completed interview) by the treatment versions (single gate vs. multiple gate questions). Then we compared the percentage of finalized breakoff cases among single rate and multiple gate questions. The data show no meaningful difference between the breakoff rates for single gate and multiple gate questions.

Respondent reactions as reported in the debriefing questionnaire. Exhibit 7.1.23 presents an analysis of items from the respondent debriefing questions to see whether multiple gate questions had any effect on respondent reactions to the computer-assisted interviewing (CAI) NHSDA interview. The exhibit includes data from five items. About 87% of the respondents who were administered the multiple gate questions indicated that they were able to enter answers easily into the computer compared to 90% of the respondents who were administered the single gate questions. There was no important difference between the responses for single gate and multiple gate questions for the other items, including respondents' difficulty in using the computer, their interest in the interview, and the accuracy and completeness of their answers.

Detailed Exhibits for Section 7.1

Exhibit 7.1.1 Lifetime Use of Cigarettes, by Contingent Questioning Structure

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	74.18%	69.38%	71.96%	70.87%
Age Group				
12 to 17	43.20%	46.04%	44.48%	36.05%
18+	77.88%	72.04%	75.18%	74.94%
Gender				
Male	75.71%	72.64%	74.25%	74.74%
Female	72.82%	66.14%	69.83%	67.25%
Race/Ethnicity				
Hispanic	64.34%	63.15%	63.68%	55.82%
Non-Hisp., Black	62.65%	66.11%	64.30%	61.62%
Non-Hisp., All Other Races	76.82%	70.65%	74.04%	73.81%
Education[1]				
< High School	84.90%	68.25%	77.83%	70.28%
High School	78.43%	81.58%	79.73%	76.65%
> High School	75.01%	67.77%	71.35%	75.27%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.2 Lifetime Use of Alcohol, by Contingent Questioning Structure

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	82.01%	84.50%	83.16%	82.15%
Age Group				
12 to 17	44.89%	44.92%	44.90%	38.01%
18+	86.45%	89.01%	87.63%	87.31%
Gender				
Male	80.10%	88.20%	83.95%	87.07%
Female	83.70%	80.81%	82.41%	77.55%
Race/Ethnicity				
Hispanic	70.87%	77.01%	74.26%	69.45%
Non-Hisp., Black	65.91%	65.13%	65.54%	72.35%
Non-Hisp., All Other Races	85.51%	88.92%	87.04%	84.98%
Education[1]				
< High School	84.79%	78.92%	82.30%	77.22%
High School	81.23%	89.67%	84.73%	85.83%
> High School	91.16%	91.17%	91.16%	91.48%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.3 Lifetime Use of Marijuana, by Contingent Questioning Structure

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	35.36%	34.98%	35.18%	34.96%
Age Group				
12 to 17	22.08%	18.70%	20.56%	16.10%
18+	36.94%	36.84%	36.90%	37.17%
Gender				
Male	39.35%	36.06%	37.78%	40.79%
Female	31.81%	33.91%	32.75%	29.51%
Race/Ethnicity				
Hispanic	38.98%	39.12%	39.06%	24.59%
Non-Hisp., Black	27.72%	34.49%	30.95%	31.43%
Non-Hisp., All Other Races	36.39%	34.63%	35.60%	36.48%
Education[1]				
< High School	26.55%	29.61%	27.85%	32.84%
High School	36.22%	35.84%	36.06%	36.60%
> High School	41.12%	39.19%	40.15%	38.91%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.4 Lifetime Use of Cocaine, by Contingent Questioning Structure				
	ACASI Treatment Groups			Comparison Group
Respondent Characteristics	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	19.18%	9.93%	14.91%	10.03%
Age Group				
12 to 17	3.64%	1.90%	2.86%	3.13%
18+	21.04%	10.84%	16.32%	10.84%
Gender				
Male	26.68%	9.70%	18.60%	12.47%
Female	12.53%	10.15%	11.46%	7.76%
Race/Ethnicity				
Hispanic	15.81%	18.52%	17.30%	7.32%
Non-Hisp., Black	15.27%	4.02%	9.90%	7.57%
Non-Hisp., All Other Races	20.08%	10.12%	15.59%	10.70%
Education[1]				
< High School	36.36%	8.47%	24.51%	9.52%
High School	17.54%	9.64%	14.27%	10.37%
> High School	18.50%	12.09%	15.26%	11.58%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.5 Lifetime Use of Any Illicit Drug, by Contingent Questioning Structure

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	46.26%	40.75%	43.72%	37.50%
Age Group				
12 to 17	33.48%	28.32%	31.15%	20.45%
18+	47.79%	42.17%	45.19%	39.49%
Gender				
Male	49.68%	37.16%	43.72%	43.44%
Female	43.23%	44.32%	43.72%	31.94%
Race/Ethnicity				
Hispanic	46.31%	42.58%	44.25%	28.63%
Non-Hisp., Black	34.56%	38.82%	36.59%	33.40%
Non-Hisp., All Other Races	48.23%	40.92%	44.93%	38.99%
Education[1]				
< High School	54.09%	32.81%	45.05%	37.22%
High School	46.31%	44.03%	45.36%	38.42%
> High School	46.79%	43.49%	45.12%	40.92%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.6 Past 12-Month Use of Cigarettes, by Contingent Questioning Structure

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	29.24%	33.46%	31.19%	33.52%
Age Group				
12 to 17	22.74%	23.05%	22.88%	25.18%
18+	30.02%	34.65%	32.16%	34.49%
Gender				
Male	38.14%	33.32%	35.85%	34.17%
Female	21.34%	33.59%	26.82%	32.91%
Race/Ethnicity				
Hispanic	34.51%	23.70%	28.55%	32.69%
Non-Hisp., Black	24.35%	36.07%	29.94%	35.19%
Non-Hisp., All Other Races	29.69%	34.01%	31.64%	33.29%
Education[1]				
< High School	57.88%	28.04%	45.20%	34.76%
High School	24.86%	51.87%	36.04%	41.55%
High School	24.44%	26.87%	25.67%	29.70%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.7 Past 12-Month Use of Alcohol, by Contingent Questioning Structure				
	ACASI Treatment Groups			Comparison Group
Respondent Characteristics	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	65.11%	71.99%	68.28%	64.95%
Age Group				
12 to 17	36.42%	36.49%	36.45%	32.74%
18+	68.54%	76.03%	72.00%	68.71%
Gender				
Male	62.83%	77.43%	69.78%	70.07%
Female	67.12%	66.58%	66.88%	60.15%
Race/Ethnicity				
Hispanic	54.39%	71.17%	63.65%	57.10%
Non-Hisp., Black	44.86%	56.29%	50.31%	53.54%
Non-Hisp., All Other Races	69.27%	75.01%	71.86%	67.64%
Education[1]				
< High School	57.94%	50.08%	54.60%	48.54%
High School	66.10%	77.46%	70.80%	68.12%
> High School	74.15%	81.74%	77.99%	75.47%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.8 Past 12-Month Use of Marijuana, by Contingent Questioning Structure				
	ACASI Treatment Groups			Comparison Group
Respondent Characteristics	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	12.47%	7.64%	10.24%	9.36%
Age Group				
12 to 17	17.56%	15.62%	16.69%	12.97%
18+	11.86%	6.73%	9.49%	8.94%
Gender				
Male	17.80%	9.62%	13.91%	13.08%
Female	7.75%	5.68%	6.82%	5.88%
Race/Ethnicity				
Hispanic	10.71%	16.66%	13.99%	10.21%
Non-Hisp., Black	8.84%	7.73%	8.31%	9.60%
Non-Hisp., All Other Races	13.21%	6.67%	10.26%	9.25%
Education[1]				
< High School	10.89%	4.19%	8.05%	14.87%
High School	9.05%	7.95%	8.59%	10.10%
> High School	14.43%	6.70%	10.52%	6.30%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.9 Past 12-Month Use of Cocaine, by Contingent Questioning Structure

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	3.69%	1.40%	2.63%	1.60%
Age Group				
12 to 17	2.23%	1.20%	1.77%	2.54%
18+	3.87%	1.42%	2.74%	1.49%
Gender				
Male	4.80%	1.16%	3.07%	2.10%
Female	2.71%	1.64%	2.23%	1.13%
Race/Ethnicity				
Hispanic	6.72%	2.91%	4.62%	1.56%
Non-Hisp., Black	4.31%	2.41%	3.40%	2.98%
Non-Hisp., All Other Races	3.37%	1.05%	2.33%	1.36%
Education[1]				
< High School	7.53%	2.01%	5.19%	1.87%
High School	1.59%	2.25%	1.86%	2.03%
> High School	4.41%	0.82%	2.59%	1.01%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.10 Past 12-Month Use of Any Illicit Drug, by Contingent Questioning Structure				
	ACASI Treatment Groups			Comparison Group
Respondent Characteristics	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	16.92%	13.10%	15.16%	10.28%
Age Group				
12 to 17	23.70%	22.22%	23.03%	14.54%
18+	16.10%	12.07%	14.24%	9.78%
Gender				
Male	20.97%	11.04%	16.24%	13.87%
Female	13.32%	15.16%	14.14%	6.93%
Race/Ethnicity				
Hispanic	17.84%	22.11%	20.20%	12.33%
Non-Hisp., Black	14.43%	10.90%	12.74%	9.99%
Non-Hisp., All Other Races	17.27%	12.56%	15.15%	10.16%
Education[1]				
< High School	15.60%	6.03%	11.53%	14.57%
High School	11.81%	16.32%	13.68%	10.63%
> High School	19.69%	11.25%	15.41%	7.71%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.11 Past 30-Day Use of Cigarettes, by Contingent Questioning Structure

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	26.61%	25.73%	26.20%	30.45%
Age Group				
12 to 17	16.62%	15.49%	16.11%	18.56%
18+	27.80%	26.89%	27.38%	31.84%
Gender				
Male	34.95%	21.26%	28.43%	31.10%
Female	19.20%	30.17%	24.11%	29.84%
Race/Ethnicity				
Hispanic	30.09%	21.46%	25.33%	28.86%
Non-Hisp., Black	18.94%	31.21%	24.79%	33.03%
Non-Hisp., All Other Races	27.65%	25.15%	26.53%	30.13%
Education[1]				
< High School	57.32%	26.00%	44.01%	34.23%
High School	23.07%	45.24%	32.25%	38.72%
> High School	21.31%	17.08%	19.17%	26.50%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.12 Past 30-Day Use of Alcohol, by Contingent Questioning Structure

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	47.35%	50.27%	48.70%	52.21%
Age Group				
12 to 17	17.43%	16.60%	17.06%	18.84%
18+	50.93%	54.11%	52.40%	56.11%
Gender				
Male	47.93%	54.42%	51.02%	60.23%
Female	46.83%	46.15%	46.53%	44.70%
Race/Ethnicity				
Hispanic	38.07%	37.65%	37.84%	40.79%
Non-Hisp., Black	27.70%	43.45%	35.21%	41.19%
Non-Hisp., All Other Races	51.31%	52.89%	52.02%	55.15%
Education[1]				
< High School	38.35%	20.71%	30.85%	40.55%
High School	43.74%	53.18%	47.65%	55.62%
> High School	60.99%	62.97%	61.99%	61.34%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.13 Past 30-Day Use of Marijuana, by Contingent Questioning Structure				
	ACASI Treatment Groups			Comparison Group
Respondent Characteristics	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	5.54%	3.82%	4.74%	5.28%
Age Group				
12 to 17	9.39%	7.12%	8.37%	7.30%
18+	5.08%	3.44%	4.32%	5.04%
Gender				
Male	8.24%	5.66%	7.01%	7.30%
Female	3.14%	1.99%	2.62%	3.39%
Race/Ethnicity				
Hispanic	6.04%	1.95%	3.79%	4.48%
Non-Hisp., Black	4.94%	4.10%	4.54%	6.05%
Non-Hisp., All Other Races	5.60%	3.96%	4.86%	5.21%
Education[1]				
< High School	6.24%	3.44%	5.05%	10.13%
High School	4.94%	3.16%	4.21%	5.92%
> High School	4.78%	3.59%	4.18%	2.86%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.14 Past 30-Day Use of Cocaine, by Contingent Questioning Structure

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	0.60%	0.53%	0.57%	0.59%
Age Group				
12 to 17	0.98%	0.26%	0.66%	1.09%
18+	0.56%	0.56%	0.56%	0.53%
Gender				
Male	0.51%	0.35%	0.43%	0.61%
Female	0.68%	0.71%	0.69%	0.56%
Race/Ethnicity				
Hispanic	1.79%	1.09%	1.40%	0.80%
Non-Hisp., Black	0.70%	1.39%	1.03%	2.09%
Non-Hisp., All Other Races	0.50%	0.31%	0.41%	0.30%
Education[1]				
< High School	2.19%	0.92%	1.65%	0.87%
High School	0.54%	0.84%	0.67%	0.89%
> High School	0.00%	0.31%	0.16%	0.18%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.15 Past 30-Day Use of Any Illicit Drug, by Contingent Questioning Structure				
	ACASI Treatment Groups			Comparison Group
Respondent Characteristics	Single Gate Questions	Multiple Gate Questions		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 3, 4	5, 6, 7, 8	Total ACASI	
Total	6.91%	4.75%	5.91%	5.44%
Age Group				
12 to 17	13.02%	10.63%	11.95%	7.98%
18+	6.18%	4.08%	5.21%	5.14%
Gender				
Male	9.30%	6.89%	8.15%	7.68%
Female	4.80%	2.62%	3.82%	3.34%
Race/Ethnicity				
Hispanic	7.30%	3.16%	5.02%	4.60%
Non-Hisp., Black	6.76%	6.76%	6.76%	6.45%
Non-Hisp., All Other Races	6.91%	4.54%	5.84%	5.33%
Education[1]				
< High School	7.66%	4.42%	6.28%	9.46%
High School	5.89%	3.97%	5.09%	5.94%
> High School	5.90%	4.05%	4.97%	3.25%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.1.16 Past 30-Day Use of Cigarettes: Percentage of Inconsistent Respondents										
	Single Gate Questions				Multiple Gate Questions					
	Multiple Use Questions Absent Treatment Versions: 1, 3 (*n*=110)		Multiple Use Questions Present Treatment Versions: 2, 4 (*n*=152)		Multiple Use Questions Absent Treatment Versions: 5, 7 (*n*=106)		Multiple Use Questions Present Treatment Versions: 6, 8 (*n*=128)		Total ACASI (*n*=496)	
Type of Inconsistency	Sample	%	Sample	%	Sample	%	Sample	%	Sample	%
Total	2	1.82	12	7.90	6	5.66	14	10.94	34	6.86
"Yes" to Period; 0 Days Reported	2	1.82	5	3.29	6	5.66	14	10.94	27	5.44
"No" to Period; Greater Than 0 Days Reported	N/A		7	4.61	N/A		0	0.00	7	1.41

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.1.17 Past 30-Day Use of Alcohol: Percentage of Inconsistent Respondents

	Single Gate Questions				Multiple Gate Questions					
	Multiple Use Questions Absent Treatment Versions: 1, 3 (*n*=171)		Multiple Use Questions Present Treatment Versions: 2, 4 (*n*=265)		Multiple Use Questions Absent Treatment Versions: 5, 7 (*n*=172)		Multiple Use Questions Present Treatment Versions: 6, 8 (*n*=213)		Total ACASI (*n*=821)	
Type of Inconsistency	Sample	%	Sample	%	Sample	%	Sample	%	Sample	%
Total	11	6.43	25	9.43	19	11.05	26	12.21	81	9.87
"Yes" to Period; 0 Days Reported	11	6.43	14	5.28	19	11.05	13	6.10	57	6.94
"No" to Period; Greater Than 0 Days Reported	N/A		11	4.15	N/A		13	6.10	24	2.92

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.1.18 Past 30-Day Use of Marijuana: Percentage of Inconsistent Respondents

	Single Gate Questions				Multiple Gate Questions					
	Multiple Use Questions Absent Treatment Versions: 1, 3 (*n*=49)		Multiple Use Questions Present Treatment Versions: 2, 4 (*n*=85)		Multiple Use Questions Absent Treatment Versions: 5, 7 (*n*=39)		Multiple Use Questions Present Treatment Versions: 6, 8 (*n*=64)		Total ACASI (*n*=237)	
Type of Inconsistency	Sample	%	Sample	%	Sample	%	Sample	%	Sample	%
Total	1	2.04	9	10.59	6	15.40	7	10.94	23	9.71
"Yes" to Period; 0 Days Reported	1	2.04	0	0.00	6	15.40	3	4.69	10	4.22
"No" to Period; Greater Than 0 Days Reported	N/A		9	10.59	N/A		4	6.25	13	5.49

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.1.19 Past 12-Month Use of Alcohol: Percentage of Inconsistent Respondents

	Single Gate Questions				Multiple Gate Questions					
	Multiple Use Questions Absent Treatment Versions: 1, 3 (*n*=0)		Multiple Use Questions Present Treatment Versions: 2, 4 (*n*=321)		Multiple Use Questions Absent Treatment Versions: 5, 7 (*n*=0)		Multiple Use Questions Present Treatment Versions: 6, 8 (*n*=247)		Total ACASI (*n*=568)	
Type of Inconsistency	Sample	%	Sample	%	Sample	%	Sample	%	Sample	%
Total	N/A		30	9.35	N/A		29	11.74	59	10.39
"Yes" to Period; 0 Days Reported	N/A		18	5.61	N/A		20	8.10	38	6.69
"No" to Period; Greater Than 0 Days Reported	N/A		12	3.74	N/A		9	3.64	21	3.70

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.1.20 Past 12-Month Use of Marijuana: Percentage of Inconsistent Respondents										
	Single Gate Questions				Multiple Gate Questions					
	Multiple Use Questions Absent Treatment Versions: 1, 3 (*n*=0)		Multiple Use Questions Present Treatment Versions: 2, 4 (*n*=110)		Multiple Use Questions Absent Treatment Versions: 5, 7 (*n*=0)		Multiple Use Questions Present Treatment Versions: 6, 8 (*n*=80)		Total ACASI (*n*=190)	
Type of Inconsistency	Sample	%	Sample	%	Sample	%	Sample	%	Sample	%
Total	NA		14	12.73	NA		12	15.00	26	13.68
"Yes" to Period; 0 Days Reported	NA		3	2.73	NA		6	7.50	9	4.74
"No" to Period; Greater Than 0 Days Reported	NA		11	10.00	NA		6	7.50	17	8.95

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.1.21 Average Time to Complete Treatment Sections, by Old and New Treatment Versions				
	Single Gate Questions (Treatment Versions: 1, 2, 3, 4)		Multiple Gate Questions (Treatment Versions: 5, 6 ,7, 8)	
Questionnaire Section	# of Cases	Time (minutes)	# of Cases	Time (minutes)
Tobacco	1,068	1.49	909	1.71
Alcohol	1,069	3.20	907	3.36
Marijuana	1,069	0.80	907	1.02
Cocaine	1,069	0.35	907	0.59
Crack	633	0.28	467	0.54
Heroin	1,069	0.19	907	0.39
Hallucinogens	1,068	1.48	907	1.42
Inhalants	1,066	2.44	906	2.41
Total for Treatment Sections	1,071	10.23	911	11.40

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.1.22 Breakoff Rates, by Old and New Treatment Versions				
	Single Gate Questions Treatment Versions: 1, 2, 3, 4		Multiple Gate Questions Treatment Versions: 5, 6, 7, 8	
Cases Included	*n*	%	*n*	%
Cases Ever Recorded as Breakoff Even if Completed Later	89	8.31	74	8.12
Cases Finalized as Breakoff	16	1.49	16	1.76
Total Number of Respondents Assigned to Treatment Versions	1,071		911	

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.1.23 Respondent Reactions to CAI Interview, by Old and New Treatment Versions					
		Single Gate Questions (Treatment Versions: 1, 2, 3, 4)		Multiple Gate Questions (Treatment Versions: 5, 6, 7, 8)	
Item		*n*	%	*n*	%
How Difficult to Use Computer?					
	Very difficult	22	2.08	20	2.23
	Somewhat difficult	75	7.10	49	5.47
	Not at all difficult	959	90.81	827	92.30
Total		1,056	100.00	896	100.00
Able to Enter Answers Easily?					
	Yes	949	89.78	775	86.50
	No	108	10.22	121	13.50
Total		1,057	100.00	896	100.00
Wanted to Change Previous Answer But Did Not?					
	Yes	195	18.52	162	18.14
	No	858	81.48	731	81.86
Total		1,053	100.00	893	100.00
	Level of interest in interview				
	Very boring	69	6.58	59	6.61
	Somewhat boring	116	11.06	111	12.44
	Neither boring nor interesting	283	26.98	224	25.11
	Somewhat interesting	364	34.70	305	34.19
	Very interesting	217	20.69	193	21.64
Total		1,049	100.0	892	100.00
Accuracy and Completeness of Answers					
	Very accurate and complete	872	82.81	748	83.76
	Fairly accurate and complete	171	16.24	133	14.89
	Not very accurate and complete	10	0.95	12	1.34
Total		1,053	100.00	893	100.00

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

7.2 Multiple Opportunities to Report Use

The idea behind allowing respondents more than one opportunity to report drug use was that respondents who were initially unwilling to report more frequent use might change their minds and report the usage in a later question. Respondents who reported use in the past 12 months were given a second opportunity to report use in the past 30 days by being routed to the 30-day frequency question. Respondents who reported use more than 12 months ago (but within the past 3 years) were given the opportunity to report use in the past 12 months by being routed to the 12-month frequency question. Subsequently, respondents who indicated 12-month use were then also routed to the 30-day frequency item. Respondents who reported use more than 3 years ago were not included in any of the multiple use routing schemes.

7.2.1 Prevalence Estimates

Exhibits 7.2.1 through 7.2.15 compare prevalence estimates between the two multiple use treatments for lifetime drug use, past 12-month use, and past 30-day use for cigarettes, alcohol, marijuana, cocaine, and any illicit drug use. In these exhibits, comparable data are included for the 1997 Quarter 4 PAPI data collection. The results do not show a consistent increase in drug use reporting for the inclusion of the multiple questioning approach. For drugs other than tobacco and alcohol, lifetime prevalence increased under the multiple use approach, but for 12-month and 30-day use, many of the drug prevalences appeared to decrease under the multiple questioning approach. Although sample sizes for some of the drugs were quite small and the associated prevalence estimates may be unstable, it seems that the multiple use approach may have introduced random variation (or "noise") into the data rather than reflect any true increase in reporting.

It should be noted, however, that the prevalence estimates in the multiple questioning approach can be manipulated by the rules used to classify the recency of use. Rules must be defined to determine how to classify each respondent based on his or her recency of use. These rules can become quite complex, particularly when combined with the consistency check treatment. In some cases, a respondent may give highly inconsistent answers, and the ensuing prevalence estimates can be generated using either a very conservative or a very liberal (or anywhere in between) classification scheme. A conservative scheme would require a respondent's data to be entirely consistent before his or her data were included in the prevalence analysis. A liberal scheme would allow for a "preponderance of the evidence" rule, which would categorize the respondent based on what the majority of his or her data indicate. For the prevalence exhibits presented here, we have chosen to use a conservative classification scheme so that the prevalence estimates will be somewhat lower than would be obtained with a more liberal classification scheme.

To illustrate how editing rules affect the results, we also created a series of weighted estimates using different editing rules. In Exhibit 7.2.16, we compare theestimates for 30-day use for youths and adults. These exhibits display the estimates using

1. the *raw variables*, which make use of the most direct question on use during the 30-day period;
2. a *most liberal* editing rule, which counts a person as a 30-day user if he or she gave any indication of use, that is,
    1. provided consistent reports of 30-day use,
    2. indicated that his or her most recent use was in the 30-day period but reported using 0 days in the past 30 days, or
    3. indicated that his or her most recent use was more than 30 days ago but reported using more than 0 days in the past 30 days;
3. and an *approximate NHSDA* editing rule, which counted some inconsistent reporters as 30-day users.

Under our *edited variables*, inconsistent 30-day users were only counted as 30-day users if they had not been consistent in reporting 12-month use.

Examining the results, we note that using liberal editing rules leads to higher reports than using either the conservative editing rules we have used or the rule that is similar to the NHSDA, particularly for those treatments employing multiple use questions. Also, it is interesting to note that the most liberal editing rule tended to track the estimates from the raw variables. This is because the number of people who reported that they used in the past 30 days and who subsequently reported 0 days of use tended to be balanced by the number who said they did not use in the past 30 days and who subsequently reported some days of use.[17]

7.2.2 Inconsistencies in Reporting

When examining the utility of the multiple questioning approach, it is useful to consider how many respondents reported more recent use as a result of being routed to a question that gave them an opportunity to do so. Exhibit 7.2.17 provides these data in two ways. It was possible that respondents could report more recent use when given a second chance to report, and there was also the possibility that respondents who initially reported 12-month or 30-day use might negate this answer as part of the multiple use questioning. For example, a respondent who reported use of marijuana during the past 12 months would then be routed to the series of questions on frequency of use in the past 12 months. This series of questions includes a response option for "no use during the past 12 months" (because some respondents would be routed to this question after reporting use more than 12 months ago but during the past 3 years). Ifthis respondent then selects the "no use" option, he or she is routed past the 12-month frequency question. So, the multiple use feature has the ability to create more recent users as well as to negate users.

Both cases are observed in Exhibit 7.2.17. The data in this exhibit are quite interesting. A total of 141 respondents gave at least one inconsistent response as a result of the multiple use treatment. Based on an overall sample size of 1,025 respondents assigned to the multiple use treatment, this means approximately 14% of the sample was affected by the multiple use treatment. On average across the four multiple use treatments, approximately two thirds of these respondents indicated more recent use, while a third suggested their use was less recent.

We also examined the inconsistencies in reporting 30-day use for alcohol, cigarettes, marijuana, hallucinogens, and inhalants by determining the proportion of consistent reporters among those who gave any indication of use in the past 30 days. We examined the PAPI data and the analysis was confined to data for the ACASI respondents in the multiple use treatment. There is only a slight indication that the ACASI respondents gave more consistent reports (data not shown in an exhibit). For alcohol, 85% of the respondents who gave any indication of use in the past 30 days were consistent reporters of use in the past 30 days for both the PAPI comparison group and the ACASI multiple use treatment group; for cigarettes, 86% of the PAPI respondents and 89% of the ACASI respondents were consistent reporters. For marijuana, the corresponding figures are 69% for the PAPI respondents and 82% for the ACASI respondents. For hallucinogens, only 37% and 38% of the respondents in PAPI and ACASI, respectively, gave consistent reports. For inhalants, among those who gave any indication of use in the past 30 days, 58% of the PAPI respondents and 43% of the ACASI respondents gave consistent reports.

These large inconsistencies in reporting indicate that more work still needs to be done to increase the saliency of the reference periods for the respondents and to assist them in deciding when an event is in or out of the reference period. For rarely used drugs, some respondents may be reporting the total number of times that they have used the substance, not the number of times in the past 30 days.

7.2.3 Operational Aspects of Reporting Under Multiple Opportunities to Report Use

Time required to complete the interview. Exhibit 7.2.18 provides descriptive data on the average length of time required to complete the treatment section of the 1997 field experiment instrument when the multiple use treatment was and was not in use. Data are included for each individual drug section and for the treatment section as a whole. From this exhibit, we see that the average time to complete each drug section was nearly identical between the two multiple use conditions. The overall average times were identical as well. It appears that the additional questions respondents must answer in the multiple use treatment have little, if any, affect on the overall length of theinterview. More detailed analysis of timing data from the field experiment are presented in Section 7.5.

Breakoff rates. We next compared the rate of incomplete interviews (commonly known as "breakoffs") by analyzing whether the multiple use treatment was in use. Breakoff cases accounted for only a small number of the total number of interviews completed for the 1997 field experiment. However, if the multiple use treatment had caused respondents not to complete the full interview, we would view this as a serious drawback of the methodology. We analyzed the breakoff data in two ways. First, we compared the percentage of cases ever reported as a breakoff (even if the interviewer went on to finalize the case as a completed interview) by multiple use status. Then, we compared the percentage of finalized breakoff cases by multiple use status. Both sets of data are presented in Exhibit 7.2.19. The data show fundamentally that the multiple use treatment had no effect on the breakoff rate for the 1997 field experiment.

Respondent reactions as reported in the debriefing questionnaire. Finally, we analyzed items from the respondent debriefing questions to see whether the multiple use treatment had any effect on respondent reactions to the CAI NHSDA interview. Exhibit 7.2.20 includes data from five items included in the respondent debriefing that we thought might vary by treatment version. Specifically, these items asked respondents to rate their difficulty in using the computer, their interest in the interview, and the accuracy and completeness of their answers. In addition, respondents were asked whether they were able to enter their answers easily and whether they wanted to change an answer to a previous question but did not. The data in Exhibit 7.2.20 show that the multiple use treatment had no effect on respondents' perceptions of the interview as measured by these five items.

Detailed Exhibits for Section 7.2

Exhibit 7.2.1 Lifetime Use of Cigarettes, by Opportunities to Report Use

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	73.87%	69.92%	71.96%	70.87%
Age Group				
12 to 17	47.37%	41.73%	44.48%	36.05%
18+	76.77%	73.45%	75.18%	74.94%
Gender				
Male	77.52%	70.61%	74.25%	74.74%
Female	70.34%	69.29%	69.83%	67.25%
Race/Ethnicity				
Hispanic	64.80%	62.62%	63.68%	55.82%
Non-Hisp., Black	72.83%	54.21%	64.30%	61.62%
Non-Hisp., All Other Races	74.81%	73.22%	74.04%	73.81%
Education[1]				
< High School	79.66%	76.50%	77.83%	70.28%
High School	78.03%	81.71%	79.73%	76.65%
> High School	75.27%	66.72%	71.35%	75.27%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.2 Lifetime Use of Alcohol, by Opportunities to Report Use

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	83.53%	82.75%	83.16%	82.15%
Age Group				
12 to 17	46.93%	42.97%	44.90%	38.01%
18+	87.53%	87.74%	87.63%	87.31%
Gender				
Male	84.98%	82.81%	83.95%	87.07%
Female	82.12%	82.70%	82.41%	77.55%
Race/Ethnicity				
Hispanic	72.09%	76.34%	74.26%	69.45%
Non-Hisp., Black	68.01%	62.61%	65.54%	72.35%
Non-Hisp., All Other Races	87.35%	86.71%	87.04%	84.98%
Education[1]				
< High School	81.52%	82.86%	82.30%	77.22%
High School	86.44%	82.74%	84.73%	85.83%
> High School	89.65%	92.95%	91.16%	91.48%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.3 Lifetime Use of Marijuana, by Opportunities to Report Use				
Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present	Total ACASI	1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8		
Total	33.38%	37.13%	35.18%	34.96%
Age Group				
12 to 17	21.79%	19.38%	20.56%	16.10%
18+	34.64%	39.35%	36.90%	37.17%
Gender				
Male	36.19%	39.56%	37.78%	40.79%
Female	30.65%	34.93%	32.75%	29.51%
Race/Ethnicity				
Hispanic	35.41%	42.55%	39.06%	24.59%
Non-Hisp., Black	37.18%	23.57%	30.95%	31.43%
Non-Hisp., All Other Races	32.50%	38.90%	35.60%	36.48%
Education[1]				
< High School	25.49%	29.56%	27.85%	32.84%
High School	35.58%	36.62%	36.06%	36.60%
> High School	36.14%	44.87%	40.15%	38.91%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.4 Lifetime Use of Cocaine, by Opportunities to Report Use				
Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present	Total ACASI	1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8		
Total	12.11%	17.92%	14.91%	10.03%
Age Group				
12 to 17	2.74%	2.96%	2.86%	3.13%
18+	13.13%	19.80%	16.32%	10.84%
Gender				
Male	13.17%	24.65%	18.60%	12.47%
Female	11.08%	11.86%	11.46%	7.76%
Race/Ethnicity				
Hispanic	13.78%	20.67%	17.30%	7.32%
Non-Hisp., Black	12.79%	6.48%	9.90%	7.57%
Non-Hisp., All Other Races	11.84%	19.59%	15.59%	10.70%
Education[1]				
< High School	14.45%	31.82%	24.51%	9.52%
High School	16.02%	12.25%	14.27%	10.37%
> High School	10.91%	20.39%	15.26%	11.58%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.5 Lifetime Use of Any Illicit Drug, by Opportunities to Report Use

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	41.59%	46.00%	43.72%	37.50%
Age Group				
12 to 17	32.80%	29.59%	31.15%	20.45%
18+	42.56%	48.06%	45.19%	39.49%
Gender				
Male	38.83%	49.18%	43.72%	43.44%
Female	44.27%	43.14%	43.72%	31.94%
Race/Ethnicity				
Hispanic	39.99%	48.32%	44.25%	28.63%
Non-Hisp., Black	43.67%	28.22%	36.59%	33.40%
Non-Hisp., All Other Races	41.34%	48.77%	44.93%	38.99%
Education[1]				
< High School	31.64%	54.79%	45.05%	37.22%
High School	46.50%	44.04%	45.36%	38.42%
> High School	42.47%	48.25%	45.12%	40.92%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.6 Past 12-Month Use of Cigarettes, by Opportunities to Report Use				
Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present	Total ACASI	1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8		
Total	29.20%	33.32%	31.19%	33.52%
Age Group				
12 to 17	27.42%	18.55%	22.88%	25.18%
18+	29.39%	35.17%	32.16%	34.49%
Gender				
Male	34.81%	37.01%	35.85%	34.17%
Female	23.77%	30.00%	26.82%	32.91%
Race/Ethnicity				
Hispanic	34.22%	23.13%	28.55%	32.69%
Non-Hisp., Black	41.84%	15.85%	29.94%	35.19%
Non-Hisp., All Other Races	26.44%	37.18%	31.64%	33.29%
Education[1]				
< High School	35.70%	52.11%	45.20%	34.76%
High School	26.67%	46.89%	36.04%	41.55%
> High School	29.73%	20.88%	25.67%	29.70%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.7 Past 12-Month Use of Alcohol, by Opportunities to Report Use

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	66.16%	70.56%	68.28%	64.95%
Age Group				
12 to 17	37.68%	35.28%	36.45%	32.74%
18+	69.27%	74.98%	72.00%	68.71%
Gender				
Male	63.43%	76.85%	69.78%	70.07%
Female	68.80%	64.89%	66.88%	60.15%
Race/Ethnicity				
Hispanic	60.90%	66.28%	63.65%	57.10%
Non-Hisp., Black	54.92%	44.87%	50.31%	53.54%
Non-Hisp., All Other Races	68.67%	75.26%	71.86%	67.64%
Education[1]				
< High School	35.59%	68.41%	54.60%	48.54%
High School	70.92%	70.67%	70.80%	68.12%
> High School	75.98%	80.37%	77.99%	75.47%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.8 Past 12-Month Use of Marijuana, by Opportunities to Report Use

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	9.12%	11.45%	10.24%	9.36%
Age Group				
12 to 17	17.03%	16.36%	16.69%	12.97%
18+	8.26%	10.83%	9.49%	8.94%
Gender				
Male	11.33%	16.77%	13.91%	13.08%
Female	6.98%	6.65%	6.82%	5.88%
Race/Ethnicity				
Hispanic	15.05%	12.98%	13.99%	10.21%
Non-Hisp., Black	8.94%	7.56%	8.31%	9.60%
Non-Hisp., All Other Races	8.67%	11.96%	10.26%	9.25%
Education[1]				
< High School	8.69%	7.58%	8.05%	14.87%
High School	6.95%	10.50%	8.59%	10.10%
> High School	9.03%	12.28%	10.52%	6.30%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.9 Past 12-Month Use of Cocaine, by Opportunities to Report Use

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	3.11%	2.12%	2.63%	1.60%
Age Group				
12 to 17	1.75%	1.79%	1.77%	2.54%
18+	3.26%	2.16%	2.74%	1.49%
Gender				
Male	3.17%	2.95%	3.07%	2.10%
Female	3.06%	1.37%	2.23%	1.13%
Race/Ethnicity				
Hispanic	4.18%	5.04%	4.62%	1.56%
Non-Hisp., Black	4.86%	1.67%	3.40%	2.98%
Non-Hisp., All Other Races	2.70%	1.93%	2.33%	1.36%
Education[1]				
< High School	8.90%	2.48%	5.19%	1.87%
High School	1.51%	2.27%	1.86%	2.03%
> High School	3.12%	1.97%	2.59%	1.01%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.10 Past 12-Month Use of Any Illicit Drug, by Opportunities to Report Use

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	14.54%	15.82%	15.16%	10.28%
Age Group				
12 to 17	24.26%	21.87%	23.03%	14.54%
18+	13.48%	15.06%	14.24%	9.78%
Gender				
Male	13.08%	19.77%	16.24%	13.87%
Female	15.96%	12.26%	14.14%	6.93%
Race/Ethnicity				
Hispanic	20.34%	20.06%	20.20%	12.33%
Non-Hisp., Black	14.42%	10.76%	12.74%	9.99%
Non-Hisp., All Other Races	14.09%	16.28%	15.15%	10.16%
Education[1]				
< High School	13.76%	9.91%	11.53%	14.57%
High School	14.73%	12.46%	13.68%	10.63%
> High School	12.59%	18.75%	15.41%	7.71%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.11 Past 30-Day Use of Cigarettes, by Opportunities to Report Use				
Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	21.75%	30.98%	26.20%	30.45%
Age Group				
12 to 17	17.31%	14.98%	16.11%	18.56%
18+	22.23%	32.99%	27.38%	31.84%
Gender				
Male	23.60%	33.82%	28.43%	31.10%
Female	19.95%	28.43%	24.11%	29.84%
Race/Ethnicity				
Hispanic	29.29%	21.55%	25.33%	28.86%
Non-Hisp., Black	36.76%	10.63%	24.79%	33.03%
Non-Hisp., All Other Races	18.34%	35.26%	26.53%	30.13%
Education[1]				
< High School	33.38%	51.74%	44.01%	34.23%
High School	22.60%	43.42%	32.25%	38.72%
> High School	19.41%	18.89%	19.17%	26.50%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.12 Past 30-Day Use of Alcohol, by Opportunities to Report Use

