2005 NATIONAL SURVEY ON DRUG USE AND HEALTH

RELIABILITY STUDY PRETEST FINAL REPORT

Contract No. 283-2004-00022
RTI Project No. 0209009.124
Deliverable No. 31

Authors:

Lanny Piper
Martin Meyer
Jeanne Snodgrass

Project Director:

Tom Virag

Prepared for:

Substance Abuse and Mental Health Services Administration
Rockville, Maryland 20857

Prepared by:

RTI International
Research Triangle Park, North Carolina 27709

July 2006

2005 NATIONAL SURVEY ON DRUG USE AND HEALTH

RELIABILITY STUDY PRETEST FINAL REPORT

Contract No. 283-2004-00022
RTI Project No. 0209009.124
Deliverable No. 31

Authors:

Lanny Piper
Martin Meyer
Jeanne Snodgrass

Project Director:

Tom Virag

Prepared for:

Substance Abuse and Mental Health Services Administration
Rockville, Maryland 20857

Prepared by:

RTI International
Research Triangle Park, North Carolina 27709

July 2006

Table of Contents

List of Tables

1. Introduction

Information on the quality of data should be a standard output of major surveys. Most surveys do a good job reporting on response rates and sampling error but not on other types of error, such as measurement error. While the sampling error provides information on the precision of the estimates, it does not indicate the accuracy with which the data are being gathered. Determining the reliability of survey measures by reinterviewing respondents provides a direct measure of response variance. In other words, the capability of the survey to provide accurate data and consequent population estimates can be determined by carrying out an assessment of its reliability. The reliability of survey data is of particular concern when the data reflect responses to questions that are sensitive in nature, that is, when respondents are asked to reveal personal characteristics and behaviors that they feel might be embarrassing or damaging to themselves if certain others were to find out about them. The National Survey on Drug Use and Health (NSDUH) gathers data on the recency and frequency of use of alcohol, tobacco, and illicit substances. Recognizing that drug use, and particularly illicit drug use, is likely to be considered a sensitive topic, NSDUH staff have incorporated state-of-the-art methods to assure respondents of their privacy and confidentiality in providing truthful data. The majority of the survey is administered in the audio computer-assisted self-interviewing (ACASI) mode so that even the interviewers are not aware of responses to the more sensitive portions of the questionnaire. However, despite these and other precautions related to privacy and confidentiality, the reliability and validity of NSDUH data continue to be questioned. In essence, assessing the reliability of the NSDUH questionnaire is seen as an important step that must be taken in order to quantify the error in key survey estimates that is due to response variance, and potentially to use this information to improve the instrument where improvements are needed. Those items that are shown to be unacceptably low in reliability can then be subjected to cognitive testing and other testing methods to determine the specific sources of measurement error and to minimize their effects.

Thus, the Substance Abuse and Mental Health Services Administration (SAMHSA) felt it was important to conduct a formal test of the NSDUH instrument's reliability. A reliability study on the NSDUH questionnaire was planned for the 2006 survey. The goal of the reliability study is to provide estimates of the reliability of the NSDUH questionnaire for three age groups: 12-17, 18-25, and 26+. The survey instrument will be administered to a sample of respondents on two occasions. The reliability of the instrument will be measured by the degree of agreement between the survey responses from the two administrations, focusing mainly on substance use and mental health questions. The NSDUH questions have reference periods for past month, past year, and lifetime experiences. However, only past year and lifetime measures are to be examined in the reliability study.

To test the reliability study field procedures, materials, and instrumentation, as well as respondents' reactions to the reinterview, a reliability study pretest with 200 reinterview respondents was planned and conducted in Quarters 1 and 2 of the 2005 survey. The pretest was scheduled and conducted across two phases. Phase 1 was conducted in January and February of 2005, and Phase 2 was conducted in April and May of 2005. The month of March was used to assess the Phase 1 experience, analyze the results, and make any refinements to the procedures,

materials, and instrumentation for the Phase 2 implementation. The original plan was to conduct Phase 2 data collection in May and June of 2005, giving us 2 months following Phase 1 to analyze the findings and make adjustments for Phase 2. However, Phase 1 went better than expected with only a few minor problems. Thus, the decision was made to accelerate the start of Phase 2. After the completion of Phase 2, we assessed the findings of both phases, prepared a report of the analyses conducted with the findings, and made appropriate refinements to the full reliability study design based on the pretest findings. The pretest was mainly intended for examining response rates, methods, and feasibility of administration. Many of the reliability analysis measures planned for the full-scale reliability study were applied to the pretest sample in preparation for and as a test of the procedures to produce the reliability study measures.