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	50.23%	47.04%	48.70%	52.21%
Age Group				
12 to 17	18.07%	16.10%	17.06%	18.84%
18+	53.75%	50.92%	52.40%	56.11%
Gender				
Male	45.82%	56.81%	51.02%	60.23%
Female	54.51%	38.24%	46.53%	44.70%
Race/Ethnicity				
Hispanic	47.53%	28.58%	37.84%	40.79%
Non-Hisp., Black	40.26%	29.24%	35.21%	41.19%
Non-Hisp., All Other Races	52.31%	51.72%	52.02%	55.15%
Education[1]				
< High School	21.18%	37.88%	30.85%	40.55%
High School	49.16%	45.90%	47.65%	55.62%
> High School	64.33%	59.22%	61.99%	61.34%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.13 Past 30-Day Use of Marijuana, by Opportunities to Report Use				
Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present	Total ACASI	1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8		
Total	5.35%	4.09%	4.74%	5.28%
Age Group				
12 to 17	9.77%	7.03%	8.37%	7.30%
18+	4.86%	3.72%	4.32%	5.04%
Gender				
Male	7.42%	6.55%	7.01%	7.30%
Female	3.34%	1.88%	2.62%	3.39%
Race/Ethnicity				
Hispanic	2.40%	5.11%	3.79%	4.48%
Non-Hisp., Black	5.05%	3.93%	4.54%	6.05%
Non-Hisp., All Other Races	5.65%	4.03%	4.86%	5.21%
Education[1]				
< High School	7.07%	3.58%	5.05%	10.13%
High School	3.83%	4.64%	4.21%	5.92%
> High School	5.04%	3.16%	4.18%	2.86%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.14 Past 30-Day Use of Cocaine, by Opportunities to Report Use

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	0.56%	0.57%	0.57%	0.59%
Age Group				
12 to 17	0.21%	1.08%	0.66%	1.09%
18+	0.60%	0.51%	0.56%	0.53%
Gender				
Male	0.30%	0.58%	0.43%	0.61%
Female	0.81%	0.57%	0.69%	0.56%
Race/Ethnicity				
Hispanic	0.68%	2.09%	1.40%	0.80%
Non-Hisp., Black	0.96%	1.12%	1.03%	2.09%
Non-Hisp., All Other Races	0.48%	0.34%	0.41%	0.30%
Education[1]				
< High School	2.35%	1.14%	1.65%	0.87%
High School	0.46%	0.91%	0.67%	0.89%
> High School	0.29%	0.00%	0.16%	0.18%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.15 Past 30-Day Use of Any Illicit Drug, by Opportunities to Report Use

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Multiple Use Questions Absent	Multiple Use Questions Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 3, 5, 7	2, 4, 6, 8	Total ACASI	
Total	6.65%	5.13%	5.91%	5.44%
Age Group				
12 to 17	13.10%	10.85%	11.95%	7.98%
18+	5.94%	4.41%	5.21%	5.14%
Gender				
Male	8.66%	7.58%	8.15%	7.68%
Female	4.70%	2.91%	3.82%	3.34%
Race/Ethnicity				
Hispanic	3.64%	6.34%	5.02%	4.60%
Non-Hisp., Black	7.15%	6.30%	6.76%	6.45%
Non-Hisp., All Other Races	6.80%	4.82%	5.84%	5.33%
Education[1]				
< High School	7.59%	5.33%	6.28%	9.46%
High School	5.16%	5.01%	5.09%	5.94%
> High School	6.08%	3.66%	4.97%	3.25%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.2.16 Comparison of 30-Day User Coding Schemes

Drugs of Interest	Single Gate Questions (Treatment Versions: 1, 2, 3, 4)				Multiple Gate Questions (Treatment Versions: 5, 6, 7, 8)				PAPI Comparison
	Raw Variable	Most Liberal	Approximate NHSDA	Edited Variable	Raw Variable	Most Liberal	Approximate NHSDA	Edited Variable	
Alcohol									
Total	50.70%	49.87%	47.64%	47.35%	55.16%	55.78%	51.78%	50.27%	46.45%
12 to 17	20.78%	21.49%	18.79%	17.43%	21.04%	21.36%	18.70%	16.60%	12.04%
18+	54.28%	53.26%	51.09%	50.93%	59.05%	59.70%	55.55%	54.11%	50.47%
Cigarettes									
Total	26.19%	26.67%	26.61%	26.61%	27.22%	26.63%	25.73%	25.73%	27.47%
12 to 17	16.41%	16.92%	16.65%	16.62%	20.30%	16.97%	15.49%	15.49%	12.83%
18+	27.36%	27.84%	27.80%	27.80%	28.01%	27.73%	26.89%	26.89%	29.19%
Marijuana									
Total	6.18%	6.47%	6.45%	5.54%	4.68%	4.67%	3.92%	3.82%	4.24%
12 to 17	9.44%	9.85%	9.73%	9.39%	8.63%	8.76%	7.63%	7.12%	5.26%
18+	5.79%	6.06%	6.06%	5.08%	4.23%	4.20%	3.50%	3.44%	4.12%
Cocaine									
Total	0.61%	0.61%	0.60%	0.60%	0.84%	0.58%	0.53%	0.53%	0.30%
12 to 17	1.02%	1.02%	0.98%	0.98%	0.64%	0.49%	0.26%	0.26%	0.71%
18+	0.56%	0.56%	0.56%	0.56%	0.86%	0.59%	0.56%	0.56%	0.26%
Any Illicit Drug[1]									
Total		7.82%	7.79%	6.91%		5.62%	4.89%	4.75%	5.44%
12 to 17		14.30%	13.95%	13.02%		12.15%	11.52%	10.63%	7.98%
18+		7.05%	7.05%	6.18%		4.87%	4.30%	4.08%	5.14%
Alcohol									
Total	52.58%	52.58%	50.23%	50.23%	52.96%	52.61%	48.82%	47.04%	46.45%
12 to 17	21.50%	21.50%	18.07%	18.07%	20.32%	21.38%	19.40%	16.10%	12.04%
18+	55.97%	55.97%	53.75%	53.75%	57.05%	56.53%	52.50%	50.92%	50.47%
Cigarettes									
Total	22.52%	22.32%	21.75%	21.75%	31.13%	31.31%	30.99%	30.98%	27.47%
12 to 17	19.54%	17.82%	17.31%	17.31%	16.84%	16.11%	15.01%	14.98%	12.83%
18+	22.84%	22.81%	22.23%	22.23%	32.92%	33.21%	32.99%	32.99%	29.19%
Marijuana									
Total	5.97%	5.97%	5.35%	5.35%	4.97%	5.28%	5.22%	4.09%	4.24%
12 to 17	10.34%	10.34%	9.77%	9.77%	7.87%	8.42%	7.84%	7.03%	5.26%

18+	5.49%	5.49%	4.86%	4.86%	4.61%	4.89%	4.89%	3.72%	4.12%
Cocaine									
Total	0.56%	0.56%	0.56%	0.56%	0.87%	0.63%	0.57%	0.57%	0.30%
12 to 17	0.21%	0.21%	0.21%	0.21%	1.46%	1.32%	1.08%	1.08%	0.71%
18+	0.60%	0.60%	0.60%	0.60%	0.80%	0.54%	0.51%	0.51%	0.26%
Any Illicit Drug [1]									
Total		7.31%	6.65%	6.65%		6.27%	6.24%	5.13%	5.44%
12 to 17		14.04%	13.10%	13.10%		12.65%	12.62%	10.85%	7.98%
18+		6.57%	5.94%	5.94%		5.47%	5.44%	4.41%	5.14%
Alcohol									
Total	51.09%	49.91%	45.37%	44.80%	54.59%	55.55%	54.15%	52.98%	46.45%
12 to 17	20.30%	20.91%	17.92%	16.16%	21.59%	22.05%	19.72%	18.12%	12.04%
18+	54.82%	53.42%	48.69%	48.27%	58.30%	59.32%	58.02%	56.91%	50.47%
Cigarettes									
Total	23.13%	22.87%	22.87%	22.87%	30.56%	30.80%	29.86%	29.86%	27.47%
12 to 17	16.52%	15.50%	15.50%	15.50%	20.08%	18.63%	16.87%	16.83%	12.83%
18+	23.93%	23.76%	23.76%	23.76%	31.74%	32.17%	31.33%	31.33%	29.19%
Marijuana									
Total	5.10%	4.56%	3.95%	3.74%	5.91%	6.82%	6.75%	5.84%	4.24%
12 to 17	8.46%	8.70%	8.22%	7.73%	9.80%	10.12%	9.44%	9.11%	5.26%
18+	4.70%	4.06%	3.43%	3.26%	5.48%	6.45%	6.45%	5.48%	4.12%
Cocaine									
Total	0.70%	0.62%	0.59%	0.59%	0.72%	0.57%	0.54%	0.54%	0.30%
12 to 17	1.03%	1.03%	1.03%	1.03%	0.64%	0.49%	0.22%	0.22%	0.71%
18+	0.66%	0.57%	0.54%	0.54%	0.73%	0.57%	0.57%	0.57%	0.26%
Any Illicit Drug [1]									
Total		5.47%	4.83%	4.66%		8.27%	8.23%	7.30%	5.44%
12 to 17		12.18%	11.67%	10.44%		14.67%	14.26%	13.71%	7.98%
18+		4.66%	4.00%	3.96%		7.55%	7.55%	6.58%	5.14%

Note: Raw variable estimates are unavailable for coding schemes for any illicit drug use.

[1] Any illicit drug includes marijuana, cocaine, crack, heroin, inhalants, hallucinogens, and nonmedical use of analgesics, sedatives, stimulants, and tranquilizers.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures, 1997 Field Experiment.

Exhibit 7.2.17 Number of Respondents Who Provided Inconsistent Reports of the Recency of Use as a Result of the Multiple Use Treatment

	Multiple Use Present				
	Single Gate Question		Multiple Gate Questions		
	Consistency Checks		Consistency Checks		
	Absent	Present	Absent	Present	
	2	4	6	8	Total
Number of Respondents	285	264	219	207	975
Total Number of Inconsistencies	48	24	41	28	141
Suggests More Recent Use	38	16	22	17	93
Percentage of Inconsistencies	79.2	66.7	53.7	60.7	66.0
Suggests Less Recent Use	10	8	19	11	48
Percentage of Inconsistencies	20.8	33.3	46.3	39.3	34.0

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.2.18 Average Time to Complete Treatment Sections, by Whether Multiple Use Was Present

	Average Time to Complete			
	Multiple Use Present		Multiple Use Absent	
Questionnaire Section	# of Cases	Time (minutes)	# of Cases	Time (minutes)
Tobacco	1,022	1.6	955	1.6
Alcohol	1,020	3.3	956	3.3
Marijuana	1,020	0.9	956	0.9
Cocaine	1,020	0.5	956	0.5
Crack	1,020	0.4	80	0.6
Heroin	1,020	0.3	956	0.3
Hallucinogens	1,019	1.5	956	1.4
Inhalants	1,019	2.5	953	2.4
Total for Treatment Sections	1,022	10.9	956	10.5

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.2.19 Breakoff Rates, by Multiple Use Status				
	Breakoff Rate			
	Multiple Use Present		Multiple Use Absent	
Cases Included	*n*	%	*n*	%
Cases Ever Recorded as Breakoff Even if Completed Later	85	8.3	78	8.2
Cases Finalized as Breakoff	17	1.7	15	1.6
Total Number of Respondents Assigned to Multiple Use Treatment	1,025		957	

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.2.20 Respondent Reactions to CAI Interview, by Multiple Use Status				
	Multiple Use Status			
	Multiple Use Present		Multiple Use Absent	
Item	*n*	%	*n*	%
How Difficult to Use Computer?				
Very difficult	23	2.3	19	2.0
Somewhat difficult	63	6.2	61	6.5
Not at all difficult	923	91.5	863	91.5
Total	1,009	100.0	943	100.0
Able to Enter Answers Easily?				
Yes	891	88.2	833	88.3
No	119	11.8	110	11.7
Total	1,010	100.0	943	100.0
Wanted to Change Previous Answer But Did Not?				
Yes	183	18.2	174	18.5
No	821	81.8	768	81.5
Total	1,004	100.0	942	100.0
Level of Interest in Interview				
Very boring	62	6.2	66	7.0
Somewhat boring	120	12.0	107	11.4
Neither boring nor interesting	259	25.8	248	26.4
Somewhat interesting	350	34.9	319	34.0
Very interesting	211	21.1	199	21.2
Total	1,003	100.0	939	100.0
Accuracy and Completeness of Answers				
Very accurate and complete	831	82.6	789	83.9
Fairly accurate and complete	163	16.2	141	15.0
Not very accurate and complete	12	1.2	10	1.1
Total	1,006	100.0	940	100.0

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

7.3 Implementation of Consistency Checks by Respondents

The consistency checks programmed for the consistency check treatment can be divided into two types. First, there are true inconsistencies, in which a respondent's answers to each of two questions cannot both be true. For example, a respondent who has reported her current age as 22 cannot logically report her first use of marijuana at age 35. Similarly, a respondent who reports drinking alcohol on 22 days during the past 30 days cannot logically report drinking alcohol on 3 days in the past 12 months. A second type of check that was programmed can be considered a verify check as opposed to a true inconsistency. In these cases, a respondent's answer is not technically inconsistent with previous information reported but rather is simply "unusual" given what we have learned from the data collected in previous NHSDAs. The two verify checks included in the 1997 field experiment prompted respondents to verify the accuracy of their response when they indicated an age at first use as equal to their current age or an age at first use less than 10. When discussing the outcome of the consistency check treatment, we first must consider the two types of consistency checks separately because they may reflect different effects on data quality.

7.3.1 Inconsistencies in the Final Unedited Respondent Data, by Presence or Absence of Consistency Checks

Exhibit 7.3.1 shows the number of respondents in each of the four versions in which consistency checks were present who triggered at least one consistency check during the course of the interview. From this exhibit, we see that approximately 28% of the respondents assigned to receive consistency checks triggered at least one check item. Of those respondents, just over half (55%) triggered a verify check as opposed to a true inconsistency item. More respondents triggered a true consistency check in Treatments 4 and 8. This may be the result of additional confusion created by the multiple use treatment. However, it may only reflect the fact that there were more consistency checks programmed into these two treatments. Interestingly, the number of respondents who triggered verify items was higher for the two treatments that used the single gate question.

Exhibit 7.3.2 shows a breakdown of the age of respondents who triggered at least one true inconsistency. Across all treatments, youths aged 12 to 17 were more likely to trigger a true inconsistency than adults aged 18 or older. In Treatment 8, the difference was especially pronounced: 23% of youths triggered an inconsistency and only 11.7% of adults did.

In addition to determining the number of respondents who triggered any of the inconsistency items, we also wanted to calculate the total number of consistency checks triggered overall and by treatment. Exhibit 7.3.3 shows these data. Combining these data with data from Exhibit 7.3.1 shows that the average number of true consistency checks triggered was approximately 1.2 per respondent. The average number of verify checks triggered was approximately 1.2 per respondent as well.

Of particular interest is whether the data coming out of the consistency checks were "cleaner" than the data going in. For the 1997 field experiment, we decided to route respondents through each consistency check only once. Thus, if a respondent reported her current age as 26 and her age at first use of cocaine as age 33, she was routed into a consistency check. If she chose to change her age at first use response to 30, her answers would still be inconsistent. However, the program was not designed to route respondents back through a consistency check a second time. We were concerned that multiple passes through a single consistency check might be overly burdensome to respondents.

Exhibit 7.3.4 provides data on the consistency of responses following a consistency check. Each resolution process was examined and classified into one of four types. First, the data could be classified as consistent; that is, after being notified of the inconsistency, the respondent revised one or both of his or her answers so that the answers did not conflict with each other. Second, the data could remain inconsistent; even after changing one or both of the answers, the respondent's answers still conflicted with one another. Some responses were classified as "indeterminate" because the nature of the resolution process did not require the respondent to fully resolve the inconsistency. For example, in Treatment 8, a respondent may have reported that he used marijuana during the past 30 days; when asked the 30-day frequency item, however, he reported 0 days. He was then routed to a resolution screen that asked him to verify that he did not use marijuana in the past 30 days. If he indicated this was correct, that was the end of the resolution process, even though he was never presented with the two conflicting items. Although it seems safe to consider this respondent a nonuser for the 30-day period, for the sake of clarity we have classified these cases separately in Exhibit 7.3.4. Finally, in a small number of situations, the respondent recorded a "don't know" or "refused" response when presented with the resolution screen. In these situations, the program routed the respondent out of the resolution series; thus, the case cannot be classified into any of the other three categories.

From Exhibit 7.3.4, we see that 30% of all inconsistencies were resolved as consistent, 17% were still inconsistent after the resolution process, and just over half (51%) the respondent confirmed an answer. If we assume that the confirmed answer is accurate, then the quality of 81% of the inconsistent responses was improved through the resolution process. Also from Exhibit 7.3.4, we find that slightly less than half of the inconsistencies resolved in Treatments 3 and 4 were determined to be consistent. In contrast, less than 20% of the inconsistencies resolved in Treatments 7 and 8 were determined to be consistent. The reason for this difference is not clear because the resolution questions do not vary significantly between the two contingent questioning structures.

We hypothesized that respondents who were overly confused by the resolution process would enter a "don't know" or "refused" response as a method of avoiding the resolution task. We found only a small number of inconsistencies were "resolved" bythe respondent in this manner (a total of four resolutions across all treatment versions). We take this as a positive sign that the resolution process was not overly confusing for respondents and believe that, overall, the improvement in data consistency is an improvement to the overall quality of the NHSDA data as well.

We also examined the resolution of the verify questions on age at first use in more detail. The results of this analysis are shown in Exhibit 7.3.5, which displays the number of respondents for the various drugs who reported very low ages at first use and the percentage who changed their answers as a result of the consistency check. Overall, 98 respondents gave ages at first use younger than 10 years old, and about 23% of these respondents indicated that this was an incorrect answer. Most of these respondents gave a new age that was consistent; however, a few did not complete the task, and we were left with the knowledge only that they had indicated that the very low age at first use was incorrect.

Exhibit 7.3.5 also shows data on those whose reported age at first use was equal to their current age. The vast majority of these respondents confirmed that this was indeed their age at first use; however, about 6% gave a revised age at first use.

7.3.2 Effect of Revised Answers on Prevalence Estimates

Based on the data presented in Section 7.3.1, it appears that the consistency check process is helpful in improving the consistency of the NHSDA data. In addition, the process seems relatively burden-free because less than a third of the respondents actually triggered a resolution screen and the average number of resolution screens triggered was quite low (1.2 for both the verify and the true inconsistency screens). Of equal importance, however, is the effect of the inconsistency resolution process on the prevalence estimates.

Exhibits 7.3.6 through 7.3.20 compare prevalence estimates between the two levels of the consistency check treatment. In these exhibits, comparable data are included for the 1997 Quarter 4 PAPI data collection. The results show a fairly consistent increase in the drug prevalence rates for both lifetime and 12-month use. The 30-day rates are less consistent, but due to the small sample sizes for many of the drugs, these rates are also likely to be less reliable.

The resolution process appears to have had a slight decreasing effect on the reports of youths aged 12 to 17, while the prevalence rates for adults aged 18 or older were increased by the presence of the resolution treatment. A large increase in the prevalence rates was also found for adults with less than a high school education when the consistency check treatment was included.

7.3.3 Operational Aspects of Reporting When Consistency Checks Are Used

Time required to complete the interview. Exhibit 7.3.21 provides descriptive dataon the average length of time required to complete the treatment section of the instrument when consistency checks were and were not in use. Data are included for each individual drug section and for the treatment section as a whole. From this exhibit, we see that the average time to complete each drug section was nearly identical between the two resolution conditions. The overall average times were identical as well. However, the majority of respondents did not trigger many, if any, consistency checks during their interviews. To present a more realistic picture of the time added by the inclusion of resolution items, we compared the time to complete the treatment section by whether even one inconsistency was resolved during the respondent's interview. The results are presented in Exhibit 7.3.22. Here we see an increase of approximately 3 minutes when at least one consistency check was triggered compared to when none was triggered. One additional method for considering the time added when the consistency checks were present is to compare the time to complete a specific section of the instrument by whether an inconsistency was detected in that section. These data are presented in Exhibit 7.3.23. In some cases, sample sizes are quite small. However, we see that for many of the drug sections, triggering a consistency check could result in a doubling of the time to complete the section compared to when no consistency checks were triggered. More detailed analysis of the timing data from the 1997 field experiment are presented in Section 7.5.

Breakoff rates. We next compared the rate of breakoffs by analyzing whether the multiple use treatment was in use. Breakoff cases accounted for only a small number of the total number of interviews. However, if the consistency checks had caused respondents not to complete the full interview, we would view this as a serious drawback of the methodology. We analyzed the breakoff data in two ways. First, we compared the percentage of cases ever reported as a breakoff (even if the interviewer went on to finalize the case as a completed interview) by the presence of consistency checks. Then, we compared the percentage of finalized breakoff cases. Both sets of data are presented in Exhibit 7.3.24. The data show fundamentally that the presence of consistency checks had no effect on the breakoff rate.

Respondent reactions as reported in the debriefing questionnaire. We analyzed items from the respondent debriefing questions to see whether the consistency checks had any effect on respondent reactions to the CAI NHSDA interview. Exhibit 7.3.25 includes data from five items included in the respondent debriefing that we thought might vary by treatment version. Specifically, these items asked respondents to rate their difficulty in using the computer, their interest in the interview, and the accuracy and completeness of their answers. In addition, respondents were asked whether they were able to enter their answers easily and whether they wanted to change an answer to a previous question but did not. The data in Exhibit 6.3.51 show that respondents in the consistency check present group were slightly more likely than respondents who did not resolve inconsistencies to report that they were able to easily enter their answers into the computer (90.0% vs. 86.6%) and that the answers they gave were very accurate and complete (84.4% vs. 82.1%). For the other items included in Exhibit 7.3.25, no differences were found between the twogroups.

Interviewer reports of respondent difficulties. Interviewers completed debriefing questions at the conclusion of each interview. For respondents who triggered at least one inconsistency item (a true inconsistency rather than only a verify item), one of these debriefing items asked the interviewer to explain why the respondent might have provided inconsistent data. A review of the open-ended responses indicates that the most common reason provided by the interviewers was the number of distractions during

the interview. These distractions included the presence of young children, the presence of other adults, the telephone ringing, the doorbell ringing, and the television playing. Other explanations mentioned by the interviewers included respondent boredom, literacy problems, respondent fatigue, and respondents who simply were not paying close attention to what they were doing.

Exhibit 7.3.26 includes data from several other interviewer debriefing items that we thought might differ based on whether the respondent was assigned to receive consistency checks. For the most part, interviewer reports were similar across the two modes. However, interviewers reported that slightly more respondents in the consistency check present group had no difficulty understanding the interview (84.4% vs. 81.4%) and were very interested in the ACASI interview (62.7% vs. 59.9%). Interviewers also reported that for slightly more of the respondents who did not receive the consistency checks, the audio component of ACASI was unnecessary (36.8% vs. 34.5%).

Detailed Exhibits for Section 7.3

Exhibit 7.3.1 Number and Percentage of Respondents Who Triggered at Least One Consistency Check

	Consistency Checks Present				
	Single Gate Question		Multiple Gate Questions		
	Multiple Use Questions		Multiple Use Questions		
	Absent	Present	Absent	Present	
Treatment Version	*3*	*4*	*7*	*8*	Total
Number of Respondents	285	264	219	207	975
Total Number of Respondents Who Triggered a Consistency Check	75	82	52	68	277
Percentage of Respondents	26.3	31.1	23.7	32.9	28.4
True Inconsistency	27	39	21	37	124
Percentage of Respondents	9.5	14.8	9.6	17.9	12.7
Percentage of Consistency Checks	36.0	47.6	40.3	54.4	44.8
Verify Only	48	43	31	31	153
Percentage of Respondents	16.8	16.3	14.2	15.0	15.7
Percentage of Consistency Checks	64.0	52.4	59.6	45.6	55.2

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.3.2 Number of Respondents Who Triggered at Least One True Inconsistency, by Age and Treatment					
	Consistency Checks Present				
	Single Gate Question		Multiple Gate Questions		
	Multiple Use Questions		Multiple Use Questions		
	Absent	Present	Absent	Present	
Treatment Version	*3*	*4*	*7*	*8*	Total
Number of Respondents	285	264	219	207	975
Total Number of True Inconsistencies Triggered	27	39	21	37	124
Percentage of Respondents	9.5	14.8	9.6	17.9	12.7
12 to 17					
Number of Respondents	157	148	118	113	536
Number Triggering True Inconsistencies	16	22	11	26	75
Percentage of Respondents	10.2	14.9	9.3	23.0	14.0
18+					
Number of Respondents	128	116	101	94	439
Number Triggering True Inconsistencies	11	17	10	11	49
Percentage of Respondents	8.6	14.7	9.9	11.7	11.2

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.3.3 Number of Consistency Checks Triggered, by Treatment					
	Consistency Checks Present				
	Single Gate Question		Multiple Gate Questions		
	Multiple Use Questions		Multiple Use Questions		
	Absent	Present	Absent	Present	
Treatment Version	*3*	*4*	*7*	*8*	Total
Number of Respondents	285	264	219	207	975
Total Number of Consistency Checks Triggered	95	101	60	83	339
True Inconsistency	31	50	24	45	150
Verify Only	64	51	36	38	189

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.3.4 Data Consistency Following the Resolution Process					
	Consistency Checks Present				
	Single Gate Question		Multiple Gate Questions		
	Multiple Use Questions		Multiple Use Questions		
	Absent	Present	Absent	Present	
Treatment Version	*3*	*4*	*7*	*8*	Total
Number of Respondents	285	264	219	207	975
Total Number of True Inconsistencies Triggered	31	50	24	45	150
Made Change: Data Consistent	15	21	4	5	45
Percentage of inconsistencies	48.4	42.0	16.7	11.1	30.0
Made Change: Data Inconsistent	5	6	7	7	25
Percentage of inconsistencies	16.1	12.0	29.1	15.6	16.7
Confirmed Last Answer	8	23	13	32	76
Percentage of inconsistencies	25.8	46.0	54.2	71.1	50.7
Don't Know/Refused	3	0	0	1	4
Percentage of inconsistencies	9.7	-	-	2.2	2.7

Source: National Household Survey on Drug Abuse: Development of Computer Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.3.5 Summary of Consistency Resolution for Checks on Age at First Use				
Drug	Very Low Age at First Use Reported	Very Low Age at First Use Changed	Reported Age at First Use Equals Current Age	Current Age Not Verified as Age of First Use
Cigarettes				
Number	30	7	17	3
Percentage		23.33		17.65
Alcohol				
Number	50	7	29	0
Percentage		14.00		0.00
Marijuana				
Number	4	1	18	0
Percentage		25.00		0.00
Cocaine				
Number	1	1	5	0
Percentage		100.00		0.00
Crack				
Number	1	0	2	0
Percentage		0.00		0.00
Heroin				
Number	2	1	2	0
Percentage		50.00		0.00
Hallucinogens				
Number	3	2	11	1
Percentage		66.67		9.09
Inhalants				
Number	7	4	4	1
Percentage		57.14		25.00
Overall Drugs				
Number	98	23	88	5
Percentage		23.47		5.68

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.3.6 Lifetime Use of Cigarettes, by Presence of Consistency Checks

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	69.70%	74.45%	71.96%	70.87%
Age Group				
12 to 17	45.49%	43.29%	44.48%	36.05%
18+	72.63%	77.96%	75.18%	74.94%
Gender				
Male	75.47%	72.81%	74.25%	74.74%
Female	63.97%	75.88%	69.83%	67.25%
Race/Ethnicity				
Hispanic	61.10%	66.72%	63.68%	55.82%
Non-Hisp., Black	63.99%	64.56%	64.30%	61.62%
Non-Hisp., All Other Races	71.32%	77.15%	74.04%	73.81%
Education[1]				
< High School	69.58%	87.21%	77.83%	70.28%
High School	74.28%	86.76%	79.73%	76.65%
> High School	72.36%	70.37%	71.35%	75.27%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.7 Lifetime Use of Alcohol, by Presence of Consistency Checks

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	82.12%	84.30%	83.16%	82.15%
Age Group				
12 to 17	45.55%	44.14%	44.90%	38.01%
18+	86.54%	88.82%	87.63%	87.31%
Gender				
Male	80.29%	88.26%	83.95%	87.07%
Female	83.94%	80.83%	82.41%	77.55%
Race/Ethnicity				
Hispanic	76.29%	71.87%	74.26%	69.45%
Non-Hisp., Black	61.36%	69.05%	65.54%	72.35%
Non-Hisp., All Other Races	85.77%	88.50%	87.04%	84.98%
Education[1]				
< High School	82.62%	81.94%	82.30%	77.22%
High School	82.80%	87.21%	84.73%	85.83%
> High School	90.68%	91.63%	91.16%	91.48%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.8 Lifetime Use of Marijuana, by Presence of Consistency Checks

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	35.04%	35.34%	35.18%	34.96%
Age Group				
12 to 17	20.73%	20.36%	20.56%	16.10%
18+	36.78%	37.02%	36.90%	37.17%
Gender				
Male	37.89%	37.66%	37.78%	40.79%
Female	32.22%	33.30%	32.75%	29.51%
Race/Ethnicity				
Hispanic	49.07%	27.31%	39.06%	24.59%
Non-Hisp., Black	26.19%	34.95%	30.95%	31.43%
Non-Hisp., All Other Races	35.15%	36.10%	35.60%	36.48%
Education[1]				
< High School	30.04%	25.36%	27.85%	32.84%
High School	35.73%	36.49%	36.06%	36.60%
> High School	39.76%	40.52%	40.15%	38.91%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.9 Lifetime Use of Cocaine, by Presence of Consistency Checks

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	11.61%	18.54%	14.91%	10.03%
Age Group				
12 to 17	2.46%	3.32%	2.86%	3.13%
18+	12.72%	20.25%	16.32%	10.84%
Gender				
Male	13.36%	24.76%	18.60%	12.47%
Female	9.88%	13.10%	11.46%	7.76%
Race/Ethnicity				
Hispanic	25.35%	7.86%	17.30%	7.32%
Non-Hisp., Black	5.46%	13.65%	9.90%	7.57%
Non-Hisp., All Other Races	11.34%	20.46%	15.59%	10.70%
Education[1]				
< High School	16.48%	33.64%	24.51%	9.52%
High School	12.61%	16.41%	14.27%	10.37%
> High School	11.59%	18.79%	15.26%	11.58%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.10 Lifetime Use of Any Illicit Drug, by Presence of Consistency Checks

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	43.43%	44.03%	43.72%	37.50%
Age Group				
12 to 17	30.82%	31.55%	31.15%	20.45%
18+	44.96%	45.44%	45.19%	39.49%
Gender				
Male	40.18%	47.88%	43.72%	43.44%
Female	46.66%	40.67%	43.72%	31.94%
Race/Ethnicity				
Hispanic	51.88%	35.29%	44.25%	28.63%
Non-Hisp., Black	28.87%	43.09%	36.59%	33.40%
Non-Hisp., All Other Races	44.90%	44.98%	44.93%	38.99%
Education[1]				
< High School	36.28%	55.02%	45.05%	37.22%
High School	50.83%	38.32%	45.36%	38.42%
> High School	43.27%	46.91%	45.12%	40.92%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.11 Past 12-Month Use of Cigarettes, by Presence of Consistency Checks				
	ACASI Treatment Groups			Comparison Group
Respondent Characteristics	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	28.56%	34.07%	31.19%	33.52%
Age Group				
12 to 17	23.82%	21.77%	22.88%	25.18%
18+	29.13%	35.46%	32.16%	34.49%
Gender				
Male	37.26%	34.20%	35.85%	34.17%
Female	19.92%	33.96%	26.82%	32.91%
Race/Ethnicity				
Hispanic	23.36%	34.64%	28.55%	32.69%
Non-Hisp., Black	23.36%	35.48%	29.94%	35.19%
Non-Hisp., All Other Races	29.80%	33.73%	31.64%	33.29%
Education[1]				
< High School	34.32%	57.58%	45.20%	34.76%
High School	32.47%	40.65%	36.04%	41.55%
> High School	24.90%	26.42%	25.67%	29.70%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.12 Past 12-Month Use of Alcohol, by Presence of Consistency Checks				
	ACASI Treatment Groups			Comparison Group
Respondent Characteristics	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	65.09%	71.79%	68.28%	64.95%
Age Group				
12 to 17	36.99%	35.83%	36.45%	32.74%
18+	68.49%	75.83%	72.00%	68.71%
Gender				
Male	65.13%	75.23%	69.78%	70.07%
Female	65.05%	68.77%	66.88%	60.15%
Race/Ethnicity				
Hispanic	65.08%	61.97%	63.65%	57.10%
Non-Hisp., Black	45.75%	54.16%	50.31%	53.54%
Non-Hisp., All Other Races	68.02%	76.26%	71.86%	67.64%
Education[1]				
< High School	51.86%	57.71%	54.60%	48.54%
High School	67.24%	75.38%	70.80%	68.12%
> High School	74.81%	81.06%	77.99%	75.47%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.13 Past 12-Month Use of Marijuana, by Presence of Consistency Checks				
Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	9.95%	10.56%	10.24%	9.36%
Age Group				
12 to 17	16.17%	17.29%	16.69%	12.97%
18+	9.20%	9.80%	9.49%	8.94%
Gender				
Male	13.43%	14.46%	13.91%	13.08%
Female	6.50%	7.15%	6.82%	5.88%
Race/Ethnicity				
Hispanic	20.85%	5.95%	13.99%	10.21%
Non-Hisp., Black	7.41%	9.06%	8.31%	9.60%
Non-Hisp., All Other Races	9.38%	11.27%	10.26%	9.25%
Education[1]				
< High School	9.98%	5.85%	8.05%	14.87%
High School	8.81%	8.31%	8.59%	10.10%
> High School	9.25%	11.75%	10.52%	6.30%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.14 Past 12-Month Use of Cocaine, by Presence of Consistency Checks

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	2.28%	3.03%	2.63%	1.60%
Age Group				
12 to 17	1.74%	1.81%	1.77%	2.54%
18+	2.34%	3.17%	2.74%	1.49%
Gender				
Male	2.42%	3.82%	3.07%	2.10%
Female	2.13%	2.34%	2.23%	1.13%
Race/Ethnicity				
Hispanic	5.52%	3.56%	4.62%	1.56%
Non-Hisp., Black	1.48%	5.02%	3.40%	2.98%
Non-Hisp., All Other Races	2.11%	2.57%	2.33%	1.36%
Education[1]				
< High School	5.07%	5.32%	5.19%	1.87%
High School	1.72%	2.04%	1.86%	2.03%
> High School	1.94%	3.22%	2.59%	1.01%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

.

Exhibit 7.3.15 Past 12-Month Use of Any Illicit Drug, by Presence of Consistency Checks				
	ACASI Treatment Groups			Comparison Group
Respondent Characteristics	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	14.57%	15.81%	15.16%	10.28%
Age Group				
12 to 17	21.22%	25.17%	23.03%	14.54%
18+	13.76%	14.75%	14.24%	9.78%
Gender				
Male	16.05%	16.47%	16.24%	13.87%
Female	13.09%	15.23%	14.14%	6.93%
Race/Ethnicity				
Hispanic	23.93%	15.81%	20.20%	12.33%
Non-Hisp., Black	9.95%	15.10%	12.74%	9.99%
Non-Hisp., All Other Races	14.45%	15.95%	15.15%	10.16%
Education[1]				
< High School	11.77%	11.27%	11.53%	14.57%
High School	15.46%	11.39%	13.68%	10.63%
> High School	13.11%	17.64%	15.41%	7.71%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.16 Past 30-Day Use of Cigarettes, by Presence of Consistency Checks

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	22.87%	29.86%	26.20%	30.45%
Age Group				
12 to 17	15.50%	16.83%	16.11%	18.56%
18+	23.76%	31.33%	27.38%	31.84%
Gender				
Male	27.94%	29.02%	28.43%	31.10%
Female	17.84%	30.60%	24.11%	29.84%
Race/Ethnicity				
Hispanic	20.83%	30.62%	25.33%	28.86%
Non-Hisp., Black	17.91%	30.58%	24.79%	33.03%
Non-Hisp., All Other Races	23.80%	29.64%	26.53%	30.13%
Education[1]				
< High School	33.42%	56.06%	44.01%	34.23%
High School	28.48%	37.10%	32.25%	38.72%
> High School	17.03%	21.24%	19.17%	26.50%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.17 Past 30-Day Use of Alcohol, by Presence of Consistency Checks

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	44.80%	52.98%	48.70%	52.21%
Age Group				
12 to 17	16.16%	18.12%	17.06%	18.84%
18+	48.27%	56.91%	52.40%	56.11%
Gender				
Male	45.13%	57.93%	51.02%	60.23%
Female	44.47%	48.65%	46.53%	44.70%
Race/Ethnicity				
Hispanic	39.24%	36.19%	37.84%	40.79%
Non-Hisp., Black	29.91%	39.68%	35.21%	41.19%
Non-Hisp., All Other Races	47.54%	57.16%	52.02%	55.15%
Education[1]				
< High School	20.37%	42.78%	30.85%	40.55%
High School	47.46%	47.90%	47.65%	55.62%
> High School	57.89%	65.95%	61.99%	61.34%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.18 Past 30-Day Use of Marijuana, by Presence of Consistency Checks

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	3.74%	5.84%	4.74%	5.28%
Age Group				
12 to 17	7.73%	9.11%	8.37%	7.30%
18+	3.26%	5.48%	4.32%	5.04%
Gender				
Male	4.89%	9.49%	7.01%	7.30%
Female	2.60%	2.65%	2.62%	3.39%
Race/Ethnicity				
Hispanic	3.82%	3.75%	3.79%	4.48%
Non-Hisp., Black	3.61%	5.32%	4.54%	6.05%
Non-Hisp., All Other Races	3.75%	6.13%	4.86%	5.21%
Education[1]				
< High School	6.06%	3.89%	5.05%	10.13%
High School	4.68%	3.59%	4.21%	5.92%
> High School	1.26%	6.99%	4.18%	2.86%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.19 Past 30-Day Use of Cocaine, by Presence of Consistency Checks

Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present		1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8	Total ACASI	
Total	0.59%	0.54%	0.57%	0.59%
Age Group				
12 to 17	1.03%	0.22%	0.66%	1.09%
18+	0.54%	0.57%	0.56%	0.53%
Gender				
Male	0.32%	0.57%	0.43%	0.61%
Female	0.87%	0.51%	0.69%	0.56%
Race/Ethnicity				
Hispanic	1.30%	1.53%	1.40%	0.80%
Non-Hisp., Black	0.70%	1.30%	1.03%	2.09%
Non-Hisp., All Other Races	0.52%	0.30%	0.41%	0.30%
Education[1]				
< High School	2.09%	1.15%	1.65%	0.87%
High School	0.59%	0.76%	0.67%	0.89%
> High School	0.00%	0.31%	0.16%	0.18%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.20 Past 30-Day Use of Any Illicit Drug, by Presence of Consistency Checks				
Respondent Characteristics	ACASI Treatment Groups			Comparison Group
	Consistency Checks Absent	Consistency Checks Present	Total ACASI	1997 Quarter 4 PAPI/SAQ
Treatment Version:	1, 2, 5, 6	3, 4, 7, 8		
Total	4.66%	7.30%	5.91%	5.44%
Age Group				
12 to 17	10.44%	13.71%	11.95%	7.98%
18+	3.96%	6.58%	5.21%	5.14%
Gender				
Male	5.89%	10.80%	8.15%	7.68%
Female	3.43%	4.23%	3.82%	3.34%
Race/Ethnicity				
Hispanic	4.87%	5.19%	5.02%	4.60%
Non-Hisp., Black	5.08%	8.17%	6.76%	6.45%
Non-Hisp., All Other Races	4.57%	7.30%	5.84%	5.33%
Education[1]				
< High School	7.29%	5.13%	6.28%	9.46%
High School	4.84%	5.42%	5.09%	5.94%
> High School	2.20%	7.64%	4.97%	3.25%

[1]Education includes only individuals aged 18 or older.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 7.3.21 Average Time to Complete Treatment Sections, by Whether Consistency Checks Are Used				
	Average Time to Complete			
	Consistency Check Present		Consistency Check Absent	
Questionnaire Section	*n*	Time (minutes)	*n*	Time (minutes)
Tobacco	972	1.6	1,005	1.6
Alcohol	972	3.4	1,004	3.2
Marijuana	972	0.9	1,004	0.9
Cocaine	972	0.5	1,004	0.5
Crack	515	0.4	585	0.4
Heroin	972	0.3	1,004	0.3
Hallucinogens	971	1.4	1,004	1.5
Inhalants	969	2.4	1,003	2.4
Total for Treatment Section	973	10.8	1,005	10.6

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.3.22 Average Time to Complete Treatment Sections, by Whether Any Consistency Checks Were Triggered Anywhere in Instrument				
	Average Time to Complete			
	Consistency Check Triggered		No Consistency Check Triggered	
Questionnaire Section	*n*	Time (minutes)	*n*	Time (minutes)
Tobacco	227	2.2	745	1.5
Alcohol	228	4.4	744	3.1
Marijuana	228	1.2	744	0.8
Cocaine	228	0.6	744	0.4
Crack	145	0.4	370	0.4
Heroin	228	0.3	744	0.3
Hallucinogens	228	1.5	743	1.4
Inhalants	227	2.5	742	2.4
Total for Treatment Section	228	13.0	745	10.1

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.3.23 Average Time to Complete Treatment Sections, by Whether Consistency Checks Were Triggered in Each Particular Section				
	Average Time to Complete			
	Consistency Check Triggered		No Consistency Check Triggered	
Questionnaire Section	*n*	Time (minutes)	*n*	Time (minutes)
Tobacco	68	3.2	904	1.5
Alcohol	137	4.8	835	3.1
Marijuana	33	2.4	939	0.9
Cocaine	9	2.0	963	0.4
Crack	4	1.5	511	0.4
Heroin	4	1.6	968	0.3
Hallucinogens	19	2.9	952	1.4
Inhalants	19	4.3	950	2.4
Total for Treatment Section	228	13.0	745	10.1

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.3.24 Breakoff Rates, by Consistency Check Status				
	Breakoff Rate			
	Consistency Check Present		Consistency Check Absent	
Cases Included	*n*	%	*n*	%
Cases Ever Recorded as Breakoff Even if Completed Later	77	7.9	86	8.5
Cases Finalized as Breakoff	16	1.6	16	1.6
Total Number of Respondents Assigned to Group Treatment	975		1,007	

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.3.25 Respondent Reactions to CAI Interview, by Consistency Check Status				
	Consistency Check Present		Consistency Check Absent	
Item	%	*n*	%	*n*
How Difficult to Use Computer?				
Very difficult	2.0	19	2.3	23
Somewhat difficult	6.4	61	6.3	63
Not at all difficult	91.6	879	91.4	907
Total	100.0	959	100.0	993
Able to Enter Answers Easily?				
Yes	90.0	864	86.6	860
No	10.0	96	13.4	133
Total	100.0	960	100.0	993
Wanted to Change Previous Answer But Did Not?				
Yes	18.2	174	18.5	183
No	81.8	783	81.5	806
Total	100.0	957	100.0	989
Level of Interest in Interview				
Very boring	6.2	59	7.0	69
Somewhat boring	10.2	97	13.2	130
Neither boring nor interesting	28.2	269	24.1	238
Somewhat interesting	34.5	329	34.4	340
Very interesting	20.9	200	21.3	210
Total	100.0	954	100.0	987
Accuracy and Completeness of Answers				
Very accurate and complete	84.4	807	82.1	813
Fairly accurate and complete	14.9	142	16.4	162
Not very accurate and complete	0.7	7	1.5	15
Total	100.0	956	100.0	990

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 7.3.26 Interviewer Reports of Respondent Difficulties, by Consistency Check Status				
	Consistency Check Present		Consistency Check Absent	
Item	%	*n*	%	*n*
Respondent's Understanding of Interview				
No difficulty	84.4	810	81.4	807
Just a little difficulty	11.7	112	12.9	128
A fair amount of difficulty	3.1	30	4.3	42
A lot of difficulty	0.8	8	1.4	14
Total	100.0	960	100.0	991
How Cooperative Was Respondent?				
Very cooperative	92.4	888	92.1	913
Fairly cooperative	6.7	64	7.2	71
Not very cooperative	0.8	8	0.4	4
Openly hostile	0.1	1	0.3	3
Total	100.0	961	100.0	991
Respondent's Interest in ACASI				
Very interested	62.7	602	59.9	592
Somewhat interested	32.0	307	34.2	338
Not very interested	5.3	51	5.9	58
Total	100.0	960	100.0	988
How Necessary Was Tutorial?				
Necessary	23.0	221	22.4	221
Useful but not necessary	50.0	479	50.5	499
Unnecessary	27.0	260	27.1	268
Total	100.0	960	100.0	988
How Necessary Was Audio?				
Necessary	17.4	167	17.9	177
Useful but not necessary	48.1	461	45.3	448
Unnecessary	34.5	331	36.8	363
Total	100.0	959	100.0	988
How Often Did Respondent Let Answers Be Known?				
None of the time	82.9	796	81.0	800
A little of the time	12.6	121	13.5	133
Some of the time	2.8	27	3.5	35
A lot of the time	1.3	12	1.5	15
All of the time	0.4	4	0.5	5
Total	100.0	960	100.0	988

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

7.4 Effect of ACASI Experimental Factors: Modeling Procedures and Effects on Prevalence

In this section, we summarize the statistical methods used to assess ACASI experimental factors on the prevalence of drug use, the time to complete different parts of the interview, and the breakoff rates. Logistic regression and Cox proportional hazards models were used to evaluate the various ACASI experimental factors and to compare ACASI versus PAPI on binary and time-to-event responses. Due to the complex sample design, we used the SUrvey DAta ANalysis (SUDAAN) software package Release 7.5 (Shah, Barnwell, & Bieler, 1998) to account for stratification, clustering, and unequal weighting, where appropriate. SUDAAN's LOGISTIC and SURVIVAL procedures were used to fit the regression models. In addition, Release 6.12 of SAS software was used to estimate weighted prevalence rates of drug use, and the SAS procedures PHREG and LIFETEST were used to estimate the unweighted median time to complete different parts of the interview.

The ACASI experimental factors evaluated in this report include the contingent questioning structure (single vs. multiple gate questions), the consistency checks (absent vs. present), and the multiple chances to report use (absent vs. present). Statistical models for evaluating the ACASI experimental factors also included the following covariates:

1. the field interviewer's (FI's) rating of interview privacy (a three-level categorical variable describing how often the respondent let the FI know his or her answers, coded 1=none of the time, 2=some of the time, and 3=all of the time),
2. the FI's rating of the degree of nonprivacy/distraction during the interview (coded 1-9, with Level 1 being the most private and Level 9 the least),
3. age in years (12 to 17 vs. 18 or older),
4. gender (male vs. female),
5. education level (less than high school, high school, greater than high school), and
6. race/ethnicity (Hispanic, black, and white).

In addition to evaluating the main effects of the three experimental factors (contingent questioning structures, multiple use questions, and consistency checks) and covariates, we also evaluated all two-way and three-way interactions among the three ACASI experimental factors.

To compare the ACASI experimental factors on prevalence of drug use, breakoff rates, and the time to complete different parts of the interview, a model containing all the main effects, covariates, and interactions was fit first:

Model 1: Main Effects + Covariates + Interactions

1. ACASI Main Effects (3 degrees of freedom): (a) contingent questioning, (b) multiple use, and (c) consistency checks.
2. Covariates (9 degrees of freedom): (a) gender, (b) age group, (c) race/ethnicity, (d) education, (e) letting FI know answers, and (f) degree of distraction/nonprivacy.
3. Interactions (4 degrees of freedom): (a) contingent questioning * multiple use, (b) contingent questioning * consistency checks, (c) multiple use * consistency checks, and (d) contingent questioning * multiple use * consistency checks.

Wald chi-square tests were used to evaluate the main effects, covariates, and interactions of interest. This interaction model was used to determine if the global 4 degrees-of-freedom interaction effect (containing all two-way and three-way interactions among the ACASI factors) was statistically significant. If the global interaction effect was not significant ($p>0.10$), all two-way and three-way interactions were removed, and a new main effects model (Model 2 below) was fit containing only ACASI main effects and covariates.

Model 2: Main Effects + Covariates

1. ACASI Main Effects (3 degrees of freedom): (a) contingent questioning, (b) multiple use, and (c) consistency checks.
2. Covariates (9 degrees of freedom): (a) gender, (b) age group, (c) race/ethnicity, (d) education, (e) letting FI know answers, and (f) degree of distraction/nonprivacy.

The ACASI experimental factors were then evaluated in the reduced main effects model. In the reduced model, each main effect was automatically adjusted for all other main effects and covariates present in the model.

However, if a significant global interaction effect (4 degrees of freedom) was detected ($p<0.10$) in Model 1, further modeling was done to evaluate specifically which two-way and possibly three-way interaction effect(s) were statistically significant. When we

determined that a particular ACASI experimental factor, *X*, significantly interacted with another ACASI factor, *Y*, then factor *X* was evaluated *separately* within each level of factor *Y* (these effects are sometimes called "simple effects" or "effect slices").

The Cox proportional hazards model was used to detect overall heterogeneity in time-to-completion distributions among the ACASI experimental factors. Median event times can be computed from the model. The Cox proportional hazards model is stated as follows:

$$\log\left(\frac{h(t|\boldsymbol{x})}{h_0(t)}\right) = \boldsymbol{X\beta},$$

where *t* represents the time to event, $\boldsymbol{X\beta}$ is the linear combination of covariate effects, and $\frac{h(t|\boldsymbol{x})}{h_0(t)}$ is the *relative risk* (or *hazard ratio*) for the event at time *t* for respondents in different ACASI treatment groups.

Using this model, we were able to quantify the relative effect of treatment combinations on the event times. The model requires the user to supply each respondent's event time, an indicator for censoring ("yes" vs. "no"), and a set of covariate values. All event times, censored and complete, were used in the analysis. A Wald chi-square test was used to evaluate the effect of treatment on event time. The logistic regression model was used to examine the results for prevalence of drug use.

Exhibit 7.4.1 summarizes the statistically significant ACASI treatment effects found. During the analyses, we looked at both unweighted and weighted data for cigarettes, alcohol, marijuana, and cocaine. Exhibit 7.4.1 contains the outcomes of interest, the model effects, the p value(s) and odds ratio(s) corresponding to the effects, and the adjusted prevalence rate. Prevalence rates and model effects were adjusted for all other variables in the model. Odds ratios greater than 1.0 indicate increased prevalence in a particular treatment group versus the reference cell, while odds ratios less than 1.0 indicate reduced prevalence in a particular treatment group.

In the discussion, statistical significance refers to $p<0.05$ for main effects and $p<0.10$ for interactions, unless otherwise specified. Marginal significance refers to $0.05< p<0.10$. The following summarizes the significant results, both unweighted and weighted:

Cigarettes

1. The presence of multiple chances to report use significantly reduced the reported prevalence of lifetime use from 59.1% to 53.7% (unweighted only).
2. The reported prevalence of past month use (weighted only) was marginally increased ($0.05< p< 0.10$) when respondents had multiple chances to report use (19.9% vs. 29.4% for absent vs. present) and when consistency checks were present (20.4% vs. 28.8% for absent vs. present).

Alcohol

1. The presence of consistency checks marginally increased ($0.05< p<0.10$) the reported prevalence of past year use (weighted only) from 67.2% to 74.6%.
2. The presence of multiple chances to report use significantly reduced the reported prevalence of past month use (unweighted only) from 29.7% to 25.2%.

Marijuana

1. The multiple gate questions significantly reduced the reported prevalence of past year use (weighted only) from 10.8% to 6.2%.
2. The presence of consistency checks significantly increased the reported prevalence of past month use (unweighted only) from 4.7% to 8.1% but only for the multiple gate version (this was reflected in a significant interaction effect between the consistency checks and the contingent questioning structure).

Cocaine

1. Lifetime use was significantly increased in the presence of multiple chances to report use (unweighted only) from 2.9% to 5.5% but only for the multiple gate questions (there was significant interaction between the multiple chances to report use and the type of contingent questioning used).
2. Lifetime use was significantly reduced when multiple gate questions were used (weighted only, 14.6% vs. 7.4% for single vs. multiple gate).
3. Past year use was significantly increased in the presence of multiple chances to report use (unweighted and weighted) but only for the multiple gate questions (there was significant interaction between the multiple chances to report use and the contingent questioning structure).
4. Past year use was significantly reduced in the presence of consistency checks (unweighted and weighted) but only in the presence of multiple chances to report use (there was significant interaction between the consistency checks and the multiple use treatment).
5. Past month use was significantly increased in the presence of consistency checks (unweighted and weighted) but only for the multiple gate questions (there was significant interaction between the consistency checks and the contingent questioning structure).
6. Past month use was marginally reduced ($0.05<p<0.10$, unweighted only) in the presence of consistency checks butonly when respondents had multiple chances to report use (there was significant interaction between the consistency checks and the multiple chances).

Any Illicit Drug Use

1. Lifetime use was marginally reduced ($0.05< p<0.10$) in the presence of multiple chances to report use, from 43.1% absent to 38.5% present (unweighted only).

Exhibit 7.4.1 Summary of Statistically Significant Modeling Results: Prevalence of Drug Use Among ACASI Treatment Groups

Outcome	Unweighted or Weighted	Effect	p Value	Odds Ratio	Adjusted Prevalence Rate
Cigarettes					
Lifetime	Unweighted	Multiple Use: Present vs. Absent	0.0144	0.80	59.1% Absent 53.7% Present
Past Month	Weighted	Multiple Use: Present vs. Absent	0.0780	1.67	19.9% Absent 29.4% Present
		Consistency Checks: Present vs. Absent	0.0675	1.58	20.4% Absent 28.8% Present
Alcohol					
Past Year	Weighted	Consistency Checks: Present vs. Absent	0.0817	1.44	67.2% Absent 74.6% Present
Past Month	Unweighted	Multiple Use: Present vs. Absent	0.0245	0.80	29.7% Absent 25.2% Present
Marijuana					
Past Year	Weighted	Multiple vs. Single Gate	0.0179	0.54	10.8% Single 6.2% Multiple
Past Month	Unweighted	Consistency Checks: Present vs. Absent	0.0237	1.93	4.7% Absent 8.1% Present

		(Routing=Multiple Gate)			
Cocaine					
Lifetime	Unweighted	Multiple Use: Present vs. Absent (Routing=Multiple Gate)	0.0165	2.45	2.9% Absent 5.5% Present
Lifetime	Weighted	Multiple vs. Single Gate	0.0280	0.47	14.6% Single 7.4% Multiple
Past Year	Unweighted	Multiple Use: Present vs. Absent (Routing=Multiple Gate)	0.0124	--	1.2% Absent 2.5% Present
		Consistency Check: Present vs. Absent (Multiple Use=Present)	0.0162	--	3.3% Absent 1.2% Present
Past Year	Weighted	Multiple Use: Present vs. Absent (Routing=Multiple Gate)	0.0079	--	0.4% Absent 1.5% Present
		Consistency Check: Present vs. Absent (Multiple Use=Present)	0.0079	--	2.3% Absent 0.6% Present
Past Month	Unweighted	Consistency Check: Present vs. Absent (Routing=Multiple Gate)	0.0397	--	0.24% Absent 0.61% Present
		Consistency Check: Present vs. Absent (Multiple Use=Present)	0.0690	--	0.88% Absent 0.26% Present
Past Month	Weighted	Consistency Check: Present vs. Absent (Routing=Multiple Gate)	0.0032	11.03	0.07% Absent 0.44% Present
Any Illicit Drug Use					
Lifetime	Unweighted	Multiple Use: Present vs. Absent	0.0633	0.88	43.1% Absent 38.5% Present

[1] Any illicit drug includes marijuana, cocaine, crack, heroin, inhalants, hallucinogens, and nonmedical use of analgesics, sedatives, stimulants, and tranquilizers.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures: 1997 Field Experiment.

7.5 Effect of ACASI Experimental Factors on the Time to Complete Different Parts of the Interview

We calculated the unweighted median time to complete different parts of the interview for each of the eight ACASI treatment combinations. Cox Proportional Hazards Regression Modeling was used to evaluate the effects of ACASI experimental factors on the completion times for the following three outcomes of interest:

1. the total core time (all CAI sections through the sedatives section),
2. the core experimental time (tobacco through inhalants sections), and
3. the remaining ACASI time (from inhalants to the end).

For each time variable, respondents who were not at risk at the start of the reference period had a missing value for that time variable and were therefore not included in the analysis. Respondents who broke off during the reference period were censored during that time and had a censoring indicator value corresponding to "yes." Their time value contains partial information that was used in this analysis.

The following information summarizes the significant results for each variable. We fit models using both weighted and unweighted data; however, only the unweighted results are presented because they apply to the actual sample. This is most appropriate for an operational measure.

7.5.1 Total Core Time, Unweighted

1. Multiple gate versions significantly delayed the time to completion compared to the single gate version (18.65 vs. 17.58 minutes with an estimated hazard ratio of 0.84, $p<0.01$).
2. Multiple chances to report use significantly delayed the time to completion (the median time to complete was 18.42 vs. 17.77 minutes with an estimated hazard ratio of 0.90, $p<0.01$).
3. There was no significant effect on time to complete based on whether consistency checks were present or absent.
4. Among the covariates, age, education, race/ethnicity, and the degree of non-privacy/distraction had significant effects on the time to complete, with significantly shortened times for those aged 12 to 17 years compared to those aged 18 or older but significantly delayed times for those with less thana high school education and high school graduates compared to those with more than a high school education, for Hispanics and blacks compared to whites, and for every unit increase in distraction.

7.5.2 Core Treatment Time (Tobacco Through Inhalants), Unweighted

1. Multiple gate versions significantly delayed the time to completion compared to the single gate version (the median time to complete was 10.6 vs. 9.77 minutes for multiple vs. single gate, respectively, with the estimated hazard ratio at 0.80, $p<0.01$).
2. Multiple chances to report use significantly delayed the time to completion (the median time to complete was 10.3 vs. 9.95 minutes for presence vs. absence, respectively, with the hazard ratio at 0.89, $p=0.014$).
3. The presence of consistency checks also significantly delayed the time to completion (the median time to complete was 10.27 vs. 9.97 minutes for the presence vs. absence of consistency checks, with an estimated hazard ratio of 0.91, $p=0.028$).
4. Among the covariates, age group, education, race/ethnicity, letting the FI know the answers, and the degree of nonprivacy/distraction had significant effects on time to complete, with significantly shortened times for those aged 12 to 17 years compared to those aged 18 years or older, but with significantly delayed times for those with less than high school education and high school graduates compared to those with more than a high school education, Hispanics compared to whites, letting the FI know the answers some of the time compared to all of the time, and for every unit increase in distraction. Hazard ratios for shortened times were greater than 1.0, while hazard ratios for delayed times were less than 1.0.
5. We examined the impact of the time required to complete the core on subsequent breakoffs. The analysis was based on only those respondents who were not censored before or during the core treatment sections ($n=1,968$) or, in other

words, at risk for subsequent breakoff. No other covariates or treatment factors were included in the model. Results indicate that subsequent breakoff after core treatment was significantly associated with a delay in the median time to complete the core treatment. The median times were 10.7 versus 10.05 minutes for the presence versus absence of subsequent breakoff, with an estimated hazard ratio of 0.84, p=0.033.

7.5.3 Remaining ACASI Time After Inhalants, Unweighted

Although ACASI treatments did not vary in this section, we examined these times to determine if the treatment assigned to a respondent in the tobacco through inhalants sections had a carryover effect.

1. None of the previously assigned ACASI experimental factors had a significant effect on the time to complete the remaining ACASI after the inhalants section.
2. Among the covariates, education, race/ethnicity, and the degree of non-privacy/distraction had significant effects on the time to complete, with significantly delayed times for those with less than a high school education and high school graduates compared to those with a high school education, Hispanics and blacks compared to whites, and for every unit increase in distraction. Hazard ratios for delayed times were less than 1.0.

Although there were multiple results that were significantly different, none of the time differences was large enough to point toward choosing one method over another.

7.6 Effect of ACASI Experimental Factors on Breakoff Rates

In this section, we describe the results of logistic regression modeling to evaluate the effects of ACASI experimental factors on the incidence of breakoffs anywhere in the interview.

Some additional information corresponding to these analyses and results follows. We present only unweighted modeling results. A table of results (Exhibit 7.6.1) contains the model effects and their corresponding p value(s) and odds ratio(s). Odds ratios greater than 1.0 indicate an increase in the probability of a breakoff for a particular group compared to the reference group, and an odds ratio less than 1.0 indicates a reduction in the probability of a breakoff. Effects are adjusted for all other variables in the model, unless otherwise indicated. Statistical significance refers to p<0.05, unless otherwise specified. The 12 to 17 year olds were classified as less than high school graduates to include all respondents in one analysis. Privacy covariates were not included in this model because their missing values would have removed all but four breakoffs from the analysis. A total of 32 breakoff cases out of 1,982 respondents were used in this analysis.

The following information summarizes the results from this analysis. None of the ACASI experimental factors significantly affected the incidence of breakoffs. None of the covariates significantly affected the incidence of breakoffs.

Exhibit 7.6.1 Logistic Regression Modeling Results: Breakoff Anywhere in the Interview (Unweighted Results)

Outcome	Effect	Contrast	*p* Value	Odds Ratio
Intstat: Breakoff indicator (1=*yes*; 0=*no*)	Interactions Involving Main Effects (4 df):	-	0.1372	--
	Contingent Questioning Structure	Multiple Gate vs. Single Gate	0.6170	1.20
	Multiple Chance	Present vs. Absent	0.8004	1.10
	Consistency Checks	Present vs. Absent	0.8636	1.05
	Gender	Male vs. Female	0.3224	1.49
	Age Group	12-17 vs. 18+	0.9633	1.03
	Education	Overall (2 df)	0.7550	

		< High School vs. > High School	0.7139	0.76
		High School vs. > High School	0.4573	0.61
	Race	Overall (2 df)	0.9692	
		Hispanic vs. White	0.8876	0.92
		Black vs. White	0.8653	1.07

Note: There were 32 break off cases out of 1,982 respondents. Privacy covariates are not included in this model because their missing values would remove all but four break off from the analysis.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

[16] Only possible for respondents assigned to the multiple use treatment.

[17] Patterns of inconsistent reporting are discussed more fully in Sections 7.1.2, 7.2.2, and 7.3.2.

8. Development and Testing of an Electronic Screener: 1997 Field Experiment and Subsequent Testing

In this chapter, we describe the development and testing of the electronic screener. The development and testing process included

1. development and small-scale testing of the basic sample selection application,

2. addition of selected case management activities to the application and testing in the 1997 field experiment,

3. revision of the application to include additional case management activities and field testing in August 1998, and

4. final development for the 1999 NHSDA.

8.1 Need for Electronic Screening: Research Issues

To achieve the desired precision for important subpopulations defined by age and race/ethnicity, a large number of households must be screened to identify enough eligible respondents in these subpopulations. Field interviewers (FIs) visit each sample dwelling unit and determine how many of the residents are eligible for the survey and their basic demographic characteristics. Based on this roster of residents, none, one, or two people are selected for the survey. Prior to 1999, the FIs consulted a complex selection table to determine who, if anybody, was selected for the survey. The electronic screener was developed to achieve the following:

1. Eliminate interviewer errors in the selection process, including accidental errors and intentional tampering with the roster by the FI to achieve certain selections.

2. Develop the capability to include more variables in the respondent selection algorithms.

3. Reduce data editing, data entry, and shipping costs.

4. Provide more detailed and timely information on field activities.

8.2 Developing and Testing of the Screener for the 1997 Field Experiment

8.2.1 Developing the Application

We first needed to choose a piece of hardware for the application. Several hardware options were explored to determine which would be the most feasible and cost effective for the desired goal. Among these were the Slate-Top, a laptop, and the Newton. After a careful review, the Newton 2000 handheld computer was selected as the best choice. The Slate-Top was considered for both the screening and the computer-assisted interviewing (CAI) and found to be unable to handle the computer-assisted personal interviewing (CAPI), and its cost was excessive for hardware that was to perform only screenings. In addition, because of its size and weight, we concluded that using a laptop was not feasible.

FormLogic from Wright Strategies, Inc., San Diego, California, was the software selected to develop the screening program. This software package offered a Microsoft Windows development platform, suitable for the screener design requirements. Furthermore, Wright Strategies was willing to work with Research Triangle Institute (RTI) programming staff to ensure timely and robust software support throughout the entire development process.

Development of the Newton NHSDA screening application began in early 1997. The development was an iterative process with extensive usability testing conducted after each cycle of development. These initial development stages focused on programming the basic rostering and sample selection activities that are done after a member of a dwelling unit has been contacted. This was the key function that had to be successfully completed for the electronic screener to be a viable alternative to the paper-based procedures. However, a number of other case management activities are integrated with the screening procedures during the NHSDA, including determination of vacancy status, record of contacts and attempted contacts and their outcomes, and so on. This additional functionality was incrementally added to the application.

In May 1997, we conducted a small field feasibility test in Durham, North Carolina, to test the Newton application in realistic field conditions. Following this feasibility test, additional development was completed, and the Newton application for the 1997 field experiment was finalized in September 1997.

The cumulative results of our testing convinced us that the basic rostering and sample selection functions were ready to be tested in the 1997 field experiment. However, near the start of the field experiment, the full application was discovered to still contain too many bugs for production use. Therefore, the application was scaled back to just the critical functions of rostering households and selecting respondents. We transferred the other tasks to a paper-and-pencil contact worksheet and to the case management system (CMS) that was on the Toshiba laptop. Also, we found that the Newton handwriting feature was not always reliable and user friendly; therefore, we made the option of entering numeric or alphabetic responses via a keypad or keyboard available to the FIs.

8.2.2 Training Interviewers on Use of Electronic Screener

During three training sessions, a total of 171 interviewers were trained in the use of the Newton in the 1997 field experiment. From September 22 through 28, 1998, 144 interviewers were trained in Los Angeles and Research Triangle Park (RTP) in North Carolina. Two additional sessions were held at RTI from November 6 to 11 and from November 30 through December 3, 1997, with 23 and 4 additional interviewers, respectively.

In the first training session, five training rooms at each site were each staffed by a lead trainer, two assistant trainers, and a computer support technician. The number of interviewers in a training room ranged from 12 to 17. The follow-up training sessions at RTI were completed in one training room with all interviewers present. Initial orientation to the Newton focused on introducing the hardware and covered general information on how to turn the Newton on and off, care for it, change its batteries,

calibrate the pen, use the pen to select responses, use the backlighting feature, and connect it to the Toshiba laptop for nightly transmissions. A basic introduction to the NHSDA project and locating sample dwelling units was then given, after which the FIs were returned to the Newton training to perform group exercises of increasing difficulty and to participate in paired mock trials.

Both trainers and interviewers were asked to evaluate the training session. Interviewers provided their input via a standardized questionnaire completed at the end of training. Input from the 30 staff, who worked as trainers, was solicited in an electronic mail message after all trainers had left the training site; trainer responses were simply open-ended text. The comments from both groups are summarized below.

In general, trainers felt the portion of the training session devoted to the Newton went well; however, they noted two specific problem areas. First, they noted that training on the Newton was difficult because the Newton screen could not be projected for the entire class to see in the same way that a laptop screen can be projected: The Newton has no screen "output" functionality. During the course of preparing for the training, several possible ways for training interviewers to use the Newton were discussed, including using no visuals during the Newton portion of training, preparing overhead transparencies of the required screens, displaying the screens using PowerPoint, and using small television cameras to display the screens. We chose to display the screens using PowerPoint because we believed this allowed for maximum flexibility at a reasonable cost. However, this decision tied the trainers to the specific scenarios prepared in advance of training. Trainers had more difficulty leading their interviewers through impromptu exercises because the interviewers could not look at a display to know that they were in the right place.

Second, nearly half the trainers thought was that the screening exercises were too complicated. In our attempt to cover all possible field situations that an interviewer might encounter, trainers felt the screening exercises became too difficult for the interviewers to handle while at the same time trying to become familiar with the Newton. Most trainers suggested that the first few screening exercises be straightforward (e.g., no addition of family members after the rostering is complete, no deleting of household members who do not really live at the housing unit). Trainers felt that this would give the interviewers a chance to learn the basic process before introducing the more complicated components. In addition, trainers noted that some interviewers new to field data collection began to worry that all household screenings would be as complicated as our practice scenarios. Trainers had to reassure these interviewers that this would not be the case and that most screening respondents would provide the roster information easily and completely. The trainers recommended that once a set of straightforward screenings has been completed, more complicated scenarios could be introduced to cover all the "what if" situations. With more time at training to cover the Newton, this progression of screening exercises should work well.

As part of the training evaluation, interviewers were asked to rate the amount of time spent on learning how to use the Newton and on how to transmit data between the Newton and the Toshiba. A majority of the interviewers felt the amount of time spent learning to use the Newton was about right. A small number of interviewers (7%) at the RTP training session felt that too much time was spent on the Newton, whereas only 1% of the interviewers trained in Los Angeles felt this way. Perhaps more of the RTP interviewers had previous interviewing experience and thus were able to pick up the Newton skills more easily. Similar results were also seen on the amount of time spent learning how to transmit data between the Newton and the Toshiba. Again, the majority of interviewers felt that the amount of time spent on this activity was about right. However, fewer interviewers felt that too much time had been spent on this activity.

Interviewers were also given the opportunity to record open-ended comments regarding the homestudy materials. A number of interviewers noted that the portion of the manual that covered the

Newton was especially useful. The screen images were included, as were step-by-step instructions, and these items were noted to be some of the more useful parts of the manual. Several interviewers suggested that even more information of this type should be included in the future. Interviewers also noted that a "cheat sheet" of Newton instructions would be a useful addition to the manual, perhaps as a detachable or laminated card that could be kept in the Newton carrying case on in the Toshiba bag. One interviewer noted that it would have been useful to receive the Newton prior to training so that she could work with it as she reviewed the manual.

Many interviewers commented that the Newton should have been introduced earlier in the training session. Several interviewers noted that the first few days of training were not particularly intensive and the last few days were exceptionally challenging because all the hardware was being covered. They felt the Newton (and the Toshiba) should be introduced on the first day of training and all procedures related to the hardware should be covered before such topics as locating the area segment, introducing the study, or refusal conversion are discussed.

8.2.3 Implementation Results: Interviewer and Respondent Reactions

To obtain information on the implementation, we conducted two debriefing teleconferences with a sample of interviewers: one during the third week of the field experiment and one following the end of the 1997 field experiment. The purpose of these calls was to hear from interviewers about their initial experiences using the Newton screening application. A total of 16 interviewers took part in the debriefings, with no interviewer taking part in any two sessions. Each call was moderated by an RTI staff member. In general, interviewers had favorable things to say about the Newton and the electronic screening application. More specific comments related to the different facets of working with the Newton are summarized below.

Respondent reactions. Some interviewers felt that the Newton gave them a more professional appearance on the dwelling unit's doorstep, which resulted in respondents being more likely to provide the screening information. Interviewers also noted that they felt more professional because they did not have to shuffle a bunch of papers at the doorstep while the respondent waited. Interviewers also commented that some of their respondents were interested in how the Newton worked and liked to watch as the interviewer entered the screening information. Finally, interviewers noted that because the screening moves so quickly on the Newton, respondents did not become annoyed by the amount of time needed to complete the screening.

Battery usage. Interviewers seemed to fall into two groups with regard to battery usage: those who had used several sets of batteries and those who were still working with their first set. The primary reason for the difference appeared to be the amount of time the Newton was set in the "backlight" mode. The backlighting mode, which causes the screen to appear brighter, drains the batteries more quickly; interviewers who use this mode frequently use more battery power.

Screen readability. All interviewers reported that the Newton screen is very difficult to read in bright sunlight. This problem was reported both by interviewers who regularly use the backlighting mode as well as by those who do not. To be able to read the screen, interviewers either had to move to a more shaded spot or try to cover the screen with their hand. Interviewers reported that this is annoying but does not seem to result in lost screenings (i.e., respondents who are unwilling to move to another location either inside or outside the dwelling unit to complete the screening). Interviewers also noted that some of the text on the Newton was too small to be easily seen. They requested that all text be printed in the same point size as the text on the introduction screen, which was a slightly larger point size.

Newton "pens." All interviewers reported that they still had the original Newton "pen" in their possession. Most interviewers were using the original pen, but a couple of interviewers reported a preference for the replacement pens that were provided at training. Several reasons were given for this preference. The replacement pen is easier to hold, the replacement pen fits easily inside the Newton carrying case, and the original pen is too expensive and they are afraid of losing it.

Carrying cases. Most interviewers seemed to like the leather carrying cases. They felt the case protected the Newton well. However, one interviewer had experienced a problem with the zipper on the side of the carrying case causing the flashcard (i.e., the device that stores the interviewer's assignment) to be popped out of her Newton. Because the Newton was in the case, the interviewer could not tell that the flashcard had come out until she was ready to screen a case. In this case, simply pushing the flashcard back in solved the problem. However, the interviewer noted that it would be wise to inform all interviewers of this possibility and to instruct staff to keep the side zipper partially unzipped to minimize the chance of the flashcard becoming dislodged. All interviewers were informed of this problem and advised of the corrective action to take.

Size and shape. Interviewers reported the Newton was easy to hold and use. Several interviewers said they rested the Newton on their clipboard to complete the screenings. Others said they simply held the Newton in their hand. The weight of the Newton did not appear to be a problem.