2. Sample Design

The reliability study pretest was conducted in two phases with separate sets of segments selected for each phase. We designed the sample for the pretest as a nonprobability-based sample. Since we felt it was important to observe some of the pretest interviews, we selected the sample in States that could be easily observed by the Substance Abuse and Mental Health Services Administration (SAMHSA) staff or by the National Survey on Drug Use and Health (NSDUH) field management staff. We selected a total of 12 retired 2004 NSDUH segments in North Carolina, Maryland, Florida, and Texas (three segments in each State) in each phase of the pretest. Thus, a total of 24 retired segments were used. As a cost efficiency measure, we selected these segments based on proximity to field interviewers (FIs) identified to work on the pretest.

We randomly selected dwelling units within these designated segments. Within sample households that completed the screening interview, the sampling algorithm programmed on the iPAQ to select eligible residents within a cooperating household was designed to select 0 or 1 person as opposed to the main study sampling algorithm that selected 0, 1, or 2 eligible residents. If a second person could be chosen in a reliability study household, then the second person could learn from the first person about the follow-up interview, and that would compromise the purpose of the experiment.

Like NSDUH, persons eligible for the pretest were civilian, noninstitutionalized persons aged 12 years old or older. This included residents of noninstitutional group quarters (e.g., shelters, rooming houses, dormitories) and civilians residing on military bases. We targeted to complete 200 reinterviews, 100 in each phase.

We also imposed a special same-versus-different-interviewer experiment on the pretest sample to allow some assessment of interviewer effect. We randomly assigned some segments to have the same interviewer do both the initial interview and the reinterview. For the remaining segments, we required that a different interviewer conduct the reinterview. In each phase of the pretest, eight segments were designated to the same interviewer substudy, and four segments were designated to the different interviewer substudy.

Tables 1 and 2 show the assumptions used in selecting the sample for the pretest. We expected that approximately 720 selected dwelling units would be needed to yield a total of 200 completed reinterviews. This assumed a 94 percent screening response rate (SRR) among eligible dwelling units, an 82 percent interview response rate (IRR) for initial interviews, and an 86 percent IRR for reinterviews. The expected overall response rate (ORR) was 66 percent.

Table 1. Reliability Study Pretest Design Parameters: Dwelling Unit Level

Dwelling Unit Level	Rate	N
Total Sample		
Segments		24
Selected Lines		717
Expected Eligible Dwelling Units	0.84	602*
Expected Completed Screening Interviews	0.94	566

*Based on prior NSDUH experience, we expected 16 percent of the selected dwelling units to be ineligible (institutional, nonresidential, etc.). Thus, the 717 dwelling units were reduced to 602 eligible dwelling units.

Table 2. Reliability Study Pretest Design Parameters: Person Level

	Overall		Ages 12-17		Ages 18-25		Ages 26+	
Person Level	**Rate**	**N**	**Rate**	**N**	**Rate**	**N**	**Rate**	**N**
Total Sample								
Expected Selected Persons (1st Interview)	0.50	283	0.50	82	0.50	89	0.50	112
Expected Completed Interviews (1st Interview)*	0.82	233	0.89	72	0.85	76	0.76	85
Expected Selected Persons (2nd Interview)	1.00	233	1.00	72	1.00	76	1.00	85
Expected Completed Interviews (2nd Interview)	0.86	200	0.92	67	0.88	67	0.78	67

*Rates were based on actual experience in Washington, DC, Maryland, Texas, and Florida in the 2003 survey.

For Phase 1 of the pretest, we selected 30 dwelling units in each of the 12 segments for a total of 360 selected dwelling units. We also selected a "reserve" sample of 6 dwelling units in each segment in case our assumptions proved to be significantly off, but we never needed to use them. We completed 112 initial interviews and 104 reinterviews. Table 3 shows the response rate information for Phase 1. We achieved an 89.5 percent screening response rate, a 79.4 percent initial interview response rate, and a 92.9 percent reinterview response rate for an overall response rate of 66.0 percent.

Table 3. Reliability Study Pretest Phase 1 Response Rates, by State

State	Overall Response Rate	Total Selected DUs	Number Screenings Completed	Number Interviews Completed	Number Reinterviews Completed	Screening Response Rate	Interview Response Rate	Reinterview Response Rate
MD	80.96	90	72	29	28	86.75	96.67	96.55
NC	78.71	90	77	30	28	92.77	90.91	93.33
FL	58.30	90	76	30	27	86.36	75.00	90.00
TX	51.01	90	72	23	21	92.31	60.53	91.30
Total	**65.98**	**360**	**297**	**112**	**104**	**89.46**	**79.43**	**92.86**

DU = dwelling unit.

Following Phase 1 of the pretest, SAMHSA decided they wanted the full reliability study for 2006 conducted as an embedded sample design. Thus, the reliability study sample was to be overlaid on the main NSDUH sample so that the initial reliability study interviews would be included with all other main study NSDUH interviews. The Phase 1 sample and procedures were

designed for a sample that was totally independent of the main study sample. This meant that FIs conducting reliability study pretest interviews knew that every interview completed in the pretest was a reliability study case and would require an attempt to conduct a follow-up interview. This would not be true when the reliability study sample was embedded with the main study sample.