Screening program. Interviewers had a number of comments regarding changes that could be made to the screening program. Several noted that it would be nice to have all information on finalized screenings resident on the Newton so that the list of pending screening cases would be accurate.[1] Interviewers also noted that there should be some way to go back into a case after the screening is completed so that corrections could be made if necessary. In general, interviewers would like to have additional functionality on the Newton, specifically, the ability to see the status of all cases and the ability to use some of the Newton's other functions to schedule appointments, organize work time, and so on. Several interviewers reported that it was difficult to use the "page down" function to scroll through their cases; they noted that it is very easy to select a case rather than scrolling down. They also noted that when new segments get added to their assignment, they may need to page down several screens to get to the case they want to work. Each time the Newton is turned off and back on, the cursor moved back to the top of the list, then the interviewer had to scroll down again; interviewers noted that this was annoying and time consuming. Interviewers also felt that time was wasted by having each screening question come up each time. They noted that after completing a few screenings they knew how to ask these questions and did not need to see the text of each question every time.

Training issues. Interviewers were asked whether there was additional information they wish had been provided at training. One interviewer mentioned that there was too much emphasis at training on doing everything correctly and not making any mistakes. She felt that interviewers should be encouraged to make mistakes during training to see what will happen so that when they are out in the field they will know what to do. Another interviewer said that additional time should have been spent discussing how to correctly screen vacant housing units. Also, some interviewers felt the screening exercises were too complex; these interviewers felt that there should have been more "straightforward" practice exercises before moving into the complex situations that are not likely to happen as often in the field.

Finally, interviewers were asked to report the thing they liked best about the Newton and the thing they liked least. The items they mentioned are listed below:

[1]This feature was added when the CMS was moved from the laptop to the Newton.

Like Best	Like Least
1. It makes the screening go faster.	1. Can't go back in to a finalized case.
2. The selections are made for you.	2. Bad glare when working in bright sun.
3. It isn't heavy/It is portable.	3. Can't remove some cases from pending status.
4. It is easy to work with.	4. Questions are too repetitive.
5. It looks professional.	5. Print on some screens is too small.

Field interviewer debriefing questionnaire. At the end of the 1997 field experiment, all interviewers were asked to complete a questionnaire concerning the overall issues of the survey as a whole. A total of 142 surveys were completed, and their results correspond very similarly to the comments given in the debriefing teleconferences.

Overall, approximately 89% of all reporting interviewers stated that some difficulty with the Newton did occur (62.9% actually having problems at the doorstep of the dwelling unit), but only 17% reported difficulty on a usual basis. There were 208 mentions of the causes of difficulty with the Newton (see **Exhibit 8.2.1**).

In regard to battery usage and lighting problems, 97% reported using their backlighting, which would account for the interviewer's concern about the decrease in battery charge life. "Worrying about changing batteries" was the third least liked aspect of the Newton as reported in the end of study survey questionnaire.

On another note, 45.4% of interviewers usually carried their Toshiba when screening a household. Of this percentage, 31.3 % reported a positive effect on the screening and 1.6% reported a negative effect. This positive effect corresponded with the interviewer's notion that the Toshiba gave an added look of professionalism and assisted with their ability to gain cooperation.

Screener debriefing questions. As part of the 1997 field experiment's CMS, interviewers were prompted to answer a short series of questions for each screening they finalized. These questions were brought up automatically by the system when the interviewer transferred data from the Newton to the laptop computer. Two of these questions related specifically to the Newton screening application. Interviewers were first asked whether they thought that the presence of the Newton influenced the respondent's decision to complete screening. If the interviewer indicated that it did, a second question asked whether the Newton influenced the respondent in a positive or negative way. Results from these two questions are presented in **Exhibit 8.2.2**.

The interviewers did not believe that the Newton had an effect on the screening outcome in nearly two thirds of the cases. However, for those cases where the Newton did seem to influence the respondent, the effect was nearly always positive (94.1% vs. 5.7%). This was seen as an important finding in that we did not want to implement a new screening procedure that forces interviewers to counter additional resistance.

8.2.4 Implementation Results: Hardware, Software, and User Problems

We monitored the field situation closely. Technical support staff entered each call received from the field into a problem report log. This log allowed us to determine the nature and number of problems being reported by our interviewing staff. Problems entered into the log after the first 5 weeks of the field experiment were reviewed, and those that relate to the Newton screener are summarized below.

A total of 177 problems were reported that were classified as related to the use of the Newton. These 177 problems were spread over 94 interviewers (approximately 55% of the original 171 interviewing staff). The largest number of problems reported by a single interviewer was seven, with most interviewers appearing on the log only once or twice. It is worth noting that 11 (including the FI with seven calls to RTI's technical support) of the 94 interviewers who reported Newton problems left the field experiment; either because they quit or were asked to resign.

At a broad level, the problems experienced by the interviewers can be coded into three categories:

1. interviewer errors,

2. hardware issues, and

3. errors due to interaction of interviewer and hardware.

Interviewer mistakes made up the largest proportion of the total number of problems. A total of 74 problems (42%) were attributable solely to the interviewer. " Hand-holding" was the most common solution to these interviewer mistakes. Nearly half of the interviewer mistakes required the technical support person to review a procedure that had been taught during training and/or was discussed in the interviewers' manual. Most of the remaining interviewer mistakes occurred when the interviewer either screened a household on the wrong line number, finalized the roster and implemented the selection process before fully reviewing the roster for accuracy, or finalized a household as ineligible by mistake (e.g., by incorrectly indicating that no one at the housing unit would live there for most of the quarter). Based on the frequency of these types of mistakes, in subsequent applications we developed procedures that allowed interviewers to "reactivate"lines or move roster records from one line to another with the permission of their supervisor.

An additional 49 problems were attributable to hardware or software problems. In some cases, the Newton suffered a complete failure and the unit had to be returned to RTI and a replacement sent to the interviewer (*n*=5) or hardware (such as a new flashcard or connector cord) had to be sent (n=5). However, the most common problem in this category (n=24) occurred when the Newton simply "died" and could not be revived by the interviewer using the procedures taught during training. This happens most often when the Newton was without power (no batteries or AC charge) for an extended period of time. In these situations, technical support staff guided the interviewers through a "Hard Reset" of the Newton using the reset button on the back of the Newton. Additionally, there were 10 reported problems involving a lost username and password, which affected the linking between the Newton and the Toshiba. These problems were fixed by the technical support staff by walking the FIs through a process to reset the information. Also, there were five other incidences of linking problems; however, there was insubstantial evidence to conclude whether it was definitely a user-related error or hardware error. Subsequently, these problems were fixed when a new link was established.

The last grouping of problems (54 in all) called into RTI's technical support staff are identified as errors that occurred as a result of both FI interaction and the hardware. A third of these problems were caused by the flashcard popping out of its slot as a result of the zipper on the carrying case pressing on the release button or the flashcard not being properly inserted. The flashcard is the device that stores the interviewer's assignment. The screening application itself is stored internally in the Newton. The decision was made to store the assignment on the flashcard so that if a Newton broke, the interviewer would be able to easily pop the flashcard out and load it into a replacement Newton. Because the carrying cases were delivered only 2 days before training began, we did not realize this would be a problem until data collection was actually under way. In addition, the cases were leather causing the Newton to overheat and interviewers had to remove the Newton from the case to transmit the information.

Different cases were purchased for the 1999 NHSDA. These new cases were not leather and did not cause the Newton to overheat, had velcro fasteners, and allowed the interviewers to transmit data without removing the case.

Some of the remaining problems (n=18) related to the Newton occurred as a direct result of the flashcard popping out. These resulted from FIs not correctly following the Newton messages that followed the popping out of the flashcard and replacing it. As a result, either the FormLogic icon was missing or all cases were "missing." Also during training, interviewers were instructed to exit out of the screening program after every completed screening. However, in several cases (n=13) FIs did not do this, which then resulted in memory problems. This problem was resolved by exiting out of the program.

8.2.5 Implementation Results: Response Rates and Data Quality

Two important considerations in measuring the success of the Newton screening application were the effect of the Newton on screening response rates and on data quality. In this section, comparisons are made between the 1997 field experiment and data collected from those same segments during Quarters 1, 2, and 3, and from data in corresponding primary sampling units (PSUs) for Quarter 4 of the 1997 NHSDA.

Data on eligibility and response rates are presented in **Exhibit 8.2.3**. In general, screening response rates for the NHSDA are quite high. We had not hypothesized that the Newton would have any significant positive impact on screening response rates. However, neither do we wish to see a significant decrease in response rates for electronic screening as compared to paper screening. The data indicate a fall in the screening response rate of the 1997 field experiment in comparison to the two other groups. However, because the interviewers used in the field experiment lacked NHSDA experience, we were unable to attribute this to the use of the Newton. This inability to distinguish between the staff experience and the new procedures was a consequence of our decision to use the ongoing NHSDA as the comparison group. Although this did provide a much larger sample for looking at differences in reporting, it did prevent us from assessing the impact of the screening instrument on the screening response rate.

Because interviewers were no longer required to complete the actual selection process, we did expect that electronic screening would result in improved data quality for the screening data. However, as can be seen from the following results, this anticipated increase in quality was not realized because we had not included the range and edit checks in the application that are the cause of the increased data quality observed in computer-assisted applications. **Exhibit 8.2.4** provides information on the quality of the electronic screener data as compared to the paper form. The percentage of non-blank responses to the demographic data items appears comparable across the three samples. Inconsistencies in the screener data were extremely low for both modes of screening. However, the 1997 field experiment data showed a higher rate of misclassifying the gender with the relationship to the householder. There were 83 inconsistencies for the field experiment as compared with 26 for Quarters 1 to 3 and 42 for Quarter 4. In addition, all three samples had occurrences of husbands and wives who were coded as "never married." In subsequent applications, additional programming logic was added so that such inconsistencies no longer occur. For example, if "daughter" is selected as the relationship, then "female" will automatically be filled for the gender.

Although also extremely rare, an additional type of inconsistency was documented for the electronic screener with regard to the ages of the household members. There were 14 cases in which the age of a household member appeared to be extremely unlikely (although perhaps not impossible) given the relationship code. For example, the householder might be listed as a 32-year-old male. Then a son would be listed with an age of 37. In these cases, it is impossible to know which, if either, of the roster items is incorrect. Also, several of these age inconsistencies were between the head of household, who happened to be male, and a son. However, the son was often classified as female, which led to the

conclusion that FIs were inadvertently tapping son instead of wife because on the relationship screen "wife" is above "son." This also constitutes a possible reason for the high number of female sons/brothers, as seen in Exhibit 8.2.4.

One additional inconsistency was noted where the age of the household member was a number significantly larger than 100. In this case, the son was listed as 120 years old. In cases of this type, it is possible that when the interviewer was writing in the age, he or she also made an additional mark that the Newton identified as a 1; the correct age for the household member was probably 20 or even maybe 12. However, this is impossible to determine after the fact. More restrictive age ranges were programmed into the Newton to eliminate some of these more egregious errors.

Finally, there was one case where the relationship appeared to be incorrect. In this situation, the household composition suggested that the interviewer had confused the relationship between the wife and the son's fiancee, thus switching their respective relationships. For example, the head of household was listed as 69, his live-in partner/fiancee was listed as 68, his son was listed as 22, and his wife was listed as 20. Most likely, the 20-year-old wife was actually the son's fiancee, making her the to-be-daughter-in-law of the householder. However, the original rostering may be correct. Remedies will be made that deal specifically with this confusion.

Exhibit 8.2.1 Causes of Newton Difficulties

Causes of Newton Difficulty	Reported Times	
	Number	Percent
Direct Sunlight	113	54.33%
Other Lighting Problems	36	17.31%
Print Size	38	18.27%
Other Non-Lighting Problems	21	10.09%

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 8.2.2 Results from Screening Debriefing Questions Related to the Newton

In your opinion, did the presence of the Newton influence the respondent's decision whether or not to complete screening?

	Percent	Number
Yes	29.2	2,780
No	68.5	6,512
Don't know	2.3	218
Total	100.0	9,510

Did it influence the respondent in a positive or negative way? That is, did they seem more or less likely to complete screening because of the Newton?

	Percent	Number
Positive way	94.1	2,617
Negative way	5.7	159
Don't know	0.1	4
Total	100.0	2,780

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 8.2.3 Analysis of Eligibility Rates and Response Rates (Unweighted)

Dwelling Unit Statistics	NHSDA Quarters 1-3 Segments	NHSDA Quarter 4 PSUs	1997 Field Experiment
# of Selected DUs	8,449	13,012	14,327
# of Eligible DUs	7,034	11,021	11,994
# of Completed DUs	6,658	10,301	10,243
Eligibility Rate	83.3%	84.7%	83.7%
Response Rate	94.7%	93.5%	85.4%

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 8.2.4Preliminary Analysis of the Quality of the Electronic Screener Data as Compared to the Paper Screener

Screener Data	NHSDA Quarters 1-3 Segments	NHSDA Quarter 4 PSUs	1997 Field Experiment
Total Rostered Persons	13,978	21,237	21,800
% of Rostered Person with a Non-Blank Response			
Relation to Head of Household	99.69	99.56	97.22
Age	100.00	100.00	100.00
Hispanicity	99.11	98.89	99.29
Race/ethnicity	99.18	98.73	98.28
Gender	99.69	99.62	98.39
Marital Status	99.06	98.82	99.27
Inconsistencies			
Total Screenings	6,658	10,301	10,243
# of Female Brothers/Sons	4	2	29
# of Male Sisters/Daughters	1	4	8
# of Female Husbands/Fathers	5	6	13
# of Male Wives/Mothers	9	12	14
# of Never Married Husbands	2	2	4
# of Never Married Wives	5	16	15
Total Inconsistencies	26	42	83

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

8.3 Revision and Additional Testing of Screener: Application Changes

After we reviewed the comments given by both interviewers and trainers and noted the problems logged by our technical support staff, we took measures to correct or minimize some of the problems associated with the Newton electronic screener. Specific problems and their respective remedies are summarized below.

Size of print on screens is too small. The print size was increased to the maximum possible on each screen. This resulted in dramatic improvements on many screens. Unfortunately, we are unable to control the print size in the "pop-up boxes."

Battery problems. Interviewers reported having to change their batteries frequently or other difficulties with the battery. In the future, rechargeable battery packs will be used in which the packs are plugged into AC outlets every night. This charge will hold well over the amount of life needed for a full day's work.

Direct sunlight. Unfortunately, no cure-all solution is available to control this environmental factor. The Newtons are currently set to the maximum amount of screen contrast before they are given to the FIs. Nothing further can be done to the hardware, but more emphasis will be placed on training interviewers to optimize their environmental surroundings to minimize any further complications due to lighting.

Flashcards. Complications were reported resulting from the flashcard "popping out" inadvertently. Alternative carrying cases were purchased for the 1999 NHSDA.

Correcting the roster after selections made. By obtaining a unique case-specific code from their field supervisor (FS), an FI can re-enter a completed screening and make corrections as needed. Some restrictions apply. Also, to help alleviate the need for changes, a roster verification screen was added before final selection could be made.

Correcting screeners performed on wrong line. As with the problem above, an FI must obtain a case-specific code from the FS from which he or she can then exchange information between two lines. To further ensure that an FI is entering data on the correct line, the line number and street address are displayed at the top of each screen.

Page scrolling and return to screen difficulties. Problems associated with page scrolling were corrected with special buttons that allow paging up or down without the use of the difficult Newton default controls. Also, currently when an FI returns to the Select Case screen, the list will be in the same place as when the FI left the screen.

Question test for roster questions is bothersome. Many FIs did not like the text of the roster questions after their first few screenings. They felt that they could recite the questions from memory and did not like having to tap "OK" to close the text box before entering the answer for each item. From a technical standpoint, it is easy to turn this presentation of questions on or off. However, RTI methodologists and field staff do not feel that it is advisable to allow FIs this option. They feel it will encourage them not to read the question text verbatim and to take "shortcuts." Upon further review, it was decided that an FI can submit a request to the FS that this option be turned off. Upon appropriate approval and demonstration of reliable and correct screening practices, via the FS's CMS, an FI's text question presentation can be turned off.

Gender/relationship/marital status mismatches. To help resolve these problems, in situations where it is obvious, gender and marital status will be coded automatically. For example, if the relationship is "wife," the person will automatically be coded female and married. If the relationship is "brother," the gender will be coded as male.

The Newton application was modified to add these additional functions. Case management activities were included in the application and the approach was modified so that the interviewers could directly transfer data from the Newton to the central office without going through the laptop computers. Records of calls and refusal reports were added as well.

In August 1998, the modified screening application was field tested. A total of 20 experienced field interviewers screened 836 households and listed 1,313 persons. Again, we closely monitored problems that were encountered. The most frequent Newton problems were problems with transmission to RTI (logged 36 times), problems with the screener (15 times), and problems with the CMS (6 times):

1. Generally, transmission problems were solved by resetting the Newton or the modem or by repeating the transmission (16 times); some were problems due to the way in which phone numbers were set up by the machine (11 times).

2. Newton screener problems included having the screen go blank or transmit slowly. These were generally solved by resetting the machine or removing/replacing the batteries. A couple of screener problems were due to the way in which phone numbers were set up by the machine.

3. About half of the Newton case management problems required intervention from either FS or technical support.

The interviewers chosen for the August 1998 field test were very skilled in using the paper-and-pencil procedures for the NHSDA. In contrast to the interviewers who participated in the 1997 field experiment, these interviewers felt that the Newton interfered somewhat with the screening. They complained that the Newton slowed down the screening process unduly, was difficult to read in bright sunlight and heat and humidity, disrupted eye contact with the respondent, and increased respondent impatience. They also had some problems with the complexity of the application. However, other comments were favorable to the technology. Interviewers noted that the Newton was light and more convenient to carry than the paper screener, that it was easier to pick up at the correct point on screening breakoffs and to get an overview of the pending interviews, that it was more convincing to the respondent that the selection was random, and that the reduction in paperwork associated with sending the data each night was welcome.

These comments and experiences were used to make final revisions in the application and training prior to fielding the application for the 1999 NHSDA.

8.4 Summary of 1999 NHSDA Electronic Screener Application

The 1999 NHSDA screening and electronic case management application contained the following features:

1. case management system,

2. household and group quarters rostering and sample selection application,

3. English and Spanish translations of the screenings,
4. collection of information that allows supervisors to verify selected final screening codes,
5. record of calls and refusal reports,
6. addition of dwelling units that were missed during the counting and listing, and
7. summary of information from the weekly production, time, and expense reports (PT&E)

The interviewer assignments were loaded onto the Newton, and the interviewers had the ability to sort the cases by status, such as all pending, pending screening, pending interview, final interview, and so on.

To screen a dwelling unit, the interviewer selected a case by tapping on a line in the list of dwelling units. An introductory screen appeared that guided the interviewer through the introduction and identification of an eligible screening respondent (someone who is 18 or older and who lives there). The interviewer then was presented with an address verification screen. Once the address was verified, the informed consent statement that was to be read to the screening respondent appeared. The application continued by displaying the basic questions on occupancy, number of residents, and questions age, relationship, gender, race, Hispanic status, and military status for all residents who were 12 years old or older. Once the roster was complete, the interviewer verified the listing and used an Edit Roster Record screen to make any necessary changes. The Make Selection button was then tapped, and the application made the random selection.

After the selection was made, a screen appeared that informed the interviewer of who, if anyone, was selected for the interview. The basic information identifying the persons selected (e.g., 18 year old, son) was displayed. This screen also displayed the questionnaire ID for each person. To begin an interview using the laptop computer, the interviewer entered this QuestID into the startup screen of the laptop computer. This QuestID provided the link between the questionnaire data, the screening data, and, thus, the sample selection probabilities. At the end of the interview, the interviewer entered a code into the Newton indicating that the interview was completed.

The CMS on the Newton was also used to monitor the visits that were necessary to conduct the interviews. A Record of Calls screen displayed the records of all prior calls to the dwelling unit. By pressing an Add button, the interviewer accessed an Add Call Record screen that allowed him or her to select a result code from a drop-down list and enter a comment. The date and time of the call were automatically recorded. If refusals were obtained, a Refusal Report screen appeared for use in coding the reason for the refusal and providing comments. Additional case management functions include the ability to edit addresses, special screening questions for group quarters, procedures for adding missed dwelling units, transmittal programs, and procedures for transferring cases from one interviewer to another.

The training for the 1999 survey on use of the electronic screening application was extensive. Interviewers received a computer manual with photographs of the parts of the Newton and explanations of its operation. The interviewer manual gave step-by-step instructions on the use of the Newton for case management and screening. During training, overhead projectors that could display the Newton screen were used which allowed the instructors to dynamically illustrate use of the Newton.

9. Operation of the 1997 Field Experiment

In this chapter, we discuss the following aspects of the operation of the 1997 field experiment:

1. the design and performance of the computer-assisted interviewing (CAI) programs, the hardware, and the supporting systems;

2. recruiting, hiring, and training of field staff; and

3. conduct of the field experiment.

Recall that the comparison group consisted of the Quarter 4 1997 NHSDA respondents in the primary sampling units (PSUs) where the field experiment was conducted. No special procedures were needed to get NHSDA substantive data from this comparison group. However, a subsample of these Quarter 4 NHSDA respondents was selected to receive an audio computer-assisted self-interviewing audio computer-assisted self-interviewing (ACASI) debriefing interview, and special interviewing procedures were needed for this subsample. Thus, this chapter discusses both the operations for the large field experiment and the comparison group debriefing subsample.

9.1 Design and Performance of the CAI Software, Hardware, and Supporting Systems

The 1997 field experiment interviewers used handheld computers for screening and laptop computers for case management, communications, and collection of interview data. Data were transmitted from the handheld to the laptop computer through a cable and from the laptop to the Research Triangle Institute (RTI) through a client-server transmission system using telephone connections. Processing took place at RTI in a networked PC environment, and operational data were made available to field supervisors (FSs) and survey managers through a protected website on the World Wide Web (WWW). Together, these computer systems allowed data to flow from the field to RTI electronically. We discussed the development and testing of the electronic screener in Chapter 8. In this chapter, we focus on the development and operation of the other software components.

The 1997 field experiment demonstrated that electronic data collection could be implemented successfully, but that care needs to be given to details to ensure good quality data. Overall, the systems worked well, although many needed improvements were identified and implemented for the 1999 NHSDA.

Software applications controlled or facilitated data movement at each stage of data collection. Programs for the CAI instrument, laptop case management system (CMS), transmission system, electronic screener, control system, and website were developed by RTI staff using commercial software packages.

The hardware for the 1997 field experiment consisted of a Toshiba Satellite 105CS laptop computer with a PCMCIA sound card and a PCMCIA modem, a set of headphones, and an Apple Newton MessagePad 2000. The Toshibas had 8 MB of RAM and a 75 MHz Pentium processor.

9.1.1 CAI Interview Software

The CAI interview for the field experiment had four components: a CAPI interview, an ACASI interview, an ACASI debriefing interview for the respondent, and a self-administered debriefing for the

interviewing staff. The two debriefing components were modified for use in the comparison group subsample.

Field experiment interview. The experimental design of the 1997 field experiment incorporated three factors with two levels each. As a result, eight distinct questionnaires had to be developed. Specifications were prepared for each questionnaire and were used by the programming staff to develop the CAI instrument. Once all question wordings were finalized, we began work on the associated audio files to be used in the ACASI portion of the interview. The eight treatments differed only for the questions included in the first eight sections of the ACASI portion of the instrument (alcohol through inhalants). Separate hard-copy specifications were prepared for these sections and included all question wordings, routing instructions, "fill" text, and definitions of new variables created as part of the logical flow of the instrument. Programmers used the electronic versions of these specifications to prepare the treatment portion of the CAI instrument. The remainder of the instrument was prepared by revising code used in the 1996 feasibility experiment.

The instrument was programmed using the Blaise 2.5 developed by the Netherlands Central Bureau of Statistics. Blaise 2.5 ran under the MS-DOS operating system, and due to questionnaire size limitations, the entire instrument could not fit into a single module. Instead, eight separate modules were developed:

1. FT2INTRO (CAPI): introduction, front-end demographics, and tutorial sections *(ACASI)*;

2. TREAT (ACASI): experimental treatment module containing tobacco, alcohol, marijuana, cocaine, crack, heroin, hallucinogens, and inhalants sections;

3. NTREAT1 (ACASI): analgesics, tranquilizers, stimulants, and sedatives sections;

4. NTREAT2 (ACASI): special drugs, risk, general drug, special topics, and treatment sections;

5. NTREAT3 (ACASI): workplace issues, drug experiences, youth experiences, and mental health sections;

6. FT2DEMOG (CAPI): back-end demographics;

7. FT2DEBRF (ACASI): respondent debriefing; and

8. FIDEBRF: field interviewer (FI) debriefing *(self-administered by the interviewer)*.

Once the FI began an interview, the CMS would then automatically write an MS-DOS command file for that case and execute it. This command file would run each of the modules in order and pass the case identification number as a parameter to each of the CAI modules.

Rather than have eight different versions of the experimental sections (TREAT) containing redundant question definitions, we programmed a single version with conditional branching based upon the experimental treatments. This approach made it easier to update the software and maintain comparability among treatments, and it also meant that one data output file was generated rather than eight.

All screens and their associated logic were tested before the audio files were incorporated. The hard-copy specifications were used during all testing. The testers followed each path through the instrument, ensuring that every response category for a particular question (including the "don't know" and "refusal" options) was being routed to the appropriate place. Programming staff made corrections, and the tester rechecked the program to verify that the corrections were made.

Within the Blaise definition of each ACASI question, the names of the audio files corresponding to the question text were identified. In some cases, such as with dates, ages, and responses to other questions, the audio file name was variable. The CAI instrument would determine the names of the audio files to be played based on previous question responses.

Once the text of every ACASI question and every response category was finalized, we identified the associated audio files. These audio files are formally designated as waveform audio files (or simply WAV files). In the simplest case, a question will require two associated WAV files: one for the question text and one for the response categories. However, in many cases, the text of a single question was split into multiple WAV files. It was necessary to split questions that include "fill" text, such as the 30-day reference date, into multiple WAV files to allow the "fill" text to change separately from the remainder of the question text. As an example, consider the following question:

> [Think specifically about the past 30 days — that is, from] [January] [5] [up to and including today.] [During the past 30 days, on how many days did you drink one or more drinks of alcoholic beverages?]

The audio for this question was split into five separate WAV files as indicated by the brackets. Similarly, a series of questions that all begin with a common stem were split into two pieces: one for the common stem and one for the text specific to that question. Splitting the text of a question into multiple WAV files reduced the total number of WAV files, thereby reducing the size of the ACASI program and enabling it to process more rapidly. There is also a downside to using multiple WAV files for a single question; namely, a pause may occur between the end of one file and the beginning of another, which can cause the question to sound choppy.

We contracted with an outside vendor for both the recording and the WAV file creation. RTI staff read the audio text from a script prepared by the programming staff. A member of the NHSDA project staff attended all recording sessions to catch any mistakes or poorly phrased readings so that a re-recording could be made on the spot.

We checked each WAV file for accuracy. The WAV files were loaded into a database along with the screen text associated with the file. Staff listened to each WAV file while comparing it to the on-screen text. Mistakes and omissions were corrected in a second round of recording, and this proved to be a valuable procedure for reducing the need to return to the recording studio. Next, we verified that the correct audio files were being played at the correct time in the interview. Testers followed every path through the questionnaire and listened to every question and every response category to make sure the audio matched what was on the screen. This detected any problems that could arise due to associating the wrong WAV file with a question. Inconsistencies were corrected in a last round of revisions.

Quarter 4 NHSDA debriefing interview. The debriefing interview for the respondents using the paper-and-pencil interview (PAPI)/self-administered questionnaire (SAQ) was designed to gather information about their experience in using the answer sheets and their opinions about using computers to answer such questions. The instrument contained

1. INTERVIEW ADMINISTRATION: seven items for the interviewer to record the screener roster number, record whether or not the respondent was willing to complete the interview, and end of the interview closing statements;

2. RESPONDENT TUTORIAL (ACASI): a tutorial that was nearly identical to the tutorial used in the field experiment sample; and

3. RESPONDENT DEBRIEFING (ACASI): questions that paralleled those used in the field experiment sample.

The questionnaire was also programmed using Blaise 2.5. After the interview was fully tested, the audio files were recorded and tested.

Performance of the CAI software. Although the CAI software performed reasonably well, some problems were encountered. Many of these involved situations where the laptop computer was turned off or lost power. In these cases, the CMS would sometimes lose track of where the respondent was when the power loss occurred. When the FI attempted to resume the interview, a "key already exists" error message would appear, indicating that the CAI module was being executed as if a new interview were being launched rather than a breakoff interview being resumed. When a low battery power message appeared, FIs would sometimes turn off their computers, and the CMS would lose track of the case.

Problems occurred if an inconsistent respondent age and birthdate data were entered during the initial demographics section. If the FI did not correctly resolve this inconsistency, the software would use the birthdate to compute the respondent's age. In some cases, sample members' ages were incorrectly computed as less than 12 years, and the CAI program was terminated inappropriately.

9.1.2 Case Management System

Design of the case management system. The CMS was developed in conjunction with the development of the Newton screening application (see Chapter 8) and the CAI questionnaire instrument. The CMS handled the following major functions:

1. case tracking,

2. launching and management of the CAI, and

3. interface with the Newton/FormLogic application and the RTI field communications system.

In the field experiment, the screening data that were collected on the Newton were transferred to the Toshiba laptops, and all data were transmitted to RTI by the CMS. The CMS was based on a set of rules and specifications for the recording of dates and events in the system. The system largely replaced the paper- and-pencil system that interviewers were using for tracking their activities associated with locating, screening, and interviewing sample housing units. Final status codes indicating a successful screening were generated by the screening application, and the CAI application generated the code and event for a final completed interview. Many other actions, such as visits to the housing unit, appointment for interview, pending refusals, and so on, were entered and coded by the interviewers. A set of final events and codes was also entered by the interviewer that reflected the last action taken at a sample dwelling unit, such as deciding the dwelling unit was vacant, that there was a language barrier, or the occupants refused to be screened or interviewed.

The user interface for entering events was a single screen design for case management that allowed the FI to move around without using the laptop's mouse. This design approach resulted in a dense and complex display that was somewhat hard to use. In addition, the FIs had to use the Alt-[letter] method of selecting actions for a very large proportion of allowed actions. The remainder are triggered in pop-up dialog boxes and require the interviewer to use push buttons to execute the actions.

The CMS was programmed in the Delphi 3.0 Windows development system. The CMS database system was MS Access 97. Programming took place in parallel with the CAI application development and the Newton screening application development. During the development phase, prototype versions were quickly created for fast-turnaround evaluation by survey managers and methodologists. This allowed the system to evolve and become more suitable to the task. Functional testing took place on an ongoing basis by the programmer and project staff. First stage usability testing also was conducted on a flow basis. Problems or suggested improvements were made quickly and returned to the tester for evaluation. Field testing of the system was confined to nine cases. Because of the dynamic nature of the development of the Newton application and the CAI application, programming of the CMS was rushed, causing some problems to emerge during the field experiment.

The CMS for the comparison group subsample allowed the interviewers to administer the interviews and then transmit the data back to RTI. This system was written in Visual Basic 5.0 and contained systems for calling the debriefing interview, status of all cases initiated on that laptop, and a system for transmitting data to RTI. All cases selected for the 1997 Quarter 4 NHSDA debriefing subsample were loaded onto the laptops so that the interviewer could initiate a case by entering the case identification number, which was verified by the program before the debriefing interview could begin.

Problems encountered with the case management system. The CMS guided the interview. Five problems were encountered with the CMS:

1. For the first few weeks of the 1997 field experiment, the CMS was using the same experimental treatment, specifically treatment 2, for the first CAI launched on each laptop. The initial treatment should have been selected randomly. A revised version of the CMS, which corrected this problem, was sent to the interviewers during the third week of data collection.

2. The CMS was very susceptible to corruption if the laptop was turned off without a normal shutdown during the Newton/Toshiba link and transmission to RTI. A total of 31 out of 195 FIs had to ship their Toshibas to RTI to have the databases repaired. Of the 232 Toshibas that were used on the field experiment, 27 were returned for repair (11.6%). This number includes mechanical failures and CMS problems. One particularly common problem related to the design of the Toshiba laptop and the CMS came to be known at the "power off" problem. The power button on the Toshiba happened to be located on the left side of the computer, precisely where one would put his or her hand to turn the machine away from the FI and toward the respondent for the ACASI portion of the interview. If this power button was accidently depressed while the ACASI interview was running, the CMS would irrevocably shut down. That is, restoring power would not rectify the situation. The entire system would be locked up and had to be returned to RTI for repair.

This frequent problem eventually was corrected via an update to the CMS.

3. Sometimes the records of contacts with the dwelling units were not in the correct order because FIs could enter a pending event with a date and time later than a final event in the Newton. After a Newton to Toshiba link, the final event would appear in the record of contacts, but because it was before the pending event, the case would remain in a pending status.

4. The FIs could circumvent some of the guidelines for entering information by using the mouse. They were instructed to use keystroke combinations but did not always do so, resulting in some difficulties determining the status of cases.

5. Some data and file problems were encountered during in-house data processing that arose from problems in the CMS, and modifications had to be made to both the in-house system and the CMS to adjust for these. The changes to the CMS were sent to the interviewers during the third week of data collection.

On the whole, we were able to recover from these difficulties; however, they clearly pointed out the necessity of making sure that there is sufficient time to program and test CMS.

9.1.3 Hardware

The hardware for the 1997 field experiment, as noted earlier, consisted of a Toshiba Satellite 105CS laptop computer with a PCMCIA sound card and a PCMCIA modem, a set of headphones, and an Apple Newton MessagePad 2000. Because the Toshibas only had 8 MB of RAM and 75 MHz Pentium processors, many of the interviewers reported the laptop systems to be slow in processing. Some of the hardware problems encountered during the 1997 field experiment included the following:

1. Toshiba Laptops: Problems with loose keyboard ribbons, batteries not holding charges, and ease of inadvertently turning off the power.

2. Newton Handhelds: Three Newtons had to be replaced. Two were dropped and one shorted out after being pluged into a faulty electrical outlet. More details concerning the operation of the handheld computers are given in Chapter 8.

9.1.4 Data Transmission and Capture Systems

To transmit data to RTI, an FI activated an automated data transfer process on his or her laptop. After entry of appropriate security information by the FI, data files were transmitted from the laptop to RTI, or vice versa, under the control of a client-server transmission system. Microsoft Visual Basic and SQL Server software, running under Windows NT or Windows 95, supported both transmission and capture. The transmission software system consisted of a central RTI database, which listed files to transfer and maintained activity logs, plus a laptop component called by the CMS. The capture process archived, expanded, and distributed files for use by other processes.

Transmission questions accounted for the second largest number of problems among interviewers, which included (a) difficulties calling from a hotel room, (b) uncertainties about the

successful transfer of cases from interviewer to interviewer, and (c) interrupted transmissions. Because of the number of transmission problems encountered, every time a laptop and Newton were returned to RTI, a transmission was done to try to capture any data that may not have been transmitted. Difficulties were encountered on approximately 1% of attempts.

Again, these problems demonstrated to us the importance of having very robust computerized field management systems.

RTI provided telephone support for the FIs during the course of the 1997 field experiment. Almost all of the FIs had a reason to contact RTI's computer support staff on at least one occasion. There were 150 FIs at work for 91 days and, on the average, problems were encountered on 5.4% of the total workdays. Interviewers were able to contact an RTI technical support person by using RTI's toll-free number during normal business hours (477 calls), calling an automated response system after hours (32 calls), or calling the "after-hours" emergency pager number (232 calls). **Exhibit 9.1.1** summarizes the calls that were received.

9.1.5 Monitoring and In-House Processing of Data

Electronic data collection shifts some aspects of data processing from the central site to the field. First, data entry is done by the interviewer or respondent and the manual pre-edit step at RTI is eliminated. Second, the cycle time is reduced in terms of receipt of data from the field, which means that central processes must respond quickly to the presence of new data to make the information available. Thus, there are new sources of error in the paperless environment, and new consistency checks must be defined so that inconsistencies and missing components can be noted. Additional errors may arise from equipment misuse or failure, and problems may be at any or all of several levels: the data items themselves, the files in which they are stored, or the software that manipulates them.

A PC-based control system, which used Microsoft Access and Visual Basic, was developed to be compatible with the CMS data files received from the field systems. Due to time constraints in implementation and the need for defining new methods of handling data and trapping errors, this system offered limited functionality as compared to the NHSDA national control system. However, they did provide valuable experience that was subsequently used to design the procedures for the 1999 NHSDA.

One gratifying result of electronic data transmission was the rapid availability of data. For example, this allowed us to detect the problem with the random assignment of respondents to treatments within the first 2 weeks of data collection. However, with expedited transmission of data from the field came a need for rapid processing. Control system operations were developed and scheduled to run periodically throughout the day and night. Thus, incoming data were added to the central database as rapidly as possible. Once in the database, the information was made available through a restricted-access site on the web. Because of this, reports and website database updates were available every morning instead of weekly. FSs and project staff members appreciated the ease of use and accessibility of the website.

Reports were viewable by authorized NHSDA staff through the use of standard browser software, either within RTI or from the field, on a 24-hour basis. FSs and RTI staff expressed frustration with the limitations of both sets of reports, and these were improved and supplemented for future use.

Based on the results, the electronic control systems were changed to include four features:

1. increased level of error checking in field systems and at the data-capture stage,

2. expanded functionality for the FS website management system,

3. expanded and improved reporting from the control system, and

4. increased data-review and correction systems in-house.

9.1.6 Problems with the 1997 Quarter 4 NHSDA Debriefing Systems

Several similar problems were noted for this component of the survey; however, because the component was much smaller and simpler than the field experiment application, the scale of problems was greatly reduced. As noted in the next section, the interviewers who participated in the debriefing did not attend training but learned how to use the laptop through home study. We were interested in how well they performed in the household and asked the respondents about laptop problems using the question shown in **Exhibit 9.1.2**. The exhibit shows that about 98% of the interviews were conducted without much trouble.

Exhibit 9.1.1 Technical Support Provided to the FIs During the 1997 Field Experiment

Type of Problem	Number of Problems Reported	Number of FIs Reporting Problem
CMS	242 (33%)	99 (26%)
Transmission	154 (21%)	74 (19%)
Screener/Newton	127 (20%)	77 (20%)
User Error	113 (15%)	61 (16%)
ACASI/Interview	33 (4%)	20 (5%)
Newton/Toshiba Link	30 (4%)	23 (6%)
Hardware	24 (3%)	16 (4%)
Miscellaneous	18 (2%)	17 (4%)
Total	741 (100%)	387 (100%)

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 9.1.2 Question Used to Evaluate the FI's Use of the Equipment

Question: How much trouble did the interviewer seem to have setting up the computer to begin this portion of the interview?

Response (n = 593)	Frequency	Percent
No trouble at all	522	87.9
Some trouble	59	9.9
A lot of trouble	12	2
Don't know	1	0.2

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

9.2 Recruiting, Hiring, and Training of Field Staff

To collect the data, RTI employed 144 FIs reporting to 10 FSs. The effort was managed by two in-house regional supervisors (RSs). To maintain consistent leadership with the 1997 NHSDA main study, the 1997 field experiment used many of the same in-house staff. Recruitment and hiring of FSs and FIs began in April 1997 and extended through August 1997.

All 10 of the FSs hired for the field experiment were recruited from RTI's database of field personnel. Each FS had served in at least one field supervisory role with RTI in the past; four had prior experience as an FS on the NHSDA. The territorial assignments were made based on PSU size, geography, and historical territorial assignments, and the assignments were divided relatively equally among the 10 FSs. The number of PSUs assigned to each FS ranged from 5 to 14, depending upon the specific characteristics of the assigned PSUs. The number of FIs working under each FS ranged from 12 to 17, again depending upon the specific makeup of the territory.

The FSs attended two training sessions to prepare them for their supervisory roles: one in late June and the second in mid-September. These sessions were conducted by RSs from both the 1997 field experiment and from the 1997 NHSDA main study. The first training session lasted 6 days and covered recruiting procedures, an introduction to the NHSDA project, including general background information, locating and contacting the sample dwelling units, and a review of the screening procedures. The final segment of FS training consisted of 2 days of training on the use of the laptop computers. A portion of the training overlapped days 1 through 3 of a 1997 NHSDA main study's FI training session in order to save costs.

The second FS training session was held for 2 days in 1997 and focused on management and monitoring.

Following the FS introductory training session, the FSs began recruiting. Their efforts focused on FI candidates who had computer experience or who were deemed likely to succeed when working in a CAI environment. Sources for FI candidates included RTI's National Interviewer File, contacts with other survey organizations, and local government employment agencies. For sample areas in which FSs had difficulty recruiting qualified candidates, advertisements were placed in local newspapers.

9.2.1 Field Interviewer Training for the 1997 Field Experiment

Training included a "Training of the Trainers" session and an FI session. A 3-day session to instruct those in-house and field management staff who would serve as trainers for the FI sessions was held on the RTI campus in September 1997. The attendees included three lead trainers, 30 staff who would subsequently serve as FI trainers, technical support staff, administrative support staff, other NHSDA project staff, and a representative from the Substance Abuse and Mental Health Services Administration (SAMHSA).

Prior to data collection activities, two simultaneous FI training sessions were held to train those interviewers retained for the 1997 field experiment on the procedures for conducting the study. One session was held in Research Triangle Park (RTP), North Carolina; the other was in Los Angeles, California. Each prospective FI hired for the 1997 field experiment was required to complete a homestudy exercise package prior to attending the in-person training session. The homestudy package included a cover memo that instructed the recipients to read the field experiment manual, answer the homestudy review questions, watch the videotape, and listen to the audiotape. Those trainees who were new to field interviewing also were instructed to read the manual on fundamentals. The completed homestudy review questions were to be turned in upon registration at the assigned training site.

Due to the complexity of the 1997 field experiment's training program, a training team consisting of a number of positions was assigned to each training site. At each site, there was a site leader, an administrative support person, a lead technical support, 4 technical assistants, 5 lead trainers, and 10 assistant trainers. A training team consisted of a lead trainer, 2 assistant trainers, and a technical support person who trained from 12 to 17 field interviewers.

During the training, interviewers were divided into groups based on their experience. Special sessions were held for those who had never done personal visit interviews; for some sessions, we formed groups with different degrees of computer experience. As a supplement to the daily training sessions, a 2-hour study session was scheduled for each evening. These sessions allowed trainees to receive further instruction.

There were some difficulties. The Newton did not have video-display capabilities that allowed active screens to be projected onto a larger screen for viewing by the group. To deal with this issue, RTI's photographers prepared images of each screen of the prescripted Newton training exercises. These were loaded into a PowerPoint presentation that resided on the trainer's Toshiba and were projected using a device designed for computer presentations. This, however, did not allow the trainer to spontaneously manipulate Newton screens or exercises to address specific questions raised by the class. For the 1999 training, projectors that could display the Newton screen were used.

There were also some problems due to the fragile state of the CMS. Some procedures were awkward, and some software bugs were discovered. Sheets summarizing the commands were developed and distributed to all trainees, and this was well received by the trainees.

A key requirement of the 1997 field experiment interviewers was that data be transmitted to RTI every night. Practicing this was important, but at each training hotel the telephone line requirements were not met. Site leaders and technical support staff worked closely with the hotel technicians to resolve this problem.

Trainers carefully monitored all attendees on overall performance and comprehension. Weak performers were given extra attention, primarily during the evening study sessions. Through these efforts, 71 out of 72 interviewers "graduated" from the Los Angeles session; 70 out of 72 interviewers graduated from the RTP session. Overall, 10 of those graduating were classified as needing extra attention from their supervisors once they began their field assignments. By the end of October, a total of 14 FIs had left or had been released from the project. Thus, a supplemental FI training session was held from November 6 through November 11, 1997, to boost the number of FIs working on the 1997 field experiment. A total of 23 FIs were trained at this session, which was held on campus at RTI. Ten (10) of the trainees were brought in as replacement FIs; the remaining 13 trainees were designated as traveling FIs who would travel as needed to boost production in particular areas. A third and final supplemental training session was held in late November 1997. During this session, four experienced NHSDA interviewers were trained to serve as additional traveling FIs. Due to their experience with the study, the training session was reduced from 6 to 3½ days and focused on the Newton and Toshiba administration components of the 1997 field experiment.

9.2.2 Field Interviewer Training for the 1997 Quarter 4 NHSDA Debriefing

A subset of the 1997 NHSDA FIs were trained to administer the ACASI debriefing interview to the subsample of PAPI respondents. A total of 71 interviewers who were working on the 1997 NHSDA main study were selected based upon their performance on the NHSDA in general, their perceived ability to conduct the ACASI debriefing portion of the interview, and their location relative to the segments selected for the debriefing sample. Field supervision of the debriefing sample was subsumed under the regular supervisory duties of the assigned FS for the 1997 NHSDA; most of the 12 supervisors had one or more interviewers working on the debriefing component. The target number of debriefing interviews to be completed was 750.

All interviewers selected for the debriefing component of the 1997 field experiment were already trained on the NHSDA PAPI administration procedures. Because the ACASI debriefing program was simple, an in-person training session for the debriefing interviewers was not necessary. Instead, they were trained via a comprehensive homestudy package that contained a booklet of study materials that covered procedures for homestudy, an overview of the debriefing study, hardware and software, a step-by-step guide through the actual debriefing interview, and a technical support information sheet. In addition, they were sent a videotape that was used in conjunction with the home study booklet. The video showed detailed, step-by-step instructions on using the Toshiba laptop computer and administering the ACASI debriefing interview. It also provided instructions on transmitting/telecommunicating completed debriefing data to RTI. Upon completing the self-study portion of the homestudy package, the interviewer was required to successfully complete a certification and transmission exercise. Receipt of the debriefing assignment was contingent upon successful completion of the certification exercise.

A laptop carrying case containing a Toshiba laptop computer and all necessary ancillary materials were also shipped to the interviewers along with packing materials and a Federal Express label for return shipment.

9.3 Data Collection

As noted above, the 1997 field experiment used many of the same in-house management staff as the 1997 NHSDA main study. The overall management approach was based on the procedures already in place for the main study; however, modifications were necessary to accommodate the differences primarily related to electronic data collection.

9.3.1 Management of Data Collection Operations

Data collection for the field experiment began on October 1, 1997, and continued through January 31, 1998. The final response rates did not reach the targeted level. Only 1,982 interviews were completed, which is 11.9% lower than the expected number of 2,250. In this section, we discuss how data were transmitted from the field, communications between field staff, the use of traveling FIs, and other field conditions. We highlight the situations that contributed to the shortfall in the number of interviews.

During training, interviewers were instructed to update all events related to their fieldwork each day and to transmit their data everyday. These daily transmissions provided the FSs and project management with frequent and consistent reports of progress made in the field. Each day, reports were generated and posted to the password-protected NHSDA website. Thus, more timely feedback could be provided to all staff. This procedure was a significant improvement in timeliness over the weekly reporting schedule used on the NHSDA main study. However, initial system problems early in the field period resulted in some inefficiencies in compiling the information necessary to manage the field staff efficiently and, in some cases, effectively. The system problems were continually addressed throughout the early weeks of data collection in an effort to stabilize and improve the automated reporting system.

FSs held a mandatory weekly conference call with each of their assigned FIs. Each call lasted a minimum of 1 hour and allowed the supervisor to work with each interviewer to monitor and improve the quality of work, rate of production, and costs associated with his or her field efforts. In addition, any other relevant topics, questions, concerns, or problems were discussed during this time. Supervisors also used the conference call to relay important information from the field management staff to the interviewers. For example, known computer problems and their corresponding solutions (if available) were relayed to the FIs during the weekly call.

Initially, we intended to use the daily and weekly production reports that were to be posted to the website to review work status with the interviewers. However, due to uncertainties in the system during the early weeks of the 1997 field experiment, a backup monitoring and management plan was put in place, and during the weekly call, the FIs reported their updated status codes for each case in their assignment. The FS used this information to discuss difficult cases with the interviewer, and it also provided a check on the information that was received during the transmission. At the beginning of the study, the FSs' hand counts most often were more accurate than the web-based reports.

The in-house regional supervisors (RSs) called their FSs each week to obtain their counts of completed and pending cases, and this became a benchmark that was used to monitor the revision of the web- based reporting system. The RS/FS calls also included discussions about FI workload and any other field production problems. In addition, the RSs conducted periodic group teleconferences with the supervisory staff to share information on how to improve respondent participation rates.

Once data collection started, FI attrition was higher than had been anticipated. At the same time, a higher-than-expected yield from the field experiment sample resulted in more cases being eligible for the study. As a result, very early in the field period it became evident that the interviewer staffing level was insufficient to complete the work as planned. To help combat this problem, a select group of specially skilled and trained FIs were identified and secured to serve as traveling FIs, or TFIs. The TFIs did not have a home-based assignment; rather, they were available for field deployment where needed. Thus, additional interviewers were trained as described above.

The 1997 field experiment was the first major NHSDA data collection effort for which all data—screening and interviewing—were collected and transmitted via computer. Prior to the start of data collection, considerable effort was expended to test the screening and ACASI components. However, a somewhat compressed development period for the CMS resulted in a product that needed a more complete and thorough testing. We identified some problems very early in the field period that required immediate corrective actions. After about 3 weeks of data collection, all FIs were instructed to stop work until they received a new version of the CMS; this was shipped out on a diskette. The need to upgrade to a new version of the CMS resulted in several days of lost field time. Although the new version of the CMS was much improved, problems still existed with the updating and posting of case-level status codes, and the CMS continued to receive a great deal of attention from the RTI technical staff. Additional problems were identified and, when possible, corrected electronically when the FIs transmitted data to RTI. For those problems that could not be readily resolved via software corrections, FIs were notified by their FSs of the known problems and of steps they could take should the problem happen to them. Again, as was noted earlier, backup hard-copy reporting systems also were implemented to provide a cross-reference on the status of work in the field.

The FSs had a version of the CMS that was a modified version of the FI system. Thus, because the FI system experienced problems early on, so too did the FS CMS. In fact, completion of the FS CMS system was delayed until the initial problems identified in the FI version were corrected. By mid-November 1997, a preliminary version of the FS CMS was operating as a web-based reporting system.

Essentially all of the problems experienced with the CMS were a result of a lack of sufficient time to thoroughly test and revise the systems prior to actual data collection. The data collection period became the testing ground, which is a less-than-satisfactory situation. The consequences included delays in field progress and resultant frustration among the field, management, and technical staff. After the field test was completed, the CMS was redesigned, enhanced, and rigorously tested to develop the version that would serve as the foundation for systems to be used in the 1999 NHSDA. In April 1998, a 2-week field test of the prototype web-based CMS was conducted, and the system performed extremely well, and some additional refinement and testing was done prior to the beginning of the 1999 survey.

Because 1997 field experiment was the first time in which screening data were collected via computer, it was unknown what impact, if any, that method of data capture would have on the screening results. Thus, the sample was designed with the intent to accommodate potential shortfalls due to failure of the screening hardware or software. However, these possible problems were not realized, and the yield was 14% higher than expected. The size of the FI staff was insufficient to handle this increased workload. To resolve this problem, at mid-quarter 1, 837 cases were removed from the sample. This required a redistribution of assignments for the interviewers and some inefficiencies in the data collection.

For the 1997 field experiment, the sampling rates were set so that 50% of the respondents were expected to be 12 to 17 years old. Because we used similar sampling algorithms as were used in the national NHSDA, this resulted in a higher than usual number of households in which more than one person was selected. Each interview takes about an hour, and the response rates are depressed in a two-interview household because of the increased burden.

The electronic screening mechanism had many advantages, particularly related to the accuracy of implementing the respondent selection routine. However, this initial screening application was relatively rigid; once an error was made by an FI, it was difficult or impossible to correct that error, particularly when standing on the respondent's doorstep. As these types of errors became known, all staff were cautioned to be particularly careful, and errors were reduced but not eliminated.

Another issue that arose was the interviewers' perception of the respondents' attitude toward the overall interview length. Although the total length of the CAI and the PAPI version used in the national study is nearly the same, many interviewers seemed to feel that respondents thought the process was excessively long. This perception apparently arose as a result of the time needed to set up and then break down the electronic equipment. Some FIs may have been less inclined to push for the interview because they felt it was excessively burdensome for the respondents. As the field period progressed, the FIs became more proficient with setting up and breaking down the equipment, and this concern subsided.

Finally, the use of the TFIs was not as effective as anticipated. Some high-quality, veteran NHSDA interviewers were selected for this assignment; however, they had not worked as traveling interviewers before and did not deliver the level of effort on the 1997 field experiment to which they had committed.

As was noted earlier, the time frame to prepare for the 1997 field experiment was relatively compressed. The final decision to introduce electronic screening for the field experiment did not occur until June 1997. Again, this limited our ability to adequately test the complete electronic system, especially as it involved interactions with the CMS.

The data collection period spanned the winter months of November and December 1997 and January 1998. This is traditionally the most difficult time for fieldwork on the NHSDA. In spite of

efforts to counteract the effects, uncontrollable factors, such as bad weather, holidays, and fewer daylight hours, had a negative impact on response rates.

In spite of the problems that were experienced, much of the overall reaction to the computerized approach was, in fact, favorable. The tremendous potential for a fully automated NHSDA data collection effort was obvious, as is evidenced in the sections that follow.

9.3.2 Field Staff Reactions to Computer-Assisted Interviewing Procedures

At the conclusion of the field experiment, we obtained input from the field staff regarding their perception of the process as a whole. Four conference calls were held in January and February 1998 to debrief the field staff. Two calls were held with the FSs, and two were held with selected FIs. RTI project management staff moderated and listened in on the calls, and SAMHSA staff participated. Prior to each call, a memo containing general instructions and a list of specific issues was distributed. Each call lasted approximately 2 hours.

Field supervisor reactions. In the discussion of training, several FSs noted that they felt inadequately prepared at the FI training session for dealing with the computers and the computer problems that occurred. They suggested that detailed explanations of potential computer problems prior to training would have been helpful (although in reality such details were not known prior to the training session). They also noted that training might be more effective if FIs with little to no computer experience were trained separately from those who had experience. They noted that the the amount of equipment/hardware involved with the field experiment was difficult for some FIs to carry.

The FSs also noted that they would have preferred to have more software and hardware capabilities available at the onset of data collection, including (a) the ability to electronically assign and transfer cases between FIs, (b) the ability to communicate via e-mail with the FIs, (c) prior training on handling the more common computer/technical problems, and (d) a more functional system of automated and accurate field status reports.

A functional electronic system of gathering and reporting data alleviates many of the mundane paperwork tasks typically completed by the FSs. However, the field experiment FSs noted that they regretted having no way to see the quality of their FIs' work. That is, with the paper NHSDA, the initial completed cases from the FIs are sent to the FS for a quality review. In the field experiment, there was no such ability to monitor the actual data being entered by the FI.

The FSs reported that they were more comfortable with hiring FIs who had computer experience. They also noted that it generally was not difficult to find candidates meeting this requirement. Some did feel that it was sometimes difficult to recruit people for the NHSDA due to the subject of the survey and the lack of respondent incentives.

Overall, FSs were generally positive. They understood that this was a test of procedures, and thus problems were to be expected. The FSs made a point to emphasize that the technical support available to the field staff was excellent, especially after the initial rush of calls during the first few weeks of data collection. Moreover, they were very positive about using the Newton for screening.

Field interviewer reactions. For the FI debriefing calls, the RSs chose a cross-section of staff from their region. Overall, the FIs also were positive. They felt that potential respondents were very receptive to the Newton and to completing the interview on the computer. The Newton generally performed well, even in adverse weather conditions. However, at times it was difficult to see the screen in

direct sunlight. Some FIs complained that the Newton ran slower or the screen fogged up in cold or rainy conditions.

Problems with the CMS constituted the largest concern expressed by the FIs. They requested simpler and more consistent procedures for the software access, management, and reporting tasks. The FIs noted that a detailed "cheat sheet" pertaining to the CMS was provided at the conclusion of training and was helpful in terms of providing a source for easy reference. Several complained about the limited battery life they experienced, especially in the Newton. Other difficulties were mentioned:

1. holding all of the required materials and equipment while at the door conducting screening;

2. a perception that the interview took much longer than 1 hour to complete;

3. the lack of an incentive for respondents;

4. difficulty in completing the second interview in a two-interview household due to the length of the first interview;

5. repetitive questions in the interview (this varied by questionnaire version); and

6. the ACASI tutorial, which, for many respondents, was seen as being a waste of time.

The FIs were asked what they did to occupy themselves while the respondent completed the ACASI portion of the interview. They reported that they completed paperwork, prepared materials, socialized with other members of the household, read personal materials, read their FI manual, or did some combination of these activities. Some FIs reported that most of their respondents used the headphones during the interview; others reported that they did not. This action seemed to depend on whether the FI encouraged the use of the headphones.

When asked about training, the FIs generally were positive about their experience. Many said they would have liked to have had more information on troubleshooting potential problems. However, they acknowledged that such information was not readily available until after problems had occurred in the field. Others suggested that there should be more time at training to practice with the computers. Similar to what was expressed by the FSs, many FIs also felt that they should have been separated by their degree of computer experience. Also similar to the FSs' response, most felt that technical support during training and while in the field was excellent. A few FIs complained about isolated problems with reaching technical support staff and with receiving a prompt response to a specific issue.

9.3.3 Management of Debriefing Sample Data Collection Operations

By design, the debriefing sample cases were assigned to field staff working on the 1997 NHSDA main study. Consequently, the FIs, FSs, and RSs taking on this additional assignment were all veteran NHSDA staff. The same 1997 NHSDA management team at RTI was in place to oversee the debriefing sample.

In those segments selected to be part of the debriefing sample, one or two people in the sample dwelling units may have been selected for the NHSDA interview. A hard-copy screening form indicated whether no, one, or two debriefing interviews were to be conducted. At the conclusion of each day's work

where they had completed one or more ACASI debriefing interviews, the FIs were to transmit the debriefing data directly to RTI.

From the FSs' standpoint, management of the debriefing sample was somewhat difficult. The FSs were not be able to determine with certainty whether a debriefing interview was to be conducted at a given dwelling unit until they received the hard-copy screening form from the FI because the information was resident on the screening forms. Thus, during their weekly conference calls, for every completed interview reported, the FS would ask if an ACASI debriefing interview had been conducted. Because the FSs had no records with which to cross-check the information, they had to rely on the FI to properly relay the debriefing information to them. Upon receiving the hard-copy information, the FS

1. reviewed the screening and selection procedures,

2. confirmed the debriefing interview assignment status debriefing interview for each eligible case, and

3. checked the debriefing sample transmission report found on the FS website and confirmed that a transmission had occurred.

Depending on the outcome of the review, one of two steps was taken:

1. If all information was correct and the ACASI interview had been transmitted, the case was considered to be complete and the FS noted this in the tracking system.

2. If a transmission did not occur, the FS contacted the FI to verify an ACASI debriefing interview had been completed. If "yes," the FI was instructed to do a manual follow-up transmission. If "no," information regarding the lack of a debriefing interview was entered into the tracking system.

Several types of problems were encountered, including missing interviews, unsuccessful transmission of cases, cases coded as completed for which the respondent had actually refused the interview so that no interview should have been sent, and extra interviews that appeared because an ACASI interview was inadvertently completed at a dwelling unit where none should have been.

In spite of the effort to make the homestudy as complete and thorough as possible, the preparations proved to be insufficient for some of FIs. A higher-than-average number of errors were associated with FIs failing to administer debriefing interviews to eligible respondents and erroneously administering the interview to ineligible respondents, as well as excessive technical difficulties related to operating the hardware/system and transmitting completed data.

In addition, the sampling rates embedded in the screeners did not yield the expected 938 cases, of which 750 would be completed. In fact, only 738 cases were yielded, which is 20% less than expected. Approximately 596 ACASI debriefing interviews were completed, 154 short of the projected target of 750. This shortfall can be explained in part by the aforementioned difficulties; however, we also felt that the large number of resources required to simultaneously manage the field experiment, the main NHSDA, and the debriefing component contributed to less than adequate attention to this latter component.

10. Willingness to Respond

In Chapter 5, it was noted that the field experiment response rates were low compared with those of the comparison group from Quarter 4 of the 1997 NHSDA. In this chapter, we examine these response rate differences in more depth, and we seek reasons for the lower response rates experienced during the 1997 field experiment.

10.1 Response Rates

The field experiment's response rates were considerably lower than the Quarter 4 NHSDA rates. **Exhibit 5.4.1** shows that the field experiment's screening response rate was 85.4% compared with 93.6% for the comparison group. **Exhibit 5.4.2** shows that the field experiment's person response rate of 62.7% was 12.9% lower than the overall person-level response rate attained in the comparison group (75.6%). Results of the field experiment for the electronic survey instruments should be evaluated by comparing them with those obtained from Quarter 4 of 1997 NHSDA using the standard paper-and-pencil interviewing (PAPI) instrumentation. However, the two surveys used separate field supervisor (FS) and field interviewer (FI) staffs, confounding the comparisons of electronic versus PAPI instrumentation differences with those due to the differences in the two interviewing staffs. The survey response rates also may have been affected by the attitudes of the interviewers and their supervisors toward the 1997 field experiment survey, as well as the attitudes of the respondents themselves toward the electronic instruments.

10.1.1 Dwelling Unit Screening

We examined the screening response rates by region, metropolitan statistical area (MSA) size, and certainty/noncertainty area. These details are shown in **Exhibit 10.1.1**. The dwelling unit (DU) ineligibility rates for the field experiment were quite similar to those for the 1997 NHSDA Quarter 4 comparison group. The same was largely the case for the ineligibility rates across the subgroups examined. A few differences in the subgroup ineligibility rates were observed, with somewhat higher vacant DU rates during the field experiment in the Northeast region and in large MSAs. Overall, however, the ineligibility rates for the field experiment and the comparison group sample were remarkably similar.

There were major differences between the two surveys in the percentage of eligible DUs that were successfully screened. The field experiment screening rate of 85.4% was 8.2 percentage points below the 93.6% rate for the comparison group. **Exhibit 10.1.2** shows that nearly all of this large shortfall was associated with much higher occurrences of field experiment screening refusals (i.e., units in which the interviewers were denied access) and the other nonresponse category. Each of these three nonresponse rates were at least twice as high for the field experiment as they were for the comparison group. The two rates of nonresponse for not at homes were nearly identical. Units not screened because of Newton screener problems, a special field experiment nonresponse code, accounted for less than 1% of the eligible DUs.

Examination of the detailed data in **Exhibit 10.1.1** reveals some contrasting patterns in the DU nonresponse subcategories. The high refusal rates in the field experiment occurred across all of the subgroups. The nonscreens due to denied access, however, occurred predominantly in the West region (8.2%), large MSAs (5.5%), and certainty primary sampling units (PSUs) (5.6%). The field experiment's denied access rates for those subgroups were many times larger than the rates for all other subgroups. Interestingly, the denied access rates for the 1997 NHSDA Quarter 4 comparison study were very low for all of these subgroups (generally 1% or less). Denied access is a highly variable reason for nonresponse

and includes gated communities and apartment buildings that do not allow interviewers to directly visit the sample house or apartment. When the survey team has been working in an area for a considerable amount of time, they develop skills and contacts that allow them to get past the gatekeepers. Given that the interviewing staff who were working in the field experiment were new to the NHSDA and worked for only a short period of time, it is not surprising that this rate was higher than it was in the comparison group.

Still another pattern of subgroup nonresponse was present in the "other nonresponse" category. The field experiment's other nonresponse rates were considerably higher than the comparison sample rates, both overall and for subgroups. The majority of the excess other nonresponse in the field experiment was due to the group quarters. The screener application did not include the group quarters screening algorithm, and these types of DUs were coded as "other" when assigning an outcome code during the screening. Overall, about 1.4% of the DUs were group quarters. Removing these DUs from the other nonresponse category results in similar rates for the field experiment and the comparison group. Reflective of an uneven distribution of group quarters, there was variability across areas: the Midwest (5.1%), the South (3.3%), non-MSA areas (7.5%), and noncertainty PSUs (4.2%) in this other nonresponse.

10.1.2 Person-Level Nonresponse

The field experiment's overall person response rate was 62.7%, which was 12.9 percentage points lower than the 75.6% response-obtained rate from the 1997 NHSDA Quarter 4 survey. The detailed person nonresponse rate information for all sample persons, from the field experiment and the comparison survey is shown in **Exhibit 10.1.3**. There was no Spanish version of the electronic questionnaire for the field experiment. Thus, for comparability of data in this report, cases that were completed in Spanish in the NHSDA comparison group were coded as language barrier nonresponses in these tabulations. We must keep this in mind when considering the field experiment's person response rates.

The lower person response rate obtained from the field experiment was primarily due to the field experiment's 19.8% person refusal rate in contrast to the 8.6% person refusal rate for the comparison survey. Although the refusal rate accounted for the largest difference in the person response rates, the field experiment also encountered somewhat higher nonresponse for physical/mental disabled and for other nonresponse reasons (see **Exhibit 10.1.4**). These were partially offset by somewhat lower field experiment nonresponse rates for the no-one-home and language-barrier categories. These differences were considerably less dramatic than the refusal rate difference described above.

The response rates for the field experiment and comparison survey exhibited many similar patterns of nonresponse across various subpopulations examined. In both surveys, the following statements can be made concerning the pattern of person-level response rates:

1. Response was negatively correlated with respondent age.

2. Response was lower for Hispanics than for blacks or whites and others.

3. Female response rate was slightly higher than male response rate.

4. Response was slightly higher when only one person was selected in the DU.

5. Response was highest for non-MSA areas and lowest for large MSA areas.

6. Response was higher in noncertainty PSUs than in certainty PSUs.

Similarly, the nonresponse categories for the field experiment and comparison sample exhibited similar patterns of nonresponse across subgroups. For example, **Exhibit 10.1.5** shows how the percentage of eligible persons refusing to be interviewed tended to vary directly with the person's age. A similar pattern was evident, with a major difference in level.

Exhibit 10.1.6 shows an increasing rate of language barrier cases, by age groups, for both the field experiment and comparison surveys. The levels were similar for both surveys, as well as the pattern of nonresponse.

To summarize, the person-level response rate was 11.0 percentage points lower in the field experiment compared with the rate for the 1997 NHSDA Quarter 4 sample. The primary reason for the difference was the higher refusal rate experienced in the field experiment. Despite the lower level of response during the field experiment, the two surveys showed many similar response and nonresponse patterns. That is, except for the significantly higher field experiment refusal rate, the two surveys showed many similar patterns across both subgroups and categories of nonresponse.

Exhibits 10.1.7 and **10.1.8** contain the detailed person response rate information for 12 to 17 year olds and for adults 18 years of age or older. There are some interesting differences between the response rates and nonresponse patterns for these two age groups. Again, the refusal rates were much higher for the field experiment than for the comparison sample. The summary information in **Exhibit 10.1.9** contrasts the results for 12 to 17 year olds with those for persons 18 years of age or older. Clearly, the lower field experiment response rates were primarily due to the field experiment's higher refusal rates for both age groups. The field experiment's nonresponse rates for both age groups also were higher than those for the comparison survey for the physical or mental problems and other nonresponse categories. The field experiment and comparison surveys both yielded approximately the same percentages of persons who could never be found at home, and this held for both age groups.

The language barrier information in **Exhibit 10.1.9** is quite interesting. The same percentage of nonresponse (10.5%) was obtained in both surveys for adults, whereas for 12- to 17-year-old persons, the field experiment found fewer language barrier cases (1.2%) than did the comparison sample (4.5%). This is likely due to the coding of 1997 NHSDA Quarter 4 persons who responded with the Spanish PAPI instrument as language barriers for the purpose of this analysis. It appears that many of the 12- to 17-year-old sample members, who might have chosen to use the Spanish instrument in the comparison sample, responded instead in the field experiment to the English CAPI instrument because they were capable of responding in either language, and English was their only choice in the field experiment.

The detailed field experiment and comparison survey response rate tabulations for the two age groups show many of the same patterns of nonresponse across the subpopulations tabulated. Both age groups in both surveys exhibited the following person-level response patterns:

1. Response was lower for Hispanics than for blacks or whites and others.

2. Female response rate was slightly higher than male response rate.

3. Response was slightly higher when only one person was selected in the DU.

4. Response was highest for non-MSA areas and lowest for large MSA areas.

5. Response was higher in noncertainty PSUs than in certainty PSUs.

To a considerable extent, several of these patterns result from the lack of Spanish-language CAPI instrumentation. This drastically lowered the computed person response rates for both age groups for Hispanics, most of whom lived in the West region, in large MSAs, and certainty PSUs.

Exhibit 10.1.10 shows for youths and adults the differences between the field experiment and 1997 NHSDA Quarter 4 response rates. This comparison of overall and subgroup response rates shows that both surveys obtained higher response rates for 12 to 17 year olds than for adults. The response rate for the younger age group was higher for both surveys overall and for each of the 16 subgroups examined.

The exhibit also shows that the negative differences between the field experiment and comparison group response rates were much larger for persons 18 years of age or older than they were for persons 12 to 17 years of age. This pattern held for the total samples, as well as for each of the 16 subgroups.

Exhibit 10.1.1 Dwelling Unit Screening Rate Percentages, by Region, MSA Size, and Certainty/Noncertainty Areas

	Total Sample		Certainty Area		Noncertainty Area			
	Total Sample		Northeast		Midwest		South	
	Field Experiment	Comparison Group	Field Experiment	Comparison Group	Field Experiment	Comparison Group	Field Experiment	Comparison Group
Ineligible Dwelling Unit	**16.28**	**15.30**	**15.77**	**12.74**	**16.80**	**18.76**		
Vacant	11.80	11.32	11.96	9.39	11.63	13.85		
Not a Primary Residence	1.63	1.79	0.74	0.92	2.53	2.93		
Not a Dwelling Unit	2.30	2.07	2.37	2.30	2.22	1.76		
Other	0.56	0.12	0.70	0.12	0.42	0.11		
Ineligible Dwelling Unit	**16.28**	**15.30**	**21.71**	**17.37**	**13.68**	**14.89**	**17.20**	**16.20**
Vacant	11.80	11.32	16.07	11.06	11.29	11.77	12.04	12.52
Not a Primary Residence	1.63	1.79	2.20	2.88	0.16	0.94	2.25	2.15
Other	0.56	0.12	0.70	0.12	0.42	0.11		
Not a Dwelling Unit	2.30	2.07	3.04	3.43	1.97	2.10	2.04	1.45
Other	0.56	0.12	0.49	0.01	0.26	0.13	0.87	0.14
Eligible Dwelling Unit, Not Screened for Eligible Persons	**14.60**	**6.37**	**17.43**	**7.61**	**11.72**	**4.64**		
Eligible Dwelling Unit, Not Screened for Eligible Persons	**14.60**	**6.37**	**10.26**	**6.58**	**17.57**	**6.62**	**12.26**	**4.80**
No One at Home	2.08	2.09	2.43	2.48	1.73	1.53		
Refusal	5.84	2.46	6.61	2.89	5.71	2.11		
Denied Access	2.91	0.59	5.56	0.89	7.07	0.17	1.86	1.79
Newton Screener Problem	0.57	0.00	0.65	0.00	0.22	0.49	5.33	1.94
Other Nonresponse (Group Quarters)[1]	3.19	1.24	2.19	0.54	3.26	0.22	1.14	0.23
Newton Screener Problem	0.57	0.00	0.32	0.00	0.52	1.32	0.64	0.00
Other Nonresponse (Group Quarters)[1]	3.19	1.24	1.53	0.67	5.06	0.00	3.29	0.90
Eligible Dwelling Units, Screened for Eligible Persons	**85.40**	**93.63**	**82.57**	**92.39**	**88.28**	**95.36**		
Eligible Dwelling Units, Screened for Eligible Persons	**85.40**	**93.63**	**89.74**	**93.42**	**82.43**	**93.38**	**87.74**	**95.14**

The Newton application used for the 1997 field experiment did not handle group quarters.

	Total Sample		Large MSA		Small MSA		Non-MSA	
	Field Experiment	Comparison Group	Field Experiment	Comparison Group	Field Experiment	Comparison Group	Field Experiment	Comparison Group
Ineligible Dwelling Unit	**16.28**	**15.30**	**15.19**	**12.52**	**17.02**	**18.19**	**17.85**	**19.07**
Vacant	11.80	11.32	11.45	9.33	11.49	13.48	13.54	13.83
Not a Primary Residence	1.63	1.79	0.62	0.81	3.36	2.73	0.70	3.33
Not a Dwelling Unit	2.30	2.07	2.42	2.27	1.68	1.81	3.34	1.90
Other	0.56	0.12	0.70	0.11	0.50	0.17	0.26	0.01
Eligible Dwelling Unit, Not Screened for Eligible Persons	**14.60**	**6.37**	**17.36**	**7.92**	**12.05**	**4.91**	**11.99**	**3.48**
No One at Home	2.08	2.09	2.47	2.66	2.07	1.60	0.91	0.87
Refusal	5.84	2.46	6.58	2.89	5.93	2.09	3.37	1.56
Denied Access	2.91	0.59	5.54	0.96	0.64	0.18	0.05	0.03
Newton Screener Problem	0.57	0.00	0.67	0.00	0.60	0.00	0.16	0.00
Other Nonresponse (Group Quarters)[1]	3.19	1.24	2.09	1.40	2.81	0.73	7.49	1.02
Eligible Dwelling Units, Screened for Eligible Persons	**85.40**	**93.63**	**82.64**	**92.08**	**87.95**	**95.55**	**88.01**	**96.52**

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures: 1997 Field Experiment. Survey on Drug Abuse: Quarter 4.

See notes at end of exhibit.