To simulate an embedded design for Phase 2 of the pretest would have required a much larger sample than planned, many time-consuming system changes, and major procedural changes. This would have taken more time than was available in order to be ready to begin the full reliability study in Quarter 1 of 2006 and would have cost significantly more to complete than planned. Therefore, the decision was made to keep the Phase 2 sample independent of the main study Quarter 2 sample by using 12 retired 2004 NSDUH segments as originally planned. However, to simulate that not every case selected would be a reliability study case, there were some sample dwelling units in Phase 2 that were not designated as a reliability study case and thus did not require a follow-up interview. This increased the number of dwelling units selected for Phase 2 to 34 dwelling units per segment for a total of 408 sample dwelling units.

Table 4 shows the response rate information for Phase 2. We achieved an 82.3 percent screening response rate, an 83.2 percent initial interview response rate, and a 90.8 percent reinterview response rate for an overall response rate of 62.2 percent.

Table 4. Reliability Study Pretest Phase 2 Response Rates, by State

State	Overall Response Rate	Total Selected DUs	Number Screenings Completed	Number* Interviews Completed	Number Reinterviews Completed	Screening Response Rate	Interview Response Rate	Reinterview Response Rate
MD	60.38	102	69	30	27	76.67	87.50	90.00
NC	63.48	102	74	29	26	82.22	86.11	89.66
FL	55.06	102	67	29	25	77.01	82.93	86.21
TX	70.33	102	79	31	30	94.05	77.27	96.77
Total	**62.20**	**408**	**289**	**119**	**108**	**82.34**	**83.23**	**90.76**

*Includes only initial interviews designated as reliability study cases.

One concern that was raised during discussions of the reliability study design was the potential impact an FI may have on the follow-up interview. To allow some assessment of the interviewer effect, an experiment was imposed on the reliability study sample where one third of the cases were assigned to have the follow-up interview conducted by an interviewer who was different than the interviewer who conducted the initial interview. The other two thirds of the cases were designated to have the follow-up interview conducted by the same interviewer who conducted the initial interview. This experiment was imposed on the pretest sample in order to design and test the procedures required to carry it out. For those cases where a different FI was designated to conduct the follow-up interview, the field supervisor (FS) responsible for the area selected the different FI from those available nearby so as to minimize the cost for the follow-up interview.

3. Data Collection Management

We assigned the same field management team for the reliability study pretest that was in place for the 2005 National Survey on Drug Use and Health (NSDUH) main study. The States selected for the pretest were all under the responsibility of one regional director (RD). The national field director and RD, along with the three regional supervisors (RSs) and four field supervisors (FSs) who supervised these States, managed the pretest data collection. This team managed production, costs, and quality using web-based case management system (CMS) reports and functions. The majority of these reports were identical to main study reports except they were populated with pretest data. There were some special CMS reports developed for the reliability study to track the follow-up interview activation, same versus different field interviewer (FI) assigned, and the time window for completing the follow-up interview. This field management team was trained on pretest field management responsibilities and the CMS reports and functions to manage the fieldwork during a special session held at the January 2005 Veteran FI Training session.

4. Field Interviewer Staffing and Training

The field management team determined that the current field interviewer (FI) workforce in each selected State could easily handle both their regular main study workload along with the pretest workload. The team identified six FIs in each State for a total of 23 FIs who were veteran National Survey on Drug Use and Health (NSDUH) interviewers with a history of good performance on the main study.

These 23 FIs were trained on the reliability study pretest data collection procedures in person in a half-day training session following the conclusion of the January 2005 Veteran FI Training session. In addition to the mechanics of administering the instrument, using the specially developed reliability study materials, making appointments for the follow-up interview, and the special procedures in the iPAQ to initiate a reinterview, emphasis was placed on the importance to the research objectives of making the initial screening and interview processes no different from the main study NSDUH processes. Interviewers were specifically directed to never mention anything about the reliability study prior to the respondent completing the initial interview. The first time any mention of a special study and a request for a follow-up interview was to occur was when the computer-assisted interviewing (CAI) instrument presented the reliability study recruitment scripts that contained scripted language about the reliability study.

5. Data Collection

Data collection was designed to be conducted in two phases. Phase 1 began right after the National Survey on Drug Use and Health (NSDUH) Veteran Field Interviewer (FI) Training in January 2005 and concluded at the end of February 2005. To allow time for analyzing and assessing the Phase 1 activities and to implement improvements based on the Phase 1 assessment, Phase 2 was scheduled to begin in May 2005. However, Phase 1 went so well that there were only minor changes required for Phase 2. Therefore, we started Phase 2 a month earlier than planned on April 1, 2005, and completed all data collection by the end of May 2005.