Exhibit 10.1.2 Dwelling Unit Screening Nonresponse Categories: Field Experiment Versus NHSDA Quarter 4 Comparison Group

	Field Experiment	Comparison Group	Difference
Small or No Difference			
No one home	2.1	2.1	
Newton screener problem	0.6	0.0	
Total	2.7	2.0	-0.7
Large Difference			
Refusal	5.8	2.5	
Denied access	2.9	0.6	
Other nonresponse[1]	3.2	1.2	

Total	11.9	4.3	-7.6

[1]Group quarters component was 1.4.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 10.1.3 Person Response Rate Percentages for All Persons, by Sociodemographic Cha[racteristics] (continued)

	Total Sample Field Experiment	Total Sample Comparison Group[1]	12 to 17 Year Olds Field Experiment	12 to 17 Year Olds Comparison Group	18 to 25 Year Olds Field Experiment	18 to 25 Year Olds Comparison Group	26 to 34 Year Olds Field Experiment	26 to 34 Year Olds Comparison Group
Nonresponse Rate	**37.34**	**26.31**	**27.33**	**18.50**	**40.28**	**27.66**	**46.39**	**28.59**
No one home/R unavailable	5.69	6.43	4.55	4.45	9.02	8.57	7.03	6.83
Physical/mental incompetent	1.80	0.95	1.37	0.65	0.60	0.49	1.33	0.53
Language barrier	5.94	8.78	1.17	4.49	7.21	9.58	11.03	10.56
Refusal	19.76	8.62	17.44	7.95	18.44	7.04	20.34	8.85
Other	4.14	1.53	2.80	0.96	5.01	2.24	6.65	1.82
Overall Response Rate	**62.66**	**74.10**	**72.67**	**81.50**	**59.72**	**72.34**	**53.61**	**71.41**

	Total Sample Field Experiment	Total Sample Comparison Group[1]	Certainty Area Field Experiment	Certainty Area Comparison Group	Noncertainty Area Field Experiment	Noncertainty Area Comparison Group
Nonresponse Rate	**37.34**	**26.31**	**43.40**	**33.09**	**30.79**	**16.36**
No one home/R unavailable	5.69	6.43	[illegible]	[illegible]	[illegible]	[illegible]
Physical/mental incompetent	1.80	0.95	[illegible]	[illegible]	[illegible]	[illegible]
Language barrier	5.94	8.78	[illegible]	[illegible]	[illegible]	[illegible]
Refusal	19.76	8.62	[illegible]	[illegible]	[illegible]	[illegible]
Other	4.14	1.53	[illegible]	[illegible]	[illegible]	[illegible]
Overall Response Rate	**62.66**	**74.10**	**56.60**	**66.91**	**69.21**	**83.64**

	Total Sample Field Experiment	Total Sample Comparison Group[1]	1 Person Selected in DU Field Experiment	1 Person Selected in DU Comparison Group	2 People Selected in DU Field Experiment	2 People Selected in DU Comparison Group
Nonresponse Rate	**37.34**	**26.31**	**36.02**	**23.82**	**38.74**	**29.98**
No one home/R unavailable	5.69	6.43	[illegible]	[illegible]	[illegible]	[illegible]
Physical/mental incompetent	1.80	0.95	[illegible]	[illegible]	[illegible]	[illegible]
Language barrier	5.94	8.78	[illegible]	[illegible]	[illegible]	[illegible]
Refusal	19.76	8.62	[illegible]	[illegible]	[illegible]	[illegible]
Other	4.14	1.53	3.87	1.36	4.43	1.79
Overall Response Rate	62.66	74.10	63.98	76.18	61.26	70.02

[1]The response rates for the comparison group that are shown in this exhibit are slightly lower than those shown in Exhibit 5.4.2 (74.1% vs. 75.6%) becaus[e] ... prior to finalizing the 1997 NHSDA screening and interviewing results. This difference does not change the conclusions.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Proced[ures] ... Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

	Total Sample Field Experiment	Total Sample Comparison Group[1]	Hispanic Field Experiment	Hispanic Comparison Group	Non-Hispanic, Black Field Experiment	Non-Hispanic, Black Comparison Group	Non-Hispanic, Non-Black Field Experiment	Non-Hispanic, Non-Black Comparison Group
Nonresponse Rate	**37.34**	**26.31**	**42.96**	**45.32**	**34.20**	**17.81**	**35.97**	**19.70**
No one home/R unavailable	5.69	6.43	4.25	6.21	7.19	7.71	5.88	5.80
Physical/mental incompetent	1.80	0.95	0.49	0.64	2.60	0.85	2.09	1.20
Language barrier	5.94	8.78	20.15	30.18	0.62	0.17	1.11	0.75
Refusal	19.76	8.62	13.11	6.34	19.95	7.32	23.24	10.81
Other	4.14	1.53	4.98	1.96	3.84	1.75	3.85	1.14
Overall Response Rate	**62.66**	**74.10**	**57.04**	**54.68**	**65.80**	**82.19**	**64.03**	**80.30**

	Total Sample Field Experiment	Total Sample Comparison Group[1]	Northeast Field Experiment	Northeast Comparison Group	Midwest Field Experiment	Midwest Comparison Group	South Field Experiment	South Comparison Group
Nonresponse Rate	**37.34**	**26.31**	**37.83**	**18.77**	**34.84**	**24.25**	**34.90**	**21.07**
No one home/R unavailable	5.69	6.43	6.86	8.07	6.38	6.83	5.58	5.39
Physical/mental incompetent	1.80	0.95	1.65	0.83	1.87	1.18	2.38	0.86
Language barrier	5.94	8.78	4.49	4.41	1.24	3.42	4.01	6.92
Refusal	19.76	8.62	20.57	4.56	23.95	11.14	19.05	6.70
Other	4.14	1.53	4.26	0.90	1.40	1.69	3.88	1.20
Overall Response Rate	**62.66**	**74.10**	**62.17**	**81.23**	**65.16**	**75.75**	**65.10**	**78.93**

	Total Sample Field Experiment	Total Sample Comparison Group[1]	Male Field Experiment	Male Comparison Group	Female Field Experiment	Female Comparison Group
Nonresponse Rate	**37.34**	**26.31**	**38.85**	**28.42**	**35.94**	**24.66**
No one home/R unavailable	5.69	6.43	5.87	7.34	5.53	5.72
Physical/mental incompetent	1.80	0.95	1.85	0.99	1.76	0.91
Language barrier	5.94	8.78	6.20	9.32	5.71	8.35
Refusal	19.76	8.62	20.58	9.26	19.00	8.12
Other	4.14	1.53	4.35	1.31	3.95	1.55
Overall Response Rate	**62.66**	**74.10**	**61.15**	**71.58**	**64.06**	**75.34**

	Total Sample Field Experiment	Total Sample Comparison Group[1]	Large MSA Field Experiment	Large MSA Comparison Group	Small MSA Field Experiment	Small MSA Comparison Group	Non-MSA Field Experiment	Non-MSA Comparison Group
Nonresponse Rate	**37.34**	**26.31**	**43.84**	**32.37**	**34.38**	**21.88**	**25.19**	**13.26**
No one home/R unavailable	5.69	6.43	6.83	7.57	4.70	5.74	4.62	3.68
Physical/mental incompetent	1.80	.95	2.07	0.98	1.22	0.94	2.31	0.82
Language barrier	5.94	8.78	8.84	12.52	4.44	5.49	0.96	1.89
Refusal	19.76	8.62	21.02	9.53	20.28	8.45	15.00	5.61
Other	4.14	1.53	5.09	1.77	3.74	1.26	2.31	1.25
Overall Response Rate	**62.66**	**74.10**	**56.16**	**67.63**	**65.62**	**78.12**	**74.81**	**86.74**

See notes at end of exhibit. (continued)

See notes at end of exhibit.(continued)

Exhibit 10.1.4 Comparison of Person Nonresponse Categories for All Persons: Field Experiment Versus Quarter 4 Surveys

	Field Experiment	Comparison Group	Difference
Small or No Difference			
Not home/response unavailable	5.7	6.4	
Physical/mental incapable	1.8	1.0	
Language barrier	5.9	8.8[1]	
Total	13.4	16.2	+2.8
Large Difference			
Refusal	19.8	8.6	

Other nonresponse	4.1	1.5	
Total	23.9	10.1	-13.8

[1]Includes respondents who completed a Spanish-language questionnaire.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 10.1.5 Refusal Rates, by Age Group and Sample

Sample	12 to 17 Years	18 to 25 Years	26 to 34 Years	35 Years or Older
Field Experiment	17.4	18.4	20.3	26.3
Comparison Group	8.0	7.0	8.8	10.6

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 10.1.6 Language Barrier Rates, by Age Group and Sample

Sample	12 to 17 Years	18 to 25 Years	26 to 34 Years	35 Years or Older
Field Experiment	1.2	7.2	11.0	12.6
Comparison Group	4.5	9.6	10.6	11.2

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 10.1.7 Person Response Rate Percentages for 12- to 17-Year-Olds by Race/Ethnicity, Gender, and Number of Sample Persons in DU

Exhibit 10.1.8 Person Response Rate Percentages for Adults 18 Years Old or Older, by Race/Ethnicity, Gender, and Number of Sample Persons in DU

	Total Sample		Hispanic		Non-Hispanic, Black		Non-Hispanic, Non-Black	
	Field Experiment	Comparison Group	Field Experiment	Comparison Group	Field Experiment	Comparison Group	Field Experiment	Comparison Group
Nonresponse Rate	**27.33**	**18.50**	**24.24**	**26.33**	**26.63**	**14.91**	**29.35**	**15.4**
No one home/R unavailable	4.55	4.45	3.73	4.30	5.62	6.56	4.55	3.3
Physical/mental incompetent	1.37	0.65	0.47	0.76	0.59	0.72	2.21	0.5
Language barrier	1.17	4.49	1.26	3.05	1.43	0.89	0.52	0.3
Refusal	17.44	7.95	13.29	5.38	17.75	6.27	19.61	10.5
Other	2.80	0.96	3.50	1.14	2.66	1.30	2.47	0.6
Overall Response Rate	**72.67**	**81.50**	**75.76**	**73.67**	**73.37**	**85.09**	**70.65**	**84.5**
Overall Response Rate	**53.20**	**70.61**	**36.71**	**46.64**	**60.34**	**81.14**	**57.35**	

	Total Sample		Male		Female	
	Field Experiment	Comparison Group	Field Experiment	Comparison Group	Field Experiment	Comparison Group
Nonresponse Rate	**27.33**	**18.50**	**26.48**	**19.14**	**28.19**	**17.89**
No one home/R unavailable	4.55	4.45	4.37	4.30	4.74	4.61
Physical/mental incompetent	1.37	0.65	1.67	0.79	1.05	0.51
Language barrier	1.17	4.49	1.29	4.75	1.05	4.25
Refusal	17.44	7.95	16.32	8.59	18.58	7.33
Other	2.80	0.96	2.83	0.72	2.77	1.20
Overall Response Rate	**72.67**	**81.50**	**73.52**	**80.86**	**71.81**	**82.11**
Other	5.41	1.76	5.96	1.88	4.95	1.67

	Total Sample		1 Person Selected in DU		2 People Selected in DU	
	Field Experiment	Comparison Group	Field Experiment	Comparison Group	Field Experiment	Comparison Group
Overall Response Rate	**53.20**	**70.61**	**48.10**	**67.29**	**57.43**	**73.01**

	Total Sample		1 Person Selected in DU		2 People Selected in DU	
	Field Experiment	Comparison Group	Field Experiment	Comparison Group	Field Experiment	Comparison Group
Nonresponse Rate	**27.33**	**18.50**	**26.10**	**15.81**	**28.55**	**20.80**
No one home/R unavailable	4.55	4.45	3.38	4.01	5.74	4.83
Physical/mental incompetent	1.37	0.65	1.56	0.64	1.17	0.65
Language barrier	1.17	4.49	0.91	3.05	1.43	5.72
Refusal	17.44	7.95	18.05	7.54	16.82	8.29
Other	2.80	0.96	2.21	0.50	3.39	1.30
Overall Response Rate	**72.67**	**81.50**	**73.90**	**84.19**	**71.45**	**79.20**
Refusal	21.96	8.88	20.07	7.87	24.06	10.76
Other	5.41	1.76	5.37	1.58	5.46	2.08
Overall Response Rate	**53.20**	**70.61**	**55.08**	**73.94**	**51.11**	**64.45**

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures: 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 10.1.9 Percentage of Eligible Persons, by Response Category and by Age Group

Age Group and Survey	Response Not Available	Physical or Mental Problems	Language Barrier	Refusal	Other Nonresponse	Response Rate
12 to 17 Year Olds						
Field experiment	4.6	1.4	1.2	17.4	2.8	72.7
Comparison group	4.4	0.6	4.5	8.0	1.0	81.5
Difference	0.2	0.8	-3.3	9.0	1.8	-8.8
Adults						
Field experiment	6.8	2.2	10.5	22.0	5.4	53.2

Comparison group	7.2	1.1	10.5	8.9	1.8	70.6
Difference	-0.4	1.1	0.0	13.1	3.6	-17.4

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

Exhibit 10.1.10 Field Experiment Versus Comparison Group Response Rates, by Subgroup

Subgroup	Response Rates for 12 to 17 Year Olds			Response Rates for 18 Year Olds or Older		
	Field Experiment	Comparison Group	Difference	Field Experiment	Comparison Group	Difference
Total	72.7	81.5	-8.8	53.2	70.6	-17.4
Hispanic	75.8	73.7	+2.1	36.7	46.6	-9.9
Black	73.4	85.1	-11.7	60.3	81.1	-20.8
White/Other	70.5	84.8	-14.0	57.4	78.6	-21.2
Males	73.5	80.9	-7.4	48.1	67.3	-19.2
Females	71.8	82.1	-10.3	57.4	73.0	-15.6
1 Selected in DU	73.9	84.2	-10.3	55.1	73.9	-18.8
2 Selected in DU	71.4	79.4	-8.0	51.1	64.4	-13.3
Northeast	70.0	86.3	-16.3	53.9	79.4	-25.5
Midwest	69.4	80.9	-11.0	61.3	73.6	-12.3
South	75.7	86.3	-10.6	55.6	76.0	-20.4
West	71.1	68.9	+2.2	38.1	51.2	-13.1
Large MSAs	67.9	77.1	-9.2	46.5	64.1	-17.6
Small MSAs	75.0	84.2	-9.2	55.8	75.6	-19.8
Non-MSA areas	79.0	90.2	-11.2	69.7	85.2	-15.5
Certainty PSUs	69.6	77.1	-7.5	45.4	63.1	-17.7
Noncertainty PSUs	75.7	87.4	-11.7	62.4	82.0	-19.6

DU = dwelling unit; MSA = metropolitan statistical area; PSU = primary sampling unit.

Sources: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment. 1997 National Household Survey on Drug Abuse: Quarter 4.

10.2 Effect of Electronic Instruments on Response Rates

Why were the field experiment's response rates so much lower than those from the comparison group sample? Was it because the field experiment used electronic survey instruments? It is not possible to answer this question directly, due to the confounding of the effect of the separate survey staffs with that of the electronic versus PAPI instrumentation. Still, several data and anecdotal indications may partially answer the question. We discuss these indications in this section.

There were some problems with the electronic instruments and the case management system (CMS). The Newton screener and the CAPI instrument were newly developed applications that, due to time limitations, had not been perfected by the beginning of the field experiment's data collection. Due to timing, the FI training also was less comprehensive than a fully developed, ideal training package. Field management staff noted that the Newton worked very well with only minor problems and that the computer-assisted personal interviewing (CAPI) software also worked quite well. Minor problems were encountered in changing some of the previous entries, and this may have contributed to the higher field experiment nonresponse rates.

Significant problems with the CMS were encountered early in the survey quarter. The FSs had difficulty transferring cases among interviewers; this was perceived as a major problem with CAPI cases, although it was not a problem with Newton cases. Headquarters and field supervisory staff also became very frustrated with the lack of accuracy in the CMS field reports, which they were using to manage the survey. The fixed time interval in which to complete the fieldwork and the frustrating problems experienced during the early data collection period may have pushed a larger fraction of the data collection into November and December 1997 and these holiday months are less conducive to effective fieldwork. Thus, such factors as the holiday months, the less than optimal flexibility of the newly developed electronic survey application, especially the CMS, and a new field staff probably increased field experiment's nonresponse.

How much of the response rate differences could be due to using the electronic instruments? At the DU screening stage, the interviewers indicated through a special code that less than 1% of the eligible DUs were not screened because of Newton problems. This indicates that there was little if any DU nonresponse because the Newtons were used for screening. The nonresponse categories that showed large differences at the screening level were refusals, denied access, and other nonresponse (Exhibit 10.1.2). It is very difficult to believe that interviewers approaching households would be more likely to be denied access to DUs, or that more sample households would be more apt to refuse to give the DU screening information because the FI carried a Newton computer. Experienced supervisors and interviewers, however, thought that there could have been another effect due to the electronic screener. Interviewers know that people are not selected in most households, and if the household has no youths and no Hispanics, that there is a low chance that anyone will be selected. When using the paper sample selection tables, interviewers can determine if anyone will be selected for certain household types prior to the screening. They could then use this information to facilitate the screening by stating, for example, "I just have a very few questions. I don't think that anyone who lives here is eligible for this survey, but I just need to ask you a few questions to be sure." This option is not available to the interviewers using the electronic screening application because the sample selection tables are not visible to the interviewer. Although this is a possibility, it is likely to be a less important factor than the lack of expertise on the part of the interviewers.

Exhibit 10.1.4 indicates that the refusal and other nonresponse categories accounted for the lower person-level response rates for the field experiment. It also is difficult to believe that more persons would refuse to participate in the survey because the interviewer planned to use a portable computer to complete

portions of the interview. Most refusals occur before interviews are begun; thus, many of the sample persons who refused may not have even been aware that the interview would involve using a computer.

Higher nonresponse in the other nonresponse category could indicate difficulties due to electronic data collection. As discussed earlier in this section, there is reason to believe that the deficiencies in the newly developed electronic survey application may have contributed to the higher field experiment nonresponse. Use of the other nonresponse category in the field experiment accounted for approximately 24% of the response rate decrease, at both the DU screening and person interviewing levels.

Therefore, although no definite conclusions can be made in this regard, it is reasonable to attribute at most 20% to 25% of the field experiment's response rate shortfall to the electronic survey environment. The deficiencies in the electronic survey application, as discussed earlier in this section, were the lack of a Spanish translation, difficulties in changing some entries in the computers, and problems in the CMS. Fortunately, all of these deficiencies can be corrected, avoiding any such response rate decreases due to the electronic survey environment in future NHSDA survey years.

10.3 Effect of Experimental Environment on Response Rates

The experimental environment in which the field experiment and comparison surveys were conducted differed in several ways. First, a completely separate staff of FSs and FIs were hired to work on the field experiment, which ensured that the 1997 NHSDA could not be negatively affected in that none of its data collection resources were diverted to conducting the field experiment survey. The field experiment's FSs were less experienced than the comparison survey's FSs. And the field experiment's supervisors also lacked the specific experience of currently working as 1997 NHSDA FSs.

Similar statements can be made in comparing the 165 FIs who worked on the pretest with the 270 FIs who worked on the comparison survey. The two groups of interviewers were remarkably similar on many of the characteristics for which data were available, including their educational level, gender, age, race, and ethnicity. However, the comparison group interviewers were more experienced in general survey interviewing and were all currently working as 1997 NHSDA FIs. Headquarters staff also commented that the field experiment's interviewers as a group lacked the level of refusal conversion expertise of the interviewers working on the comparison survey. This is obviously a critical skill for gaining access to sample DUs and convincing reluctant individuals that they should cooperate with the survey.

The two field staffs were both responsible to the same headquarters survey management staff. The central office, FSs, and FIs all exhibited very positive attitudes toward the field experiment study. Still, there may have been some unavoidable tendency to see the field experiment as a onetime add-on to the continuing NHSDA "bread and butter" project. There are no data to suggest that there was any effect on the study response rates due to the attitudes of any project staff.

Returning to the interviewing experience deficiency of the field experiment interviewers, it may be observed in Exhibits 10.1.2 and 10.1.4 that the categories of both DU screening and person interviewing that accounted for small response rate differences are the categories that one would not expect to be affected by interviewer experience levels. That is, one would not expect interviewer experience or expertise to have a significant effect on the percentage of dwellings in which no one was found at home, the eligible persons who were unavailable when calls were made, who were physically or mentally unable to respond to the survey, or who were unable to participate due to language capability.

Conversely, the categories of nonresponse that are logically related to large response rate differences in Exhibits 10.1.2 and 10.1.4 are those that are probably related to interviewer experience levels (refusals, denied access, and other nonresponse).

In addition, anecdotal information indicates that the traveling interviewer strategy used for the field experiment was not as effective as expected. Interviewers selected for this assignment were experienced veteran interviewers but lacked computer expertise and were not accustomed to working for extended periods away from their homes. Once the traveling interviewer team was assembled and trained, they had difficulty working effectively due to the remaining data collection time and complications of the holiday season.

No conclusive experimental data clearly show whether the experimental environment of the field experiment was responsible for the lower response rates. But it appears that the lower level of general survey experience and specific NHSDA experience, plus the refusal conversion experience deficit of the field experiment's interviewing staff, may have been the main factors causing the response rate shortfall. Thus, it seems reasonable to conclude that roughly 75% of the shortfall was due to the experimental environment.

10.4 Effect of Electronic Instruments on Respondent Attitudes and Self-Reported Willingness to Respond

Chapter 6 contains a detailed examination of the field experiment (CAPI) and comparison group (PAPI) respondents' attitudes toward the surveys. This analysis, based on respondent debriefings conducted following both surveys, concluded the following:

1. The CAPI respondents were more likely to feel that the interviewers did not see any of their answers.

2. With respect to protecting their privacy, only about 11% of the respondents to both surveys felt that answer sheets were the best method of data collection.

3. The CAPI survey respondents found the survey somewhat more interesting than the PAPI respondents did.

4. A slightly higher percentage of the CAPI respondents said they would be likely to participate in the survey again.

5. The CAPI respondents were slightly more likely to feel that their survey information would never be linked with their names.

Thus, the respondent debriefings confirmed the results of the 1996 feasibility experiment: The CAPI respondents tended to be more positive than PAPI respondents on privacy and confidentiality issues; they preferred to use the computer to respond to the survey; and those responding by computer said they would be more likely to participate in the survey if it were done again. All of these results seem to indicate respondent attitudes toward the electronic survey methods were very positive, and that there is no reason to believe that the electronic survey environment will have any negative effects on respondents' willingness to completely respond to future NHSDAs once they have begun the interview. However, most decisions to respond are made prior to the subject encountering the instrument.

11. Effect of Interviewers on Data Quality

In this chapter, we report the results of two investigations of the effect of the interviewers on data quality. We examine the relationship between interviewer characteristics and performance measures and look at the effects of the mode of interview on interviewer variance, which is the increase in variance due to the correlation of response errors within interviewers.

11.1 Relationship Between Interviewer Characteristics and Interviewer Performance

In this section, we describe the characteristics of the 1997 field experiment's field interviewers (FIs) and examine the relationship of those characteristics to a number of FI performance measures. This chapter describes an exploratory study to determine whether the quality of the NHSDA interviewer workforce can be improved by selecting them on the basis of their background characteristics.

11.1.1 Field Interviewer Characteristics

There were 165 FIs involved in the 1997 field experiment. For most of the interviewers, data from a personal data sheet (PDS) were available. For the purposes of this study, the following variables were created from the PDS: gender, age in years, a four-category ethnicity variable (black, white, Hispanic, and other), and a three-category education-level variable (high school, some college, and 4-year college graduate).

Eighty-one percent (133) of the interviewers were female. Their average age was 50 years, with a standard deviation of approximately 12. Most were white (73%), 17% were black, and the remaining 10% were Asian, Pacific Islander, American Indian, Alaskan Native, or Hispanic. Some 21% attended high school or were high school graduates; 41% had some college or a 2-year degree; and 36% were college graduates or had postgraduate degrees.

The other set of background variables related to field interviewing experience were as follows: a six-level ordered categorical variable reflecting years of survey interviewing experience (none, up to 1 year, 1 to 2 years, 2 to 5 years, 5 to 10 years, more than 10 years). Of the 143 interviewers who provided information on survey interviewing experience, only 21% had no experience, 48% had up to 5 years of experience, and 31% had 5 or more years of experience.

Five binary indicator variables indicated whether the FI had experience as a data collector or supervisor of data collectors in the following five areas: household surveys with face-to-face interviews; listing area segments; computer-assisted personal interviewing (CAPI); computer-assisted telephone interviewing (CATI) or other telephone interviewing; and field reporting with a microcomputer. These field interviewing experience variables are summarized in Exhibit 11.1.1.

Most of the FIs had either household survey face-to-face interviewing experience (70%) or CAPI experience (58%). Also, 79% of the interviewers had either household survey or CAPI experience. Some 49% had experience in both areas.

11.1.2 Methodology

Each interviewer completed a 23-item computer beliefs and attitudes questionnaire at three time points: prior to training, immediately after training, and at the end of the 1997 field experiment study. The time between the pre- and post-administrations was approximately 2 weeks, and the time between the post- and end-of-study administration was approximately 3 months. The 23 items were scored using a 5-point Likert scale: 5 (strongly agree), 4 (agree), 3 (neither agree nor disagree), 2 (disagree), and 1 (strongly disagree).

Twenty-one of these items came from an instrument used by the U.S. Bureau of the Census to study their FIs. Research Triangle Institute (RTI) added two items concerning the use of the Newton in screening. We decided to drop these two items so that our results might eventually be compared with those of the Census Bureau. The 21 items measured, for the most part, beliefs and attitudes about the role of computers in general and specifically with field interviewing. The correlation matrices for the 21 items were factor analyzed separately for each of the three time points. For each administration, two through five factors were extracted and rotated by Varimax. We felt that the best criterion for selecting a particular factor solution (i.e., number of factors) at each time point should be based on how effectively they predicted the FI performance measures.

The factor analyses of the beliefs and attitude questionnaires were undertaken (a) to reduce the number of independent variables in the interviewer performance regression models, (b) to address the problem of multicollinearity, and (c) to simplify the interpretation of results from the regression analyses. The unit of analysis was the interviewer. About 140 interviewers met the

criteria to be included in the regression analyses. Therefore, we wanted to minimize the number of independent variables to be included in the regression models.

To model interviewer performance, we needed to control for the segment characteristics in which each interviewer worked. Nine segment characteristics were selected for this purpose:

1. Quarter 1 through 3 successful screening percentage (number of households screened divided by number of eligible households),
2. Quarter 1 through 3 successful interview percentage (number of completed interviews divided by the number of people who were selected to be interviewed),
3. median rent,
4. median value of owner-occupied dwelling,
5. percentage of noninstitutionalized Hispanics 12 years of age or older,
6. percentage of noninstitutionalized black, non-Hispanics 12 years ofage or older,
7. percentage of urban population in segment,
8. a binary urbanized area indicator, and
9. an ordered three-level population density variable.

The performance measures for the interviewers were based upon a single segment, which was the first clean segment in which the interviewer worked. A clean segment was one in which only one interviewer worked. The reason for this definition was to avoid contaminating the performance measures for a particular interviewer with the performance of one or more other interviewers. In most cases, an interviewer worked in only one segment, and that segment was usually clean.

The following interviewer performance measures were used as the five dependent variables in the regression modeling: the percentage of eligible dwelling units that were successfully screened; the percentage of completed interviews from selected persons; the probability of the interviewer having trouble setting up the computer as judged by the respondent; the probability of the interviewer having trouble with the computer during the interview as judged by the respondent; and the number of days from assignment of segment to beginning work in the segment.

Logistic regression analysis was used to model the four binomial outcome variables. For each interviewer, we had the number of trials (e.g., the number of persons selected for an interview) and the number of successful outcomes or events (e.g., the number of completed interviews). All of the logistic regression analyses were adjusted for overdispersion because the interview completion events within interviewers were not independent. A Cox proportional hazards model was used to model the time from assignment to beginning work in the assigned segment.

Because there were eight different factor solutions (two- to five-factor solutions for both the pre-training and post-training beliefs and attitudes questionnaires), we modeled each of the five interviewer performance measures as a function of each of the eight sets of factor scores. Thus, there were a total of 40 regression analyses. In general, the attitude and belief factors were not very predictive of interviewer performance. Overall, the post-training five-factor solution seemed to be the best set of factors for predicting the five performance measures, perhaps because these factors were measured after training and because they were also measured closer to the fieldwork activities on which the performance measures were based. Also, five factors explain more variation in the 21-questionnaire items than two-, three-, or four-factor solutions.

The five rotated factors accounted for 60% of the total variation in the 21 attitude and belief questions. The first factor had high loadings from 10 of the 21 items ranging from 0.55 to 0.78. It explained the most variation in the attitude items. We interpret this factor as a general expression of positive beliefs and attitudes toward computers both in general and specifically withrespect to their utility for interviewing.

The remaining four factors explained considerably less variation and had only between two and four variables loading high (i.e., about 0.50 and above) on them. Factor 2 reflected a disagreement with negative beliefs, such as beliefs that computers threaten interviewers' jobs and the feeling that the interviewer is being watched all the time by using a computer. Factor 3 was defined by three questionnaire items that reflected not being afraid of the responsibility for taking care of the computer and other equipment. Factor 4 was defined by three items that reflected lack of concern about computer breakdowns and keying in data. The last factor, Factor 5, was defined by two items that indicated that the interviewer believed that the respondents will enjoy the interview and that carrying the computer around will not be tiring.

Two logistic regression models, described below, were developed for each of the following four performance measures: the probability of a completed interview, the probability of a successful screening, the probability of having trouble setting up the

computer, and the probability of having trouble with the computer during the interview. The independent variables came from three variable sets: nine segment characteristics, eight interviewer background and experience characteristics, and the five-factor post-training belief and attitude factors. Using the same independent variables, the Cox proportional hazards model was used to model the number of days from segment assignment to start of work in the segment.

Because there were a large number of independent variables relative to the number of observations, the following strategy was adopted for the first logistic regression model. The five post-treatment belief and attitude factors were included in all of the models. CAPI experience also was included in all four models. A stepwise logistic regression was run separately for the nine segment characteristics to select segment characteristics for inclusion in one of the two final models for each performance measure. Likewise, a stepwise logistic regression analysis was run on the eight interviewer background and experience variables to select interviewer characteristics for inclusion in one of the two final models for each performance measure. Variables significant at the 0.10 level for each of the two stepwise regression analyses along with the five attitude and belief factors (as well as CAPI experience) were included in a final model for each of the four performance measures.

The second model-building strategy was to do a single stepwise logistic regression analysis for each of the four interviewer performance measures. The variables were selected from the entire pool of 22 independent variables (i.e., the nine segment characteristics, the eight interviewer background and experience characteristics, and the five post-training factors). For all of the stepwise regressions, the 0.15 significance level was chosen for entry into the model, and 0.20 was chosen for removal from the model.

Thus, for each of the four performance measures, two models were developed. One model involved selecting the variables for inclusion in the final model (henceforth referred to as Model 1) by selecting the significant variables separately from the segment and interviewer background characteristics sets by stepwise logistic regressions and then combining the significant variables from these two stepwise regressions with the five post-training attitude and belief factorsand CAPI experience to estimate a final model. The other final model (henceforth referred to as Model 2) involved selecting the variables from a single stepwise logistic regression analysis using the single pool of all 22 variables from all three variable classes.

11.1.3 Results

For modeling the probability of completing an interview, both models indicated that the most important variables were segment characteristics. In both models, there was a highly significant negative relationship between the percentage of Hispanics in the segment and the probability of a completed interview; moreover, there was a highly significant positive relationship between the percentage of black, non-Hispanics in the segment and the probability of a completed interview. For Model 1 (see Exhibit 11.1.2), there were no significant parameters for any of the interviewer background variables or for any of the five belief and attitude factors. For Model 2 (see Exhibit 11.1.3), two interviewer variables were significant: education level and household survey experience. Education level was positively associated with the probability of a completed interview, and household survey experience was negatively associated. (The significance levels were 0.04 and 0.05, respectively.) Factor 2, reflecting disagreement with the belief that computers are watching the interviewer, and reflecting disagreement with the belief that computers may threaten interviewers' job security, was negatively associated (p=.02) with the probability of a completed interview. The direction of two of these relationships (household survey experience and Factor 2) was counterintuitive.

For modeling the probability of completing a household screening, once again both models indicated that the most important variables were segment characteristics. For Model 1 (see Exhibit 11.1.4), percentage Hispanic and median rent were negatively associated with the probability of a completed screening, while median value of owner-occupied housing and the Quarters 1 to 3 completed screening rate were positively associated with the screening rate. None of the interviewer background variables was significant. However, Factor 2 was once again significant, but in the expected direction. Interviewers who disagreed with the negative belief statements had a higher probability of completing household screenings. For Model 2 (see Exhibit 11.1.5), percentage Hispanic and median rent were negatively associated with the probability of a completed screening as in Model 1. The percentage of the segment that was urban had a positive relationship with the screening rate. In this model, interviewers who had computer-assisted interviewing experience had about two times the odds of a completed screening (p=.04) compared to those with non-CAPI experience. Factor 2 was once again significant (p=.01) in the expected direction. A one-scale point increase in Factor 2 increased the odds of a completed screening by about 44%.

Model 1 (see Exhibit 11.1.6), for modeling the probability of having trouble in setting up the computer for the CAPI, indicated that none of the 22 candidate variables was significant. The overall level of significance of the model was only 0.44. For Model 2 (see Exhibit 11.1.7), only two variables entered the model (percentage urban in the segment and interviewer's age). However, neither was significant at the 0.05 level.

For Model 1 (see Exhibit 11.1.8), in modeling the probability of having computer troubleduring the interview, only two segment characteristics were marginally significant. Percentage Hispanic had a negative parameter (p=0.07), and population density had a positive parameter (p=0.09). Factor 5 had a positive relationship (p=0.03) with the probability of having computer problems during the interview. For Model 2 (see Exhibit 11.1.9), percentage Hispanic and population density had negative and positive parameter, respectively, and were significant at the 0.09 and 0.02 levels, respectively. Age and Factor 5 had positive relationships with computer problems. (The significance levels were 0.03 and 0.01, respectively.) Older interviewers had a higher probability of experiencing computer trouble. Factor 5 was defined primarily by the belief that respondents will enjoy the interview and the belief that carrying computers around will not be tiring. It is puzzling why this factor should be positively associated with computer troubles during the interview.

The Cox proportional hazards models indicated that the time from segment assignments to starting work in the segment could not be predicted from these same independent variables.

11.1.4 Discussion

Segment characteristics were the most important predictors of the probability of a successful screening and of the probability of a completed interview. In general, the six survey experience variables were not predictive of interviewer performance. The only exception was the finding that CAPI experience was positively related to the successful screening rate. The five attitude and belief factors were marginally predictive, but in some cases the direction of the relationship was counterintuitive.

In general, there was not a lot of variation in the dependent variables. The screening rates were generally in the 0.90s, the interview rates were in the high 0.70s, and the two computer problem variables were in the 0.10s. Nevertheless, the screening and interview rates were fairly predictable from the segment characteristics. One of the reasons that the survey experience variables were not predictive of performance is that five of the six variables simply indicated whether or not they had any experience in that particular area, but did not indicate the amount of experience with respect to either years or number of surveys in each of the five areas.

Another problem was the large number of independent variables relative to the number of interviewers. With four performance measures and 22 independent variables, we would expect four or five parameters to be significant at the 0.05 level just by chance. The situation is even worse when variables are selected by stepwise procedures. This could be a reason why the relationship of the attitude and belief factors to the performance measures is in some instances counterintuitive. Because the independent variables came from different sources, missing data further reduced the sample sizes for the regression analyses.

Exhibit 11.1.1 Interviewer Experience

Area of Experience	Experience		No Experience	
Listing Area Segments	58	(35%)	107	(65%)
Computer-Assisted Personal Interviewing	96	(58%)	69	(42%)
Computer-Assisted Telephone Interviewing	72	(44%)	93	(56%)
Field Reporting with a Microcomputer	33	(20%)	132	(80%)
Household Surveys-Face-to-Face Interviews	115	(70%)	50	(30%)

Exhibit 11.1.2 Probability of Completed Interviews: Model 1

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	PR > Chi-Square	Standardized Estimates	Odds Ratio
Intercept	1	2.061	0.304	45.979	0.000	0.000	0.000
Hispanic % of Civilian, Noninstitutionalized Population ≥12 years	1	-0.012	0.004	10.732	0.001	-0.141	0.989
Black, Non-Hispanic % of Civilian, Noninstitutionalized Population ≥12 years	1	0.009	0.004	6.416	0.011	0.148	1.009
Urbanized Area Indicator	1	-0.967	0.281	11.811	0.001	-0.200	0.380
Segment Median Value, Owner-Occupied HUs	1	0.000	0.000	2.535	0.111	0.090	1.000

Computer-Assisted Interview Experience	1	-0.272	0.213	1.632	0.202	-0.065	0.762
Post-Training Factor #1: Positive Toward Computers	1	0.104	0.104	0.989	0.320	0.048	1.109
Post-Training Factor #2: Disagreement with Negative Attitudes	1	-0.177	0.117	2.293	0.130	-0.085	0.838
Post-Training Factor #3: Not Afraid of Caring for Computers	1	-0.003	0.110	0.001	0.975	-0.002	0.997
Post-Training Factor #4: Lack of Concern Re Computer Operations	1	-0.030	0.108	0.076	0.782	-0.013	0.971
Post-Training Factor #5: Rs Like Interview; Easy to Do	1	-0.149	0.101	2.199	0.138	-0.072	0.861

HU=housing unit.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 11.1.3 Probability of Completed Interviews: Model 2

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	PR > Chi-Square	Standardized Estimates	Odds Ratio
Intercept	1	-2.883	2.096	1.892	0.169	.	.
Hispanic % of Civilian, Noninstitutionalized Population ≥12 years	1	-0.016	0.004	16.375	0.000	-0.207	0.984
Black, Non-Hispanic % of Civilian, Noninstitutionalized Population ≥12 years	1	0.015	0.005	11.028	0.001	0.236	1.015
Quarters 1-3 Segment Person Response Rate	1	0.009	0.006	2.002	0.157	0.075	1.009
Quarters 1-3 Segment DU Screening Rate	1	0.037	0.022	2.892	0.089	0.104	1.038
Segment Median Rent, Renter-Occupied HUs	1	-0.002	0.001	2.751	0.097	-0.174	0.998
Segment Median Value, Owner-Occupied HUs	1	0.000	0.000	8.555	0.003	0.329	1.000
Urbanized Area Indicator	1	-0.943	0.318	8.793	0.003	-0.202	0.389
Education	1	0.260	0.126	4.254	0.039	0.118	1.297
Household Survey Experience	1	-0.600	0.307	3.824	0.051	-0.112	0.549
Post-Training Factor #2: Disagreement with Negative Attitudes	1	-0.290	0.127	5.261	0.022	-0.142	0.748

DU=dwelling unit; HU=housing unit.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 11.1.4 Probability of Completed Screenings: Model 1

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	PR > Chi-Square	Standardized Estimates	Odds Ratio
Intercept	1	-2.607	2.498	1.089	0.297	.	.
Hispanic % of Civilian, Noninstitutionalized Population ≥12 years	1	-0.016	0.006	7.493	0.006	-0.098	0.984
Quarters 1-3 Segment DU Screening Rate	1	0.073	0.026	7.795	0.005	0.116	1.076
Segment Median Rent, Renter-Occupied HUs	1	-0.004	0.001	14.307	0.000	-0.250	0.996
Segment Median Value, Owner-Occupied HUs	1	0.000	0.000	2.812	0.094	0.100	1.000
Percent Urban Population in Segment	1	0.012	0.004	8.582	0.003	0.124	1.012
Population Density	1	-0.580	0.262	4.894	0.027	-0.111	0.560
Computer-Assisted Interview Experience	1	0.352	0.280	1.584	0.208	0.047	1.422
Post-Training Factor #1: Positive Toward Computers	1	0.200	0.145	1.892	0.169	0.052	1.221
Post-Training Factor #2: Disagreement with Negative Attitudes	1	0.444	0.118	14.255	0.000	0.126	1.559
Post-Training Factor #3: Not Afraid of Caring for Computers	1	-0.097	0.153	0.403	0.526	-0.025	0.907
Post-Training Factor #4: Lack of Concern Re Computer Operations	1	-0.070	0.149	0.222	0.638	-0.019	0.932
Post-Training Factor #5: Rs Like Interview; Easy to Do	1	-0.145	0.134	1.182	0.277	-0.042	0.865

DU=housing unit.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 11.1.5 Probability of Completed Screenings: Model 2

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	PR > Chi-Square	Standardized Estimates	Odds Ratio
Intercept	1	2.598	0.470	30.560	0.000	.	.
Hispanic % of Civilian, Noninstitutionalized Population ≥12 years	1	-0.015	0.007	5.302	0.021	-0.080	0.985
Segment Median Rent, Renter-Occupied HUs	1	-0.002	0.001	5.319	0.021	-0.093	0.998

Percent Urban Population in Segment	1	0.011	0.005	5.909	0.015	0.100	1.011
Computer-Assisted Interview Experience	1	0.737	0.346	4.522	0.034	0.083	2.089
Post-Training Factor #2: Disagreement with Negative Attitudes	1	0.363	0.124	8.606	0.003	0.090	1.437

HU=housing unit.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 11.1.6 Probability of Problems Setting Up Computer: Model 1

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	PR > Chi-Square	Standardized Estimates	Odds Ratio
Intercept	1	-3.517	1.145	9.439	0.002	.	.
Percent Urban Population in Segment	1	0.008	0.007	1.374	0.241	0.108	1.008
Age	1	0.014	0.018	0.640	0.424	0.063	1.014
Field Reporting Microcomputer Experience	1	-0.686	0.633	1.173	0.279	-0.095	0.504
Computer-Assisted Interviewing Experience	1	0.091	0.424	0.046	0.831	0.016	1.095
Post-Training Factor #1: Positive Toward Computers	1	-0.072	0.220	0.106	0.745	-0.024	0.931
Post-Training Factor #2: Disagreement with Negative Attitudes	1	0.012	0.180	0.004	0.948	0.004	1.012
Post-Training Factor #3: Not Afraid of Caring for Computers	1	-0.075	0.205	0.134	0.714	-0.025	0.928
Post-Training Factor #4: Lack of Concern Re Computer Operations	1	-0.216	0.197	1.204	0.273	-0.071	0.806
Post-Training Factor #5: Rs Like Interview; Easy to Do	1	-0.052	0.193	0.071	0.790	-0.019	0.950

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 11.1.7 Probability of Problems Setting Up Computer: Model 2

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	PR > Chi-Square	Standardized Estimates	Odds Ratio

Intercept	1	-4.442	1.077	17.007	0.000	.	.
Percent Urban Population in Segment	1	0.013	0.008	3.118	0.077	0.182	1.013
Age	1	0.023	0.016	2.063	0.151	0.100	1.023

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment

Exhibit 11.1.8 Probability of Computer Problems During Interview: Model 1

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	PR > Chi-Square	Standardized Estimates	Odds Ratio
Intercept	1	-2.901	0.400	52.523	0.000	.	.
Population Density	1	0.319	0.175	3.331	0.068	0.125	1.376
Hispanic % of Civilian, Noninstitutionalized Population ≥12 years	1	-0.011	0.007	2.976	0.085	-0.164	0.989
Computer-Assisted Interviewing Experience	1	-0.114	0.280	0.165	0.684	-0.031	0.892
Post-Training Factor #1: Positive Toward Computers	1	0.061	0.152	0.158	0.691	0.032	1.062
Post-Training Factor #2: Disagreement with Negative Attitudes	1	0.102	0.159	0.406	0.524	0.054	1.107
Post-Training Factor #3: Not Afraid of Caring for Computers	1	-0.143	0.143	0.993	0.319	-0.074	0.867
Post-Training Factor #4: Lack of Concern Re Computer Operations	1	0.033	0.140	0.056	0.813	0.017	1.034
Post-Training Factor #5: Rs Like Interview; Easy to Do	1	0.291	0.136	4.559	0.033	0.159	1.337

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

Exhibit 11.1.9 Probability of Computer Problems During Interview: Model 2

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	PR > Chi-Square	Standardized Estimates	Odds Ratio
Intercept	1	-4.773	0.866	30.399	0.000	.	.
Hispanic % of Civilian, Noninstitutionalized Population ≥12 years	1	-0.013	0.008	2.793	0.095	-0.197	0.987
Population Density	1	0.404	0.178	5.150	0.023	0.165	1.498
Age	1	0.031	0.014	4.581	0.032	0.210	1.031
Post-Training Factor #1: Positive	1	0.340	0.186	3.348	0.067	0.182	1.405

Toward Computers							
Post-Training Factor #5: Rs Like Interview; Easy to Do	1	0.377	0.142	7.086	0.008	0.217	1.458

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

11.2 Effect of Mode of Interview on Interviewer Variance

Interviewer variance or correlated response variance refer to an increase in the variance of estimates due to correlation of response errors within interviewers. It has often been assumed that this correlation arises because the "interviewers influence respondents to make response errors of a similar type" (Groves, 1991, p. 16). Well-developed self-administered questionnaires (SAQs) should be subject to less influence by the interviewers. This is because the interviewer would have less interaction with the respondent during the question-answering process and, thereby, less chance to influence his or her response errors. Thus, as we become more successful in developing questionnaires that respondents can complete without help from the interviewer, we should expect interviewer variance to decrease. ACASI does allow the respondents to operate more independently than the paper answer sheets. Other chapters of this report document the fact that ACASI increases reporting of sensitive behaviors and that respondents find it to be a more private response mode. We also examined how it affected interviewer variance.

Ideally, to assess the impact of mode of interview on correlated measurement errors, we would use a design in which subsamples of the sample segments were randomly assigned to interviewers and, within interviewers, randomly assigned to an interviewing mode (either ACASI or PAPI). The 1997 field experiment did not meet these requirements; however, because the sample segments within primary sampling units (PSUs) were randomly assigned to one of four quarters of data collection and because we were using segments from Quarters 1 through 3 for the 1997 field experiment, the assumption that segments were randomly assigned to mode of interview within PSUs is valid. Segments were not randomly assigned to interviewers.

For evaluating the impact of the interviewing method on the interviewer variance component, we fit a general linear mixed model with fixed and random effects of the form

$$Y = X\beta + Z\gamma + \varepsilon$$

where Y is an outcome variable, β represents fixed effects with respect to explanatory variables X, γ represents random effects with coefficients Z, and ε represents random error in each individual observation. For fixed effects, we used the following independent variables: (a) stratum, (b) gender, (c) race, and (d) age category. For the random effects, we considered the following components: (a) PSU and (b) field interviewer ID. We initially tried fitting a single model, which included mode of interview. Because of the fact that overall people gave different responses under ACASI and PAPI interviews, however, the individual or residual variance components for the same response variable were different for the two interview types. Therefore, the model was applied separately for ACASI and PAPI data. We examined several outcome variables using the SAS procedure MIXED. The results are displayed in Exhibit 11.2.1. For the mental health-related outcomes, individuals under the age of 17 years were dropped and the educationvariable was added to the model.

It is not possible to compare the interviewer components directly because the individual and PSU components are different for the two interview types. Thus, for each method, we computed the intraclass correlation to estimate the relative effect of interviewers on variance. This represents the proportion of the population variance that is accounted for by interviewers.

Of the 15 variables examined, for 14 of the variables correlation for PAPI was greater than the correlation for ACASI. For only one of the variables, *percentage who did not gain or lose weight when depressed*, the PAPI correlation was slightly smaller than ACASI, but both values were very close (0.09869 and 0.09207).

The results indicate that the ACASI method reduced the influence of interviewers on the variance. This is another positive impact of the ACASI methodology.

Because of the difficulties associated with this analysis and the fact that the design was not as well suited as it might be to investigate this issue, we recommend that this analysis be repeated with 1999 NHSDA split-sample data.

Exhibit 11.2.1 Variance Components and Interviewer Intraclass Correlation for the ACASI and PAPI Interviews for Selected Variables				
	Variance Components			
	Individual	**PSU**	**Interviewer**	**Rho**
ACASI	2247.89	29.293	1.089	0
PAPI	2021.09	6.333	55.252	0.02653
Ever used alcohol				
ACASI	2048.74	0.877	27.861	0.01341
PAPI	1587.99	15.024	94.959	0.05592
Ever used marijuana or hashish				
ACASI	1764.13	11.639	0	0
PAPI	1863.07	8.530	42.258	0.02208
Ever used hallucinogens				
ACASI	808.159	6.061	6.799	0.0083
PAPI	676.95	7.265	6.559	0.0095
Ever used inhalants				
ACASI	1020.15	0	15.415	0.01489
PAPI	484.96	3.228	17.242	0.03411
Ever felt sad or depressed for more than 2 weeks				
ACASI	1967.36	2.963	0.0	0
PAPI	1241.28	11.222	38.832	0.03007
Did not gain or lose weight when depressed				
ACASI	1896.63	0	207.677	0.09869
PAPI	1994.01	0	202.195	0.09207
Felt anxious for more than a month				
ACASI	1554.02	0	18.871	0.012
PAPI	910.86	0	40.743	0.04282
Number of days smoked in last 30 days				
ACASI	133.723	0	4.525	0.03273
PAPI	121.621	0.796	12.385	0.09188
Number of days had a drink in last 30 days				
ACASI	36.218	0.135	0.612	0.01656
PAPI	43.284	0	1.876	0.04154
Number of days used marijuana in last 30 days				
ACASI	89.921	0	0	0
PAPI	87.313	0	0.688	0.0078
Number of days used hallucinogens in last 30 days				
ACASI	0.455	6.469	0	0
PAPI	11.281	0	0	0
Number of days used inhalants in last 30 days				
ACASI	23.568	0	0	0
PAPI	2.389	11.015	14.729	0.523549

ACASI=audio computer-assisted self-interviewing; PAPI=paper-and-pencil interviewing; PSU=primary sampling unit.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1997 Field Experiment.

12. Refinement of Procedures: 1998 Laboratory and Field Testing and Final 1999 NHSDA CAI

The research conducted in 1996 and the 1997 field experiment and its subsequent analysis in early 1998 provided most of the information that was used to make decisions on the structure and content of the 1999 NHSDA. However, several issues remained and some new ones arose that required additional testing in the summer of 1998. Because the electronic screener application that was tested in the 1997 field experiment had not included the case management system (CMS) on the Newton, this component of the application needed to be tested under field conditions. This testing was done during the August 1998 field test and is described in Chapter 8.

The main reason that additional testing was needed was that the Department of Health and Human Services (DHHS) decided that a special youth tobacco module should be included in the 1999 NHSDA and that this module needed to be administered using audio computer-assisted self-interviewing (ACASI). This decision was made in the late spring of 1998 to address data requirements in proposed legislation related to tobacco use. Before this decision was made, SAMHSA had been planning to phase in the CAI implementation over a two-year period with the 1999 NHSDA conducted using PAPI methodology. The decision to include the new tobacco module in 1999 made it necessary to conduct the entire 1999 NHSDA using CAI. This required (a) developing and testing a new tobacco module that would provide the data required by the legislation and (b) making decisions as to the structure and content of the remainder of the CAI application. To investigate the former, we conducted a series of laboratory studies that focused on the ability of youths to complete a new tobacco module that asked about brand of tobacco use. To assist in decisions concerning the remainder of the application, timing tests of a draft of the instrument were conducted, and a version of the interview was field tested in August 1998.

12.1 Development and Testing of the Tobacco Module

The tobacco module was to be designed to provide more detailed information on the use of tobacco products by youths, including the brands that they used. Although the major interest was in the use of tobacco by youths, SAMHSA decided that the detailed tobacco module would be administered to all respondents and that it would become the tobacco component of the core NHSDA interview. Thus, the Research Triangle Institute (RTI) was instructed to develop this module as an ACASI interview and to do this very quickly. In addition, because the main concern was that the module be designed so that youths could complete it, the testing and development focused on youths rather than adults.

Several activities were undertaken to develop the module, including (a) a review of the suggested content and wording by an expert panel, (b) two rounds of cognitive laboratory testing, and (c) testing of the module in the August 1998 NHSDA field test.

12.2 Tobacco Module Development and Evaluation: Overview

12.2.1 Content of the Tobacco Module

To have time for testing, the content and basic approach of the module needed to be drafted very quickly. Based on a meeting between RTI and SAMHSA staff on June 8, 1998, an initial draft of the questionnaire was completed on June 10, 1998, and circulated for review. A conference call was held on June 16, 1998, and comments were gathered. Additional comments were gathered by e-mail, and the final changes for the initial round of laboratory testing were made on June 22, 1998.

The draft tobacco module included separate questions on cigarettes, chewing tobacco, snuff, cigars, and pipe tobacco. The cigarette section included questions on a variety of topics:

1. ever smoking in one's lifetime;
2. likelihood of initiating smoking for those who had never smoked;
3. smoking within the past 30 days;
4. recency of use for those who reported lifetime use but no use in the past 30 days;
5. days smoked in the past 30 days for 30-day smokers;
6. average number smoked per day on days smoked in the past 30 days;
7. brand of cigarette smoked most often in the past 30 days, which was presented as a pick-list;
8. open-ended question on brand for those who had used a brand not on the pick-list;
9. verification question on brand chosen;
10. use of roll-your-own-tobacco cigarettes;

11. daily smoking for 30 days or more; and
12. age when daily smoking began.

The cigarette questions were followed by an introduction that distinguished between chewing tobacco and snuff. This was followed by a series of questions on lifetime use of chewing tobacco, age at first use, use in the past 30 days, recency of use, days used in the past 30 days, brand used most often during the past 30 days, and verification ofthe reported brand. Similar questions were then asked for snuff and cigars. The module ended with two questions on using a pipe to smoke tobacco.

12.2.2 Expert Review of Tobacco Module Questions

A panel of seven independent consultants who are experts on measurement of cigarette use and selected substance use experts at RTI were asked to comment on the reliability and validity of the tobacco use questions and to make suggestions for improving the tobacco module. These expert opinions were documented in a memorandum that was submitted to SAMHSA and discussed by DHHS staff (including representatives from the DHHS Office of the Assistant Secretary for Planning and Evaluation [ASPE], the Centers for Disease Control and Prevention's [CDC's] Office on Smoking and Health [OSH], the CDC's National Center for Health Statistics [NCHS], the National Cancer Institute [NCI], and SAMHSA) in a meeting convened on July 2, 1998. At that meeting, the comments were discussed and decisions were made as to the wording that was to be used in the August testing.

The expert reviewers were generally positive about the questions. Although several reviewers suggested additional questions and editorial changes, no reviewer indicated that the proposed questions failed to meet current standards in terms of reliability and validity. The experts suggested additional questions that they thought would improve the utility of the NHSDA to provide information on the epidemiology of tobacco use by youths:

1. more detailed information on the initiation of smoking,
2. smoking different brands of tobacco products,
3. versions of brand smoked (e.g., Regular, Lite, Ultra Lite),
4. issues of youths' access to tobacco products, and
5. "blunting" (joint use of tobacco and other drugs, such as smoking marijuana in a cigar wrapper).

Because of the need to limit the overall length of the NHSDA interview, additional questions were not added for the 1999 NHSDA.

The performance of the tobacco module was evaluated by means of both laboratory and field testing. Two rounds of cognitive interviews were conducted in RTI's Laboratory for Survey Methods and Measurement. A number of cognitive interviews were also conducted outside of laboratory locations in order to interview hard-to-locate Spanish-speaking respondents. Field testing was done as part of the August 1998 field test.

12.2.3 Programming the Tobacco Module

Programming began on June 12, 1998, and WAV files for the laboratory version were recorded between June 18 and June 22, 1998. We tested the ACASI program from June 18 to 22, 1998, and the final version tested in Round 1 was fully programmed on June 24, 1998.

12.3 Tobacco Module: Round 1 Laboratory Testing (Cognitive Issues)

The new questions went beyond the previous core tobacco questions by asking respondents to identify and select the brands of tobacco products they used. Laboratory testing focused on the ability of respondents to use the computer to select the brand of cigarettes or other tobacco products that they used as well as their ability to answer the other tobacco use questions. The testing focused on the following issues:

1. Comprehension: Is the language of the questionnaire simple and easy to understand?
2. Usability: Do respondents understand how to use the computer to complete the questionnaire (entering/changing answers, backing up, etc.)?
3. Brand Identification: Can respondents consistently choose the brand of cigarettes/tobacco products that they use? Can they distinguish among brands? Do the questions enable respondents to describe their smoking preferences and choices? What is the best way to array lists of brands so that respondents can quickly identify the one they use?

4. Smokeless Tobacco: Can respondents distinguish between chewing tobacco and snuff? Do the terms and descriptions provided in the questionnaire adequately define smokeless tobacco products? Are brand identification issues the same for smokeless tobacco as for cigarettes?

Respondents for the first round of laboratory testing were recruited by posting flyers at various locations in the Triangle area of North Carolina, including Raleigh, Durham, Cary, and Chapel Hill. Flyers were posted at locations where teenagers congregate during summertime (parks, recreation centers, shopping malls, restaurants and well-known spots for "hanging out"). The flyers directed interested volunteers to call in to enlist as participants. Upon calling, the laboratory interview procedures were explained to potential respondents, preliminary interview times were discussed, and the names and phone numbers of parents were elicited. After parental permission was obtained in a separate phone call made by RTI staff, final scheduling of respondentsvisits was confirmed.

A laboratory interview protocol was designed to elicit information from respondents about key cognitive issues relating to understanding and usability. The protocol called for using think-aloud interviewing techniques. Specific probe questions designed to elicit information about the above topics were asked, and answers were noted by interviewers on a paper copy of the interview as the respondent used the computer. RTI's Institutional Review Board (IRB) reviewed the laboratory procedures, as well as respondent recruiting methods, to make sure that respondents were able to give their informed consent to participate and that parental permission was adequately obtained for interviewing minor respondents.

A total of 10 respondents participated-5 males ranging in age from 13 to 17 years old and 5 females ranging in age from 14 to 17 years old.

The results clearly indicated that respondents were able to identify the brands of cigarettes, cigars, and (in one case) smokeless tobacco using the brand identification pick-list that was programmed. Respondents selected the brand of tobacco product they used most often from an alphabetical list of brand names. Separate lists were used for each type of tobacco product (i.e., cigarettes, chewing tobacco, snuff, and cigars). Respondents all found the lists easy to use; the alphabetical order made it easy for them to quickly find the brand they were looking for. Respondents were also able to type in names of brands not listed.

The only difficulty that the respondents had selecting the brand they used most often arose not from the way in which the question was asked but from their uncertainty about which brand was used most. Youths who borrowed cigarettes from others were not always certain what brands they were given. The only problem identified with the listing was that respondents sometimes failed to see the "some other brand" option and, instead, chose the brand among those presented that they used most. However, when the option was pointed out, they were able to use it correctly. This was later revised to make the response option clearer.

Youths found the computer screen easy to read and had no difficulty with listening to audio files or using the keyboard. They had little difficulty answering questions about how many of the past 30 days they smoked and the average number of cigarettes they smoked on those days.

Other surveys had used the term "smokeless tobacco" to refer to both chewing tobacco and snuff. The respondents were queried about this term, and many of them expressed uncertainty or confusion when asked what the term meant. They either did not recognize the term or else considered it to apply to part of the general category. They had no difficulty with the definitions of chewing tobacco and snuff that were provided in the questionnaire.

12.4 Tobacco Module: Round 2 Laboratory Testing

A second round of cognitive interviews with teenagers was planned for August 1998. Round 2 testing was aimed at checking the effectiveness of several refinements in the wording, attempted to include more youths who had used smokeless tobacco, and included a Spanish version of the module.

12.4.1 Changes in Content of the Tobacco Module from Round 1

The tobacco module tested was similar to that used in the first round with a few changes:

1. Additional brands of cigarettes and cigars were presented, and there was a change in the way in which cigarette brand names were presented. Two alphabetical lists of brand names were presented instead of the one that was used in the prior round of testing. The second list of more rarely used brands was presented to respondents who did not identify

their most used brand in the first list. As was done in Round 1, respondents whose cigarette brand was not found on either the first or second list were asked to type in the brand name.

2. At the end of all brand name lists, the last choice was changed from "some other brand" (Round 1 wording) to "a brand not on this list."
3. An additional question about smoking 100 cigarettes in one's lifetime was asked.
4. Respondents who gave an age of initiation for any tobacco product that was the same or 1 year less than their present age were asked the year and month of first use.
5. The definition of snuff was presented a second time for respondents who answered questions about chewing tobacco, prior to asking about snuff use. For respondents who did not answer chewing tobacco questions, the definition of snuff was not repeated.
6. A Spanish-language version was developed.

12.4.2 Round 2 Research Design

Cognitive issues for Round 2. We examined several questions during Round 2:

1. Was the modified method of presenting cigarette brands usable? The modified method involved using two lists of cigarettes, the second one shown if the target brand was not on the first list, followed by a provision to type in any unlisted brand using the keyboard.
2. Was the redefinition of snuff after answering questions about chewing tobacco helpful?
3. Was the language of the tobacco module questions understandable? Specifically, how the phrase "all or part of a cigarette" was interpreted received specific attention, as did the phrase "on the days that you smoke." Although not part of the tobacco module wording, respondents were also asked how they would interpret a question reading "In the past 30 days, have you smoked any cigarettes?"
4. Could respondents read and understand the computer screen, keyboard, and sound equipment?

Round 2 data collection procedures. The data collection plan for Round 2's laboratory interviews called for interviewing 50 youths: 40 in English and 10 in Spanish. Of the 40 English interviews, 10 were to use smokeless tobacco (in the past 30 days, if possible, but recently enough to be able to identify the brand they used) and 10 respondents who were not tobacco users. Additionally, half were to be with youths 12 to 14 years old and half 15 to 17 years of age.

Recruiting and cognitive interview procedures for Round 2 were the same as those for Round 1, except for the Spanish interviews. These interviews took place either in respondents' homes or at a local Hispanic cultural fair. Although the original research plan called for interviewing only teenagers (12 to 17 year olds), several adults were interviewed in order to expedite the completion of enough interviews to make recommendations. Furthermore, only 4 of the 10 interviews were done using the ACASI computerized interviewing method. Because the remaining interviews were conducted at a fair, it was inconvenient to use the computer in that setting. However, because the primary interest in the Spanish version of the questionnaire involved the wording of the questions, rather than use of the computer, we felt that valuable information could still be gained without the computer. These remaining respondents were interviewed using a translated version of the latest tobacco module.

Recruiting of youth respondents for Round 2 proved more difficult than anticipated and more difficult than for the Round 1 testing. It may have been that the beginning of the school year, in the middle of August, made students unsure of their schedules and unwilling to devote their free time to something that sounded like work. Our original plan was to interview 50 youths, but we only interviewed about two thirds of that number.

12.4.3 Round 2 Testing Results

A total of 36 respondents who met the criteria of age and parental consent were interviewed. English- language respondents were 8 females and 18 males ranging in age from 13 to 17. Three respondents (all males) between the ages of 12 and 14 were interviewed, as well as five 15 year olds. An additional 10 respondents completed a Spanish-language version of the cognitive interview. The age range for Spanish interviews was from 12 to 57, with 8 of the 10 interviews being conducted with respondents in the target age range.

Brand identification. The results clearly indicated that respondents were able to identify the brands of cigarettes, cigars, and smokeless tobacco using the brand identification questions programmed for the youth tobacco module. The English and Spanish versions of these questions performed equally well. These questions were asked by having respondents select the brand of tobacco product they used most often from one of two alphabetical lists of brand names of cigarettes, and single lists for chewing

tobacco, snuff and cigars (separate lists for each type of tobacco product). Brand names for the Spanish instrument were not translated, although the ACASI recordings were made using Spanish-accented pronunciation.

Respondents all found the lists easy to use. The alphabetical order made it easy for them to quickly find the brand they were looking for. Respondents were also able to type in names of brands not listed. Respondents who did not smoke cigarettes, or who had only smoked one or two, were as easily able to locate specific brand names as were regular smokers. Again, the only difficulty respondents had selecting a brand stemmed from the uncertainty about which brand was used most among the youths who borrowed cigarettes from others.

Those youths who had experience with smokeless tobacco found identifying brands of those products just as easy as identification of cigarettes and cigars.

Wording and terminology.

1. The change of "some other brand" to "a brand not on this list" seemed to make the intention of that response option clearer.
2. The wording "all or part of a cigarette" was almost uniformlyinterpreted as referring to any level of smoking beyond passive second-hand smoke. However, the wording "... smoke any cigarettes" was interpreted by a substantial minority as not referring to smoking only part of a cigarette.
3. The restatement of the definition of snuff drew mixed reviews. Although some respondents were not bothered by it, a couple said it was too long and the restatement might be shortened and emphasize only the differences from chewing tobacco.
4. The question of smoking 100 cigarettes in a lifetime was easy to answer by our respondents who had either smoked for a long time or else were non-smokers, or youths who had only smoked one or two times.

Several of the Spanish translations of tobacco-related terms were not well understood.

1. In Spanish, *cigarro* can mean cigar or cigarette (also called *cigarillo*, which in English is a form of small cigar). Central and South Americans preferred the term *puro* to *cigarro* for a cigar. *Puro* means "pure," referring to the fact that cigars are made completely out of tobacco, rather than tobacco wrapped in paper, like cigarettes. Puerto Ricans preferred *cigarros* to *puros*. Clearly, both terms (along with appropriate definitions) are needed in referring to cigars and cigarettes in the general population.
2. The concepts of *plug* and *twist* of chewing tobacco (translated literally as *tapado* and *atado* in the Spanish translation) were not understood either by younger respondents or adult smokers (some of whom had experience with chewing tobacco). A suggested solution would be to leave the English words in quotation marks and simply describe, in Spanish, the look of a plug/twist. An alternative would be to provide a picture of these forms of chewing tobacco along with the definitions.

Use of the computer. Respondents had no difficulty using the computer either for categorical answers or for typing in numerical or text answers. Respondents also had no difficulty providing the year and month of first use of tobacco products.

12.5 Tobacco Module: Discussion of Laboratory Results

The results of the laboratory testing clearly indicated that youths were able to answer questions about their own tobacco use using the tobacco module. Most could identify all the brands they had ever smoked and even brands that they never smoked (brands used by friends or brands seen in ads). This finding was clearly evident even though the target number of respondents was not reached.

Identifying smokeless tobacco brands was no different than for cigarette or cigar brands. However, respondents were not always correct in identifying a smokeless tobacco product as chewing tobacco vs. snuff. One respondent thought *Kodiak* was a brand of chewing tobacco and entered it by typing in the chewing tobacco section. However, when he completed the snuff section, he found it again. That made it somewhat confusing to answer the final smokeless tobacco question about which was used more: the brand of chewing tobacco or the brand of snuff. Although this finding might suggest that presenting both chewing tobacco and snuff brands together should be considered, it does not indicate that the alphabetical listing of brand names was difficult for respondents to use.

Based on these findings, it was recommended that the tobacco module be used in more or less the same form as it was tested. The findings from the cognitive laboratory testing regarding specific issues suggested the following:

1. The use of alphabetical lists of brand names should be retained. There was no evidence that any other means of presenting brand name information, such as pictures, was necessary.
2. The use of "... smoke all or part of a cigarette" wording should be retained. Some respondents did not think the "...smoke any cigarettes" wording referred to smoking part of a cigarette.
3. The other wording changes introduced in Round 2 worked well enough, although the second definition of snuff could be shortened. They also suggested that "a brand not on this list" was an improvement over the "some other brand" wording used in Round 1.
4. The computer screen was easy to read, and youths had no difficulty with listening to audio files or using the keyboard. They had little difficulty answering questions about how many cigarettes they smoked over a variety of time periods.

12.6 Tobacco Module: August 1998 Field Test Results

Data on tobacco use from 177 respondents who completed the August field test are available (see Exhibit 12.6.1). There were 91 youths (12 to 17 years old) and 86 adults (18+ years old). Among the youths, there were 53 males and 38 females, while among the adults, 35 were men and 51 were women. The following highlights the results for tobacco use.

1. Among the 62 youths who had never smoked, 53 (85%) said that they would not smoke if one of their best friends offered them a cigarette; 50 (81%) said that they would definitely not smoke a cigarette in the next 12 months. (Adults were not asked these questions.)
2. Among the 29 youths who had ever smoked a cigarette, the youngest age at which they first smoked was 5 years and the oldest was 16. Fifteen (52%) were 12 or under when they first smoked a cigarette. Six entered an age of first use of 10 or below, and this triggered a consistency check that they were all able to complete. Age of initiation for the 58 adult lifetime smokers ranged from 7 to 22 years, and 12 (21%) were 12 or younger when they first smoked a cigarette.
3. Of the 12 youths who had smoked within the past 30 days, 10 named a brand on the list, one did not know the brand, and one person typed in Black and Mild (a brand of cigar). Of the 31 adult 30-day smokers, 30 smoked a named brand and one typed in a brand name (Pyramid).
4. Among the 12 youths who had smoked in the past 30 days, three gave "don't know" answers to the questions on the number of days smoked, one gave a "don't know" answer to the number smoked on the days that they smoked, and one did not know the brand. These may have been youths who smoked very little and obtained the cigarettes that they smoked from others.
5. One youth reported smoking a roll-your-own brand as did six adults.
6. A total of 3 youths and 12 adults reported smoking cigarettes every day for 30 days at some time in their life.

A series of respondent and interviewer debriefing questions was also included in the August 1998 field test. Of note is the following:

1. Among past 30-day smokers, 10 youths and all 31 of the adults reported that they had no difficulty answering the questions on the brand of cigarettes that they smoked most often in the past 30 days. One youth reported some difficulty, and the other youth did not get to this section.
2. Only 2 of the 91 youths were rated by the interviewer as having some difficulty answering the tobacco questions on the computer. One did not understand the distinction between chewing tobacco and snuff, and one could not remember the information asked for. Of the 86 adults, 8 reported having at least some difficulty answering the tobacco questions on the computer: 2 had difficulty understanding some of the questions, 1 had difficulty remembering the required information, 1 had trouble responding appropriately and 4 had some other difficulty.
3. Two youths and one adult made comments on tobacco legislation.

Overall, the field test respondents appeared to have little procedural difficulties answering the questions in the tobacco module. However, a total of 20 "don't know" or "refused" answers were entered. These were most frequently entered for questions asking about the number of cigarettes (five "don't know" answers) and ages when first used tobacco products (four total). Thus, the difficulties that respondents had appeared to be more associated with the inability to recall answers rather than an inability to complete the procedures required to answer the questions. Respondents provided responses on frequency of tobacco use that were comparable to previous NHSDA data, although the small sample size makes any statistical comparisons unstable. The tobacco module appeared to perform well in the field, bolstering the conclusions from the laboratory testing that the questions were adequate for the purpose for which they were designed. Certainly, the brand use questions, the major change from the prior tobacco modules, performed well.

Exhibit 12.6.1 Reported Use of Tobacco Products from August 1998 Field Test

Tobacco Product Usage	Youth	Adults
Recency of Smoking Cigarettes (Frequencies)		
Within past 30 days	12	29
More than 30 days, but less than 12 months ago	5	5
More than 12 months, less than 3 years ago	9	8
More than 3 years ago	3	13
Never	62	29
Cigarette Smoking Frequency (Days Smoked in Past 30 Days)		
1 - 10 days	6	7
11 - 20 days	2	1
21 - 30 days	1	21
Number of Cigarettes Smoked (Past 30 Days)		
<1 cigarette	6	2
1 - 5 cigarettes	5	8
6 - 15 cigarettes	0	5
16 - 25 cigarettes	0	13
26 - or more cigarettes	0	2
Recency of Chewing Tobacco Use		
Within past 30 days	0	3
More than 30 days, but less than 12 months ago	1	1
More than 12 months, less than 3 years ago	1	2
More than 3 years ago	3	5
Never	86	73
Recency of Snuff Use		
Within past 30 days	0	3
More than 30 days, but less than 12 months ago	0	0
More than 12 months, less than 3 years ago	3	0
More than 3 years ago	2	3
Never	86	80
Recency of Smoking Cigars		
Within past 30 days	6	5
More than 30 days, but less than 12 months ago	5	9
More than 12 months, less than 3 years ago	2	9
More than 3 years ago	2	7
Never	75	56
Frequency of Smoking Cigars (Days Smoked in Past 30 Days)		

1 - 10 days	5	5
11 - 20 days	1	0

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1998 Field Test.

12.7 Additional Field and Laboratory Testing

Other components of the 1999 CAI application were tested in the summer of 1998.

12.7.1 Field Testing of the Full Instrument

The August 1998 field test was initially designed as the final test of the software and procedures for using the Apple Newton 2100 for screening households and case management. However, when it was decided to implement the CAI in 1999 to accommodate the addition of the expanded tobacco module, the testing plan was revised to allow testing of a CAI application that included this new module and a draft CAI application for other sections of the interview. The original plan for field testing called for employing 4 field supervisors (FSs) to direct a total of 20 field interviewers (FIs) to collect data in areas that cut across all four U.S. time zones and regions. Subsequent to the decision to include testing of the full CAI application including the new tobacco module, a total of 8 FSs were selected and the field force was expanded to include 40 FIs. The sample design was modified to yield an expected total of 150 CAPI/ACASI interviews. Of those interviews, half were to be with 12 to 17 year olds and half with respondents 18 years old or older.

In the following paragraphs, we describe the findings for the CAI interview. Newton results are discussed in Chapter 8. Note that an earlier version of the Apple Newton was used for the 1997 field experiment.

Response rates. Overall, 836 households were screened; 683 (81.7%) were eligible; within those households, 1,313 individuals (2.06 persons per household) were listed using the Newton application. From these eligible individuals, 256 sampled persons were selected, and 183 (71.5%) of them completed interviews. A higher response rate could have been achieved; however, interviewing was ended early because yields were higher than expected. Exhibit 12.7.1 displays the age distribution of the people who were eligible and sampled, as well as those who completed interviews.

Hardware and software performance. The computer support system at RTI logged problems that were handled during the course of the August 1998 field test.

Problems with the laptop application, which used a Gateway computer, included e-mail communication (14 times), data transmission (13 times), and closing the ACASI program (14 times):

1. E-mail communication problems were mainly due to phone number problems or problems entering the correct software commands.
2. Most of the transmission problems were the result of FI errors (usually problems with passwords).
3. Problems closing the ACASI program were generally solved with a couple of keystrokes or by restarting the interview and jumping to the end.

Interview data analysis. Due to the small sample size of the field test, little could learned about the population distribution of behaviors that are as rare as substance use. Thus, we examined operational aspects of the survey. There were no unusual patterns of "don't know" and "refusals" (such responses were quite rare), and the only questions where "other-specify" was an answer were the race/ethnicity questions, for which a total of 10 answers were specified by respondents.

The completion time data for the ACASI interview were analyzed for the sample as a whole and separately for youths (aged 12 to 17) and adults (18 or older). The results of these analyses are displayed in Exhibit 12.7.2. Overall, the entire interview took approximately 47 minutes, averaged across all field test respondents. The introductory sections averaged about 8 minutes, the core ACASI substance use section took about 14 minutes, the non-core ACASI section took about 15.5 minutes, and the final interviewer administered section took about 9.5 minutes. Youths averaged about 5 minutes more than adults. The difference was in the introductory section (a little over 1 minute more for youths) and the non-core ACASI section (more than 5 minutes more, which was mainly due to the Youth Experience module). The core and the final interviewer-administered sections took slightly

less time for youths. The maximum total time was nearly 85 minutes, and the minimum total time was 16 minutes. Three fourths of the interviews were conducted in less than 55 minutes.

Raw, unweighted frequencies of drug use were examined to make sure that there were no surprises in this new application. In general, these showed no differences from prior implementations of the NHSDA.

Interviewer debriefing analysis. During the first 2 weeks of August 1998, conference calls were held with the FIs, and a final call was held during the first week of September. Five calls were held each week with groups of interviewers attending each call. The purpose of these calls was to discuss experiences and problems with data collection.

At the outset of the weekly calls, FIs spontaneously voiced their concerns about the computerized data collection procedures. Concerns about the use of the laptop computer for the CAI interview included the following:

1. worry that the laptop computer would overheat and be uncomfortable to use if operated in the respondent's lap;
2. reluctance of respondents to use the headphones (respondents' concerns about cleanliness or refusal to use them);
3. disliking passing the computer back and forth more than once during transitions between CAPI and ACASI portions of the interview;
4. complaints that the tutorial is too slow, is insulting to respondents, or other negative reactions to the tutorial;
5. questions as to the necessity of having separate pill cards and other show cards instead of having computerized presentation of that information; and
6. various software problems with the laptop (mainly entering final codes).

There were also many cases of problems in transmission between the laptop computer and the mainframe computer at RTI.

Interviewers suggested shortening the tutorial or making it optional, getting rid of calendar and pill cards, simplifying questionnaire language, putting all CAPI questions together, and moving certain additional questions to the ACASI section. They also suggested changes to the way the computers are programmed to transmit data and putting games on the Newtons to amuse FIs during ACASI sections of the interview. The testing was helpful in identifying changes in procedures, particularly transmission of data, and training that addressed FI concerns.

12.7.2 Laboratory Testing of Full Instrument

The full instrument was also tested in the laboratory during July and August 1998. This was necessary because the 1997 field experiment had used eight different versions of the core drug use questions, and testing of a version close to that planned for the 1999 survey was needed. The respondents for the laboratory full interview testing were a convenience sample, divided between clients at drug treatment centers and RTI employees and their acquaintances. A total of 53 people participated in the laboratory testing of the full ACASI instrument. Although some of the interviews were conducted in RTI's Laboratory for Survey Methods and Measurement, the majority were conducted by having RTI staff meet participants at convenient places (either at treatment centers for treatment center clients or in the respondent's home).

The laboratory testing of the full instrument was conducted primarily for the purpose of investigating procedural issues (difficulties with question wording, time to completion) rather than for estimating the prevalence of substance use. Because aconvenience sample was used, we do not report levels of substance use but rather report only the data the interviews were designed to study.

Time to completion. Time to completion was measured in two ways: (a) by calculating completion time from the time stamps that the computer was programmed to record, and (b) by having the interviewer note the beginning and end times of the interview. The first way is more precise and the only way to segment the total time into completion times for different sections of the interview. However, the laboratory version of the full interviews contained modifications in the instrument that were not present in the field test instrument. There was limited time to complete testing of the computer program before beginning the laboratory interviews. Programmers were not certain that the time-recording function was successfully programmed. Therefore, the second method of timing was also used as a backup for possible failures of the programmed timing, even though it yielded only relatively rough estimates of the length of the entire interview.

Laboratory interview sampling was done in a way that would tend to inflate the average interview completion time. Half of the respondents were clients at substance use treatment centers. These respondents would be expected to take longer because they would be more likely to complete sections on at least some substances were bypassed by nontreatment respondents, as well as sections on substance use treatment. These respondents were recruited in order to maximize the likelihood that all sections of the

instrument would be completed by at least some respondents. The small size of the field test made it improbable that any drug users would be interviewed. This deliberate heterogeneity made it difficult to generalize the time to completion of the full interview to other populations.

Upon retrieval of the laptop computers used in the interviewing, it was found that there were timing data for only 38 of the 53 participants. Exhibit 12.7.3 displays the time to completion results for these respondents. It is not clear whether the failure to record the times for all interviews was due to equipment malfunctioning, to problems with programming, or to procedural irregularities by interviewers.

Completion times for the laboratory interviews were longer than for the field test[19] (about 47 minutes for field test interviews vs. 64 minutes for laboratory interviews). Some of the difference can be attributed to the presence of treatment center respondents (whose interviews averaged over 70 minutes). However, the completion times for the nontreatment laboratory respondents were more than 15 minutes longer than the comparable field interview sample: respondents 18 years of age or older (all nontreatment laboratory interviews with timing information were from that age group). Most of the difference was in the initial CAPI questioning (approximately 7 minutes) and the non-core ACASI section (approximately 6 minutes, despite the absence of parenting questions answered by the nontreatment laboratory interview respondents). It may be that the interviewers for the laboratory interviews, who did not receive formal training on the use of the laptop computer, may have been less adept at using it, making the initial section longer. Furthermore, respondents in laboratory interviews were encouraged to discuss any problems they may have experienced, which would also serve to extend the duration of the interviews.

Specific issues from planned debriefing interviews. During the debriefing of the laboratory respondents, we asked about comprehension of specific questions or instructions in the interview and about selected procedural issues. None of the answers suggested that there were any problems with understanding the specific terms that were investigated ("nonmedical drug use," "treatment"). Results regarding procedural issues were as follows:

1. Length of the interview: All nontreatment respondents, and a majority (15 of 19) of the treatment clients, did not consider the interview to be overly long. The four clients who did consider it to be too long thought 30 to 60 minutes would be a reasonable length.
2. Adequacy of the tutorial: The tutorial was not perceived as overlong, and it seemed adequate at instructing nontreatment respondents how to use the computer. Some problems (pushing the wrong keys or inability to go back and change answers) did occur infrequently. Treatment clients found the computer tutorial an adequate introduction to using the computer, although a couple thought it could be shorter. Some respondents did not remember how to do some of the functions addressed in the tutorial, so it would be helpful to have labels for the function keys to serve as memory aids.
3. Usefulness of the calendar: For nontreatment respondents and treatment clients, reactions to the calendar were mixed. About 33% used the calendar and found it useful, the remainder did not consider it useful and did not use it. Other criticisms of the calendar mainly involved seeing it as too small and hard to read.

Other difficulties observed or mentioned by respondents. There were a few suggestions that the vocabulary for the survey may still be at a relatively high level in a few places. However the ACASI format seemed to make the questions easy enough to answer so that no respondents asked for clarification during the interviews. Most of the other difficulties mentioned by respondents related to their unique circumstances rather than the shortcomings of the instrument (e.g., a respondent who lived in a shelter, who answered that he lived with 65 other people, had this answer rejected as being out of range by the computer).

Exhibit 12.7.1 Number of Persons Listed, Selected, and Responding, by Age

Status/Age in Years	Sample Size	Percent
Eligible Persons Listed in Eligible Household	**1,313**	100.0
12-17	147	11.2
18-25	165	12.6
26-34	218	16.6
35-49	366	27.9
50+	416	31.7
Persons Selected (percent = % of total eligibles by age group)	**256**	**19.5**
12-17	129	87.8

18-25	69	41.8
26-34	25	11.5
35-49	17	4.6
50+	16	3.8
Completed Interviews (percent = % of sampled persons by age group)	**183**	**71.5**
12-17	93	72.1
18-25	47	68.1
26-34	14	56.0
35-49	15	88.2
50+	13	81.2

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1998 Field Test

Exhibit 12.7.2 Field Test: Interview Completion Time (Minutes)

Module/Answer Sheet	n	Min	Q1	Median	Q3	Max	Mean
Total Sample							
Total Interview	**170**	**16.02**	**36.90**	**45.83**	**54.68**	**84.68**	**46.93**
Pre-Answer Sheet	**170**	**1.65**	**5.18**	**7.05**	**9.88**	**26.85**	**7.92**
Total Core	**170**	**2.21**	**9.10**	**12.89**	**17.37**	**33.86**	**13.85**
Tobacco	170	0.37	1.77	2.81	4.03	14.53	3.08
Alcohol	170	0.03	0.37	1.08	2.52	8.75	1.77
Marijuana	170	0.03	0.10	0.18	0.35	3.33	0.35
Cocaine	170	0.03	0.07	0.10	0.17	1.57	0.16
"Crack"	11	0.05	0.07	0.08	0.15	1.35	0.21
Heroin	170	0.05	0.17	0.27	0.55	2.32	0.40
Hallucinogen	170	0.18	0.60	0.96	1.47	3.72	1.14
Inhalants	170	0.28	1.25	1.98	2.83	5.83	2.14
Analgesics	170	0.23	0.80	1.36	2.13	4.78	1.56
Tranquilizers	170	0.17	0.55	0.91	1.52	4.12	1.14
Stimulants	170	0.25	0.50	0.87	1.53	3.68	1.11
Sedatives[1]	170	0.22	0.55	0.81	1.26	3.34	0.98
Non-Core ACASI	**170**	**4.39**	**10.75**	**14.28**	**19.30**	**34.21**	**15.48**
Special Drugs	170	0.08	0.21	0.31	0.49	1.28	0.37
Risk Assessment	170	1.02	2.32	2.79	3.98	12.03	3.37
Drug Experience	93	0.25	0.75	1.22	1.95	6.20	1.56
Special Topics	170	0.13	0.35	0.52	0.85	4.90	0.70
Treatment	169	0.08	0.15	0.22	0.37	7.73	0.37
Youth Experience	84	3.58	9.97	12.02	15.75	24.33	12.73
Social Environment	86	2.45	4.00	5.32	7.40	18.02	6.29
Parenting	6	1.63	2.18	2.69	3.77	3.88	2.81
Interviewer-Administered	**170**	**2.75**	**6.98**	**9.15**	**11.80**	**27.22**	**9.68**
12 to 17 Sample							

Total Interview	**84**	**21.40**	**39.87**	**47.23**	**56.91**	**81.17**	**49.50**
Pre-Answer Sheet	**84**	**1.65**	**5.46**	**8.08**	**10.77**	**26.85**	**8.47**
Total Core	**84**	**2.21**	**9.55**	**12.89**	**16.66**	**31.74**	**13.65**
Tobacco	84	0.47	1.88	2.90	4.03	6.58	2.99
Alcohol	84	0.03	0.20	0.58	1.39	6.07	1.17
Marijuana	84	0.03	0.11	0.18	0.28	1.75	0.27
Cocaine	84	0.03	0.07	0.10	0.15	0.55	0.13
"Crack"	0	.	.	.	.	.	.
Heroin	84	0.05	0.17	0.30	0.69	2.32	0.47
Hallucinogen	84	0.18	0.68	1.04	1.65	3.72	1.27
Inhalants	84	0.28	1.33	2.08	3.18	5.83	2.23
Analgesics	84	0.30	0.82	1.47	2.42	4.78	1.75
Tranquilizers	84	0.20	0.58	1.01	1.66	4.12	1.22
Stimulants	84	0.25	0.46	0.92	1.58	3.42	1.18
Sedatives	84	0.22	0.58	0.82	1.26	2.90	0.97
Non-Core ACASI	**84**	**5.78**	**14.02**	**17.96**	**20.60**	**34.21**	**18.12**
Special Drugs	84	0.08	0.22	0.31	0.49	1.11	0.37
Risk Assessment	84	1.03	2.33	2.95	4.26	10.10	3.41
Drug Experience	25	0.25	0.75	1.12	1.62	6.20	1.49
Special Topics	84	0.13	0.34	0.45	0.76	4.90	0.64
Treatment	84	0.08	0.17	0.25	0.40	0.93	0.30
Youth Experience	84	3.58	9.97	12.02	15.75	24.33	12.73
Social Environment	0	.	.	.	.	.	.
Parenting	0	.	.	.	.	.	.
Interviewer-Administered	**84**	**3.42**	**6.06**	**7.95**	**11.43**	**27.22**	**9.26**
18+ Sample							
Total Interview	**86**	**16.02**	**33.82**	**44.23**	**52.05**	**84.68**	**44.41**
Pre-Answer Sheet	**86**	**1.92**	**5.03**	**6.81**	**9.10**	**20.92**	**7.38**
Total Core	**86**	**4.24**	**8.66**	**12.95**	**17.85**	**33.86**	**14.04**
Tobacco	86	0.37	1.73	2.72	4.25	14.53	3.17
Alcohol	86	0.07	0.90	1.93	3.23	8.75	2.36
Marijuana	86	0.05	0.10	0.21	0.62	3.33	0.43
Cocaine	86	0.03	0.07	0.11	0.18	1.57	0.20
"Crack"	11	0.05	0.07	0.08	0.15	1.35	0.21
Heroin	86	0.07	0.15	0.22	0.40	2.07	0.33
Hallucinogen	86	0.22	0.53	0.78	1.32	3.00	1.02
Inhalants	86	0.47	1.13	1.92	2.62	5.20	2.06
Analgesics	86	0.23	0.78	1.21	1.80	3.87	1.37
Tranquilizers	86	0.17	0.53	0.81	1.42	4.10	1.05
Stimulants	86	0.25	0.52	0.82	1.52	3.68	1.05
Sedatives	86	0.24	0.54	0.79	1.26	3.34	0.98
Non-Core ACASI	**86**	**4.39**	**8.58**	**11.16**	**15.50**	**28.37**	**12.90**
Special Drugs	86	0.09	0.21	0.30	0.49	1.28	0.38

Risk Assessment	86	1.02	2.17	2.78	3.77	12.03	3.34
Drug Experience	68	0.35	0.74	1.28	1.98	5.83	1.59
Special Topics	86	0.13	0.40	0.63	1.00	3.60	0.77
Treatment	85	0.08	0.13	0.18	0.30	7.73	0.43
Youth Experience	0	.	.	.	.	.	.
Social Environment	86	2.45	4.00	5.32	7.40	18.02	6.29
Parenting	6	1.63	2.18	2.69	3.77	3.88	2.81
Interviewer-Administered	**86**	**2.75**	**7.73**	**9.87**	**11.93**	**27.10**	**10.10**

[1]There was no time stamp between the sedative and special drugs sections, so the recorded time included both sections. Times for each section were interpolated by splitting the total time into the proportions of the 1997 Field Test #2, Treatment Version3, for the two sections.

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1998 Field Test.

Exhibit 12.7.3 Full Laboratory Interviews: Interview Completion Time (Minutes)

Module/Answer Sheet	n	MIN	Q1	MEDIAN	Q3	MAX	MEAN
Total Sample							
Total Interview	**39**	**32.1**	**46.3**	**60.3**	**78.1**	**121.9**	**64.1**
Pre-Answer Sheet	**39**	**3.3**	**9.1**	**11.3**	**15.0**	**40.4**	**13.38**
Total Core	**39**	**7.9**	**12.0**	**16.5**	**20.6**	**43.4**	**17.98**
Tobacco	39	0.8	2.0	3.6	5.6	9.1	4.10
Alcohol	39	0.3	1.6	2.1	3.1	8.3	2.45
Marijuana	39	0.1	0.3	0.5	1.0	3.2	0.72
Cocaine	39	0.0	0.1	0.3	1.1	4.7	0.81
"Crack"	23	0.01	0.2	0.8	1.1	2.1	0.77
Heroin	39	0.1	0.2	0.4	0.7	2.9	0.59
Hallucinogen	39	0.4	0.7	0.9	1.4	3.9	1.17
Inhalants	39	0.9	1.4	1.8	2.6	3.6	1.95
Analgesics	39	0.6	1.0	1.4	3.0	5.6	1.98
Tranquilizers	39	0.5	0.9	1.3	2.2	5.8	1.63
Stimulants	39	0.5	0.7	1.00	1.3	3.2	1.16
Sedatives	39	0.4	0.6	0.8	1.1	2.7	0.96
Non-Core ACASI	**39**	**7.1**	**11.7**	**18.1**	**28.0**	**60.2**	**21.53**
Special Drugs	39	0.2	0.2	0.3	1.0	4.6	0.81
Risk Assessment	39	1.6	2.2	2.9	4.0	14.8	3.57
Drug Experience	39	1.0	1.8	3.7	5.8	20.2	4.99
Special Topics	39	0.0	0.5	0.8	1.7	4.8	1.20
Treatment	39	0.0	0.4	1.8	5.0	17.0	3.34
Social Environment	38	0.0	5.0	7.7	9.1	12.8	7.18
Interviewer-Administered	**39**	**5.5**	**8.6**	**11.7**	**14.0**	**19.5**	**11.24**
Treatment Center Sample							

Total Interview	**12**	**34.3**	**48.9**	**65.4**	**89.4**	**121.9**	**70.44**
Pre-Answer Sheet	**12**	**5.2**	**7.5**	**10.2**	**12.1**	**29.3**	**11.20**
Total Core	**12**	**8.3**	**14.1**	**19.6**	**27.7**	**43.4**	**21.24**
Tobacco	12	2.0	3.3	4.9	7.4	9.0	5.22
Alcohol	12	0.3	1.3	1.6	3.4	8.3	2.59
Marijuana	12	0.1	0.3	0.6	0.9	3.2	0.78
Cocaine	12	0.2	0.3	0.7	1.9	2.8	1.05
"Crack"	10	0.2	0.6	0.9	1.2	2.1	0.99
Heroin	12	0.1	0.3	0.6	1.6	2.9	0.96
Hallucinogen	12	0.4	0.8	1.1	1.7	3.4	1.32
Inhalants	12	0.9	1.5	2.0	2.7	3.6	2.06
Analgesics	12	0.8	1.0	1.8	3.3	5.6	2.25
Tranquilizers	12	0.5	0.7	1.4	2.6	4.4	1.73
Stimulants	12	0.6	0.8	1.0	2.7	3.2	1.35
Sedatives	12	0.5	0.7	0.9	1.4	2.7	1.12
Non-Core ACASI	**12**	**11.0**	**14.2**	**25.9**	**34.5**	**60.2**	**27.61**
Special Drugs	12	0.2	0.3	0.7	2.3	4.6	1.34
Risk Assessment	12	1.6	2.2	3.4	4.0	8.8	3.81
Drug Experience	12	1.4	3.7	4.8	7.3	20.2	6.71
Special Topics	12	0.3	0.7	1.6	2.1	4.2	1.66
Treatment	12	1.7	2.3	4.1	9.8	17.0	6.29
Social Environment	12	3.4	4.8	8.4	9.2	12.8	7.49
Interviewer-Administered	**12**	**6.7**	**7.6**	**9.2**	**12.9**	**17.9**	**10.40**
Nontreatment Sample							
Total Interview	**27**	**32.1**	**45.6**	**60.3**	**74.1**	**110.8**	**61.33**
Pre Answer Sheet	**27**	**3.3**	**9.3**	**12.4**	**15.4**	**40.4**	**14.35**
Total Core	**27**	**7.9**	**11.3**	**16.5**	**20.1**	**34.2**	**16.54**
Tobacco	27	0.8	1.8	3.4	4.4	9.1	3.61
Alcohol	27	0.7	1.8	2.2	3.0	5.4	2.39
Marijuana	27	0.1	0.3	0.4	1.0	3.0	0.70
Cocaine	27	0.0	0.1	0.2	0.9	4.7	0.71
"Crack"	13	0.0	0.1	0.5	1.1	1.4	0.60
Heroin	27	0.1	0.2	0.4	0.5	1.4	0.43
Hallucinogen	27	0.4	0.6	0.9	1.4	3.9	1.10
Inhalants	27	1.0	1.4	1.7	2.4	2.9	1.90
Analgesics	27	0.6	1.1	1.4	2.6	4.2	1.86
Tranquilizers	27	0.6	0.9	1.2	1.8	5.8	1.59
Stimulants	27	0.5	0.7	1.0	1.3	2.6	1.07
Sedatives	27	0.4	0.6	0.8	1.1	2.2	0.89
Non-Core ACASI	**27**	**7.1**	**11.6**	**15.2**	**25.2**	**55.9**	**18.82**
Special Drugs	27	0.1	0.2	0.3	0.5	2.4	0.57
Risk Assessment	27	1.7	2.2	2.9	3.8	14.8	3.46
Drug Experience	27	1.0	1.6	3.1	4.4	14.6	4.22

Special Topics	27	0.0	0.5	0.6	1.2	4.8	1.00
Treatment	27	0.0	0.4	0.6	3.5	10.0	2.04
Social Environment	26	0.0	5.0	7.5	8.7	12.8	7.03
Interviewer-Administered	**27**	**5.5**	**9.2**	**11.9**	**14.0**	**19.5**	**11.62**

Source: National Household Survey on Drug Abuse: Development of Computer-Assisted Interviewing Procedures; 1998 Field Test.

12.8 The 1999 NHSDA CAI

The NHSDA CAI fielded in 1999 incorporated those features that were tested throughout the prior two years of development, as described in this report, that were determined to decrease respondent burden and increase data quality. Decisions resulting from this research regarding survey methodology and subject matter content for the NHSDA beginning in 1999 are described in the next two sections.

12.8.1 Methodology

Major features of NHSDA methodology implemented in 1999 are summarized below:

1. Electronic screening and case management were implemented using the Apple Newton hand-held computer (see Chapters 5.3 and 8 for a full description of the electronic screener development and testing). A Spanish translation of the screener instrument was added.

   Screening is conducted as follows: at the beginning of the data collection period, each field interviewer's (FI's) list of assigned cases (case ID numbers and addresses) is transmitted via phone lines to his or her Newton computer from RTI's central office. On each visit or attempted visit to an address, whether for screening or interviewing, the FI enters the result code for the visit and any relevant notes into the Newton. If a household respondent is contacted, the FI asks the appropriate screening questions, enters responses, and the Newton computer selects zero, one, or two sample person(s) for interviewing.

2. The CAPI/ACASI methodology was adopted for the NHSDA interview**.** Both software and hardware were upgraded for the August, 1998 pretest and subsequent field work beginning in January, 1999. The interview was programmed in Blaise 4.0 for Windows and loaded on Gateway Solo™ 5100 Multimedia Notebook laptop computers. This computer has a 14" monitor, built-in sound card, 56K modem, 64 MB of RAM, 4 GB hard disk space, and a 233 MHZ Pentium II processor.

   A Spanish translation of the questionnaire was added.

   The CAI interview is conducted as follows: for each sample person selected, a unique questionnaire ID number is generated by the Newton, which must be entered into the laptop computer by the FI in order to start the interview. The FI selects the language to be used (English or Spanish) and conducts the initial CAPI portion of the interview, then turns the laptop over to the respondent, pointing out the keys the respondentwill use and giving instructions in use of the headphones.

   The function keys that respondents are instructed to use for backing up one question at a time, turning off the sound, replaying the sound, and entering "don't know" and "refused" responses are labeled with a template.

   Before the actual ACASI portion of the interview begins, the respondent is presented with the short, interactive ACASI tutorial described earlier that provides basic instructions and practice in entering responses to different types of questions, changing responses, and having questions repeated. The FI makes every effort to ensure that no person other than the respondent can see or hear the questions. After the respondent completes the ACASI portion, the FI administers the remaining CAPI portion of the interview. Finally, the FI enters responses to a few questions about his/her impressions of the interview. These "FI debriefing items" were revised in 1999 to include an assessment of the effect of the computer on the respondent's decision to participate and on the privacy of the interview.

Each day, FIs transmit all screening work from the Newton and interviews from the laptop to RTI headquarters over telephone lines. Screening data are updated daily and made available on the project website to RTI and SAMHSA staff for monitoring field progress. (See Chapters 9.1.4 and 9.1.5 for a description of the data transmission, capture and monitoring systems.)

3. Experimental Factor 1: Structure of contingent questioning in the CAI questionnaire: a single gate question for all core substances was selected over multiple gate questions. (See Chapters 5 and 7.1.)
4. Experimental Factor 2: Data quality checks were implemented throughout the CAI questionnaire, including range edits and inconsistency resolution in CAPI and ACASI. (See Chapters 5 and 7.2.)
5. Experimental Factor 3: A single opportunity to report 30 day and 12 month substance use was selected over multiple opportunities to report use. (See Chapters 5 and 7.3.)
6. A two-stage question was implemented for asking 12-month frequency of use, in which respondents first choose the metric that is easiest for them (days per year, month or week) and then report number of days.
7. The voice used for the ACASI sections of the interview was the preferred female voice from our cognitive laboratory study (see Caspar and Edwards, 1997).

12.8.2 Questionnaire Content

Significant new questionnaire content was introduced and tested in the final pretest in August 1998, and fielded in January, 1999. This new content is summarized below.

1. An entirely new tobacco module was fielded, with questions on several smoking-related topics:

   - usual (past 30 day) brand of cigarettes, snuff, chewing tobacco, and cigars: wording for cigarettes and cigars was changed from "even a few puffs" to "part or all of a _____," and snuff and chewing tobacco were asked about separately;

   - questions on use of pipe tobacco and "roll-your-own" cigarettes; and

   - a set of questions for adolescents who do not smoke, asking about the likelihood of initiating smoking in the next year.

2. Response categories for the recency of use questions were shortened from four to three: "within the past 30 days," "more than 30 days ago but within the past 12 months," and "more than 12 months ago".
3. A new question on "month of first use" of all drugs was added for better incidence data on persons recently initiating substance use.
4. The "pill cards" (showing prescription drugs) were updated.
5. Questions on non-medical use of prescription drugs were revised and tested, and the term "pain killer" for analgesics was changed to "pain reliever."
6. New questions were added to the drug dependence section to assess *withdrawal symptoms* resulting from use of specific drugs. These questions were based on criteria for withdrawal as described in the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) published by the American Psychiatric Association (1994).
7. Two questions traditionally asked in the NHSDA core drug sections were deleted as a result of field and lab testing: lifetime number of days a substance was used and number of days in past 12 monthsthat a respondent got very high or drunk. The former question was determined to be too difficult for most respondents to answer accurately, and the latter was poorly understood, based on lab testing and on data from consistency checks in the field experiment.
8. Because of increasing concerns about *methamphetamine* abuse, new questions were added to the core "stimulants" section to estimate the incidence (age at first use, including month and year for recent initiates) and recency of use of methamphetamines.
9. New content in the socio-demographic sections includes additional questions on employment status, country of birth, and the opportunity to report multiple racial and Hispanic groups.

12.8.3 Summary

We believe the new CAI instrument and procedures implemented in 1999 represent a significant improvement over the PAPI instrument and procedures employed from 1971 through 1998. The new methodology was implemented in 1999 on a sample of approximately 67,000 respondents, with an additional PAPI sample of approximately 14,000. A report on the implementation and assessment of the new design, "Methodological Issues in the 1999 Redesign of the NHSDA" will be published later in 2001. This

report discusses particular problems and issues addressed during the redesign, as well as the effect of interview mode (CAI vs. PAPI) on the reporting of drug use data. Additional information on the NHSDA, including the 1999 questionnaire, details on all data collection procedures, and the Summary of Findings from the 1999 NHSDA can be accessed through the World Wide Web, at **http://www.oas.samhsa.gov**.

[19] Recall that no formal hypothesis testing of comparisons of means was made.

References

Abma, J., Chandra, A., Mosher, W., Peterson, L., & Piccinino, L. (1997, May). Fertility, family planning, and women's health: New data from the 1995 National Survey of Family Growth. Vital and Health Statistics, 23(19), 1-114.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.) Washington, DC: Author.

Aquilino, W. (1994). Interview mode effects in surveys of drug and alcohol use. Public Opinion Quarterly, 58, 210-240.

Aquilino, W., & LoSciuto, L. (1990). Effect of interview mode on self-reported drug use. Public Opinion Quarterly, 54, 362-395.

Aquilino, W.S. (1997). Privacy effects on self-reported drug use: Interactions with survey mode and respondent characteristics. In L. Harrison & A. Hughes (Eds.), The validity of self-reported drug use: Improving the accuracy of survey estimates (NIH Publication No. 97-4147, NIDA Research Monograph 167, pp. 383-415). Rockville, MD: National Institute on Drug Abuse.

Bearman, P.S., Jones, J., & Udry, J.R. (1997). The National Longitudinal Study of Adolescent Health: Research design, [On-line]. Available: http://www.cpc.unc.edu/projects/addhealth/design.html. [1999, November].

Caspar, R. (1992). Follow-up of nonrespondents in 1990. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929, pp. 155-176). Rockville, MD: National Institute on Drug Abuse.

Caspar, R., & Edwards, T. (1997, June 12). 1997 National Household Survey on Drug Abuse: Results from cognitive laboratory testing to develop procedures for Field Test #2: Final report (prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-96-0001). Research Triangle Park, NC: Research Triangle Institute.

Cox, B.G., Witt, M.B., Traccarella, M.A., & Perez-Michael, A.M. (1992). Inconsistent reporting of drug use in 1988. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929, pp. 109-154). Rockville, MD: National Institute on Drug Abuse.

Dijkstra, W., & van der Zouwen, J. (1982). Response behaviour in the survey interview. New York: Academic Press.

Duffer, A., Lessler, J., Weeks, M., & Mosher, W. (1996). Impact of incentives and interviewing modes: Results from the National Survey of Family Growth Cycle V pretest. In R. Warnecke (Ed.), Health survey research methods (pp. 147-152). Hyattsville, MD: National Center for Health Statistics.

Duffy, J.C., & Waterton, J.J. (1984). Under-reporting of alcohol consumption in sample surveys: The effect of computer interviewing in fieldwork. British Journal of Addiction, 79, 303-308.

Forsyth, B.H., Lessler, J.T., & Hubbard, M.L. (1992). Cognitive evaluation of the questionnaire. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929, pp. 13-52). Rockville, MD: National Institute on Drug Abuse.

Gfroerer, J. (1992, August). An overview of the National Household Survey on Drug Abuse and related methodological research. In Proceedings of the Section on Survey Research Methods (pp. 464-469). Alexandria, VA: American Statistical Association.

Gfroerer, J., Lessler, J., & Parsley, T. (1997a). Studies of nonresponse and measurement error in the National Household Survey on Drug Abuse. In L. Harrison & A. Hughes (Eds.), The validity of self-reported drug use: Improving the accuracy of survey estimates (NIH Publication No. 97-4147, NIDA Research Monograph 167, pp. 273-295). Rockville, MD: National Institute on Drug Abuse.

Gfroerer, J., Wright, D., & Kopstein, A. (1997b). Prevalence of youth substance use: The impact of methodological differences between two national surveys. Drug and Alcohol Dependence, 47, 19-30.

Gfroerer, J.C., & Hughes, A.L. (1991). The feasibility of collecting drug abuse data by telephone. Public Health Reports, 106, 384-393.

Gfroerer, J.C., & Hughes, A.L. (1992). Collecting data on illicit drug use by phone. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929, pp. 277-295). Rockville, MD: National Institute on Drug Abuse.

Groves, R.M. (1991). Measurement errors across disciplines. In P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, & S. Sudman (Eds.), Measurement errors in surveys (pp. 1-28). New York: John Wiley & Sons, Inc.

Harrell, A. (1997). The validity of self-reported drug use data: The accuracy of responses on confidential self-administered answer sheets. In L. Harrison & A. Hughes (Eds.), The validity of self-reported drug use: Improving the accuracy of survey estimates (NIH Publication No. 97-4147, NIDA Research Monograph 167, pp. 37-58). Rockville, MD: National Institute on Drug Abuse.

Harrell, A.V. (1985). Validation of self-report: The research record in self-report methods of estimating drug use. In B.A. Rouse, N.J. Kozel, & L.G. Richards (Eds.), Self-report methods of estimating drug use: Meeting current challenges to validity (DHHS Publication ADM 85-1402, NIDA Research Monograph 57, pp. 12-21). Rockville, MD: National Institute on Drug Abuse.

Hubbard, M. (1992). Laboratory experiments testing new questioning strategies. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929, pp. 53-84). Rockville, MD: National Institute on Drug Abuse.

Hubbard, M.L., Pantula, J., & Lessler, J.T. (1992). Effects of decomposition of complex concepts. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929, pp. 245-266). Rockville, MD: National Institute on Drug Abuse.

Kelly, J.E., Mosher, W.D., Duffer, A.P., Jr., & Kinsey, S.H. (1997, October). Plan and operation of the 1995 National Survey of Family Growth. Vital and Health Statistics, 1(36), 1-89.

Kessler, R.C., McGonagle, K.A., Zhao, S., Nelson, C.B., Hughes, M., Eshleman, S., Wittchen, H.U., & Kendler, K.S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Archives of General Psychiatry, 51, 8-19.

Lessler, J.T., & Durante, R. (1992). 1992 National Household Survey on Drug Abuse: Findings of first quarter skip test (prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 271-91-5402, RTI/5071). Research Triangle Park, NC: Research Triangle Institute.

Lessler, J.T., Weeks, M., & O'Reilly, J. (1994). Results from the National Survey of Family Growth Cycle V Pretest. In Proceedings of the Section on Survey Research Methods (pp. 64-70). Alexandria, VA: American Statistical Association

Nicols, B., Grulley, N., Thalji, L., & Tourangeau, R. (1997, February). The Prospective Longitudinal Study: In-home wave I final report. Chicago, IL: National Opinion Research Center.

O'Reilly, J., Hubbard, M., Lessler, J., Biemer, P., & Turner, C. (1994). Audio and video computer assisted-self interviewing: Preliminary tests of new technologies for data collection. Journal of Official Statistics, 10(2), 197-214.

Parsley, T. (1993). Report on 1990 NHSDA-Census Match (prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 271-89-8333, RTI/4469-15). Research Triangle Park, NC: Research Triangle Institute.

Research Triangle Institute. (1997, June 19). 1997 National Household Survey on Drug Abuse: Electronic instruments field test analysis, final report: Field Test 1: Fall 1996: Deliverable No. 11 (prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-96-0001). Research Triangle Park, NC: Author.

Research Triangle Institute. (1998, November 5). National Household Survey on Drug Abuse: Development of computer-assisted interviewing procedures: 1997 field experiment (revised draft report prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-96-0001, RTI Project No. 6616). Research Triangle Park, NC: Author.

Schober, S.E., FeCaces, M., Pergamit, M., & Branden, L. (1992). Effects of mode of administration on reporting of drug use in the National Longitudinal Survey. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929, pp. 267-276). Rockville, MD: National Institute on Drug Abuse.

Schwarz, N., & Hippler, H. (1991). Response alternatives: The impact of their choice and presentation order. In P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, & S. Sudman (Eds.), Measurement errors in surveys (pp. 41-56). New York: Wiley and Sons.

Shah, B.V., Barnwell, B.G., & Bieler, G.S. (1998). SUDAAN user's manual: Version 7.5. Research Triangle Park, NC: Research Triangle Institute.

Substance Abuse and Mental Health Services Administration. (1996a, April). Development and implementation of a new data collection instrument for the 1994 National Household Survey on Drug Abuse (DHHS Publication No. SMA 96-3084, prepared under Contract No. 283-93-5409 by L.A. Kroutil, L.L. Guess, M.B. Witt, L.E. Packer, J.V. Rachal, & J.C. Gfroerer). Rockville, MD: Office of Applied

Studies.

Substance Abuse and Mental Health Services Administration. (1996b). Mental health estimates from the 1994 National Household Survey on Drug Abuse (Advance Report No. 15). Rockville, MD: Office of Applied Studies.

Substance Abuse and Mental Health Services Administration. (1997). National Household Survey on Drug Abuse: Main findings 1995 (DHHS Publication No. SMA 97-3127). Rockville, MD: Office of Applied Studies.

Substance Abuse and Mental Health Services Administration. (2000). Summary of findings from the 1999 National Household Survey on Drug Abuse (DHHS Publication No. SMA 00-3466, NHSDA Series H-12, available at http://www.drugabusestatistics.samhsa.gov/). Rockville, MD: Office of Applied Studies.

Turner, C., Ku, L., Sonenstein, F., & Pleck, J. (1996). Impact of ACASI on male-male sexual contacts: Preliminary results from the 1995 National Survey of Adolescent Males. In R. Warnecke (Ed.), Health survey research methods (pp. 171-176). Hyattsville, MD: National Center for Health Statistics.

Turner, C.F., Lessler, J.T., & Devore, J.W. (1992a). Effects of mode of administration and wording on reporting of drug use. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929, pp. 177-220). Rockville, MD: National Institute on Drug Abuse.

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

Turner, C.F., Lessler, J.T., George, B.J., Hubbard, M.L., & Witt, M.B. (1992b). Effects of mode of administration and wording on data quality. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929, pp. 221-244). Rockville, MD: National Institute on Drug Abuse.

Turner, C.F., Lessler, J.T., & Gfroerer, J.C. (Eds.). (1992c). Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929). Rockville, MD: National Institute on Drug Abuse.

Witt, M.B., Pantula, J., Folsom, R.E., & Cox, B.G. (1992). Item nonresponse in 1988. In C.F. Turner, J.T. Lessler, & J.C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies (DHHS Publication No. ADM 92-1929, pp. 85-108). Rockville, MD: National Institute on Drug Abuse.

Wojcik, M., Bard, S., & Hunt, E. (1991). Training field interviewers to use computers: A successful CAPI training program. Paper presented at the Annual Conference of the American Association for Public Opinion Research, Phoenix, AZ.