Data collection began with the initial (T1) interview conducted exactly as a main study interview was conducted. At the end of the T1 interview, the computer-assisted interviewing (CAI) questionnaire program prompted the FI with wording designed to recruit respondents for a follow-up (T2) interview. The interviewer made no mention of the follow-up interview prior to this point. The CAI questionnaire program also determined if the T2 interview was to be conducted by the same FI or a different FI. For the same FI substudy, the FI went ahead and set the T2 appointment for anywhere from 5 to 15 days in the future. For the different FI substudy, the interviewer told the respondent that another FI would contact him or her to set up an appointment for the T2 interview and asked for the respondent's phone number so that the other FI could call to make an appointment. Administrative procedures at the end of the T1 and T2 interviews were identical to the main study—the respondent was given a cash payment and a signed incentive receipt and was asked to complete a Quality Control Form. At the close of the T1 interview, the respondent received a $30 cash payment just as was done in the main study. The payment at the end of the T2 interview was $50.

Both T1 and T2 interviews were identical to the main study except at the very end. The T1 questionnaire ended with the recruitment scripts mentioned above, and the T2 questionnaire ended with additional questions for the respondent and the FI. After completing the standard main study back-end computer-assisted personal interviewing (CAPI) portion of the interview, the FI gave the laptop back to the respondent to answer some extra "debriefing" questions in the audio computer-assisted self-interviewing (ACASI) format. Afterwards, the laptop was returned to the FI to complete the interview. The FI had three extra debriefing questions as well. In Phase 1, 112 T1 interviews and 104 T2 interviews were completed, and in Phase 2, 119 T1 interviews and 108 T2 interviews were completed. (See Tables 3 and 4 in Chapter 2 for detailed sample sizes and response rates).

Interview verification procedures also mirrored those of the NSDUH main study. Since this was a pretest with a small sample and we wanted to obtain as much information from respondents as feasible regarding the pretest process, we decided to verify all pretest interviews in both Phase 1 and Phase 2. All reliability study respondents received a telephone call expressing appreciation for their participation in the study. Each respondent was also asked to answer a few questions verifying that the interview took place, that proper procedures were followed, and that the amount of time required to administer the interview was within expected parameters. Cases where both T1 and T2 interviews were completed were only verified once following the completion of the T2 interview. Respondents who agreed to the T2 interview but were unable to be contacted for the T2 interview or who refused to complete the T2 interview at

any point were asked some additional questions during verification to try to determine why the respondent chose not to participate in the reinterview. All verification and refusal results are presented in the final reliability study pretest analysis report (Hunter, Feder, Granger, Piper, & Chromy, 2006).

Table 5 presents the primary changes between Phase 1 and Phase 2 for each task associated with the study. There were no changes to data collection materials, instruments, or staff between the two phases.

Following each phase, telephone debriefings were conducted with the FIs to probe them on their experiences with conducting the reliability study interviews and to obtain their opinions and suggestions for what worked well and what improvements could be made. There were three debriefing sessions for each phase with approximately six interviewers in each debriefing session. The input from the debriefings was used to augment the pretest analysis findings and was used to refine materials and procedures for the 2006 reliability study. Details of the debriefing findings can be found in the final reliability study pretest analysis report (Hunter et al., 2006).

Table 5. NSDUH Reliability Study Pretest Design and Process Changes between Phase 1 and Phase 2: 2005

Task	Phase 1		Phase 2	
	Description	**Issues/Results**	**Description**	**Reason for Change**
Methodology	In-person CAI using the 2005 NSDUH methodology (including a $30 incentive) and questionnaire with special scripts at end of regular interview to solicit participation in follow-up interview 5-15 days later. A $50 incentive given for the follow-up interview. No mention of the follow-up interview made until prompted following completion of the initial interview.	Completed 112 initial interviews and 104 follow-up interviews for a 92.9 percent follow-up response rate.	Same as Phase 1.	No change.
Sample	Purposive sample of retired segments from the 2004 NSDUH sample in four States: North Carolina, Maryland, Florida, and Texas. These States were selected because of their stable interviewing staffs, convenience for RTI and SAMSHA staff to observe, convenience for assigning different FIs to some segments, variety of large urban and rural areas, and convenience of having a local regional supervisor who could observe interviews. Randomly selected DUs from those not selected for 2004 NSDUH. Select 0 or 1 person in eligible sample dwelling units. Every person selected was also selected for the reliability study follow-up interview.	Selected and screened 360 DUs with an 89.5 percent screening response rate. The screenings yielded 141 selected persons; 112 selected persons completed the initial interview (79.4 percent). Completed 104 follow-up interviews (92.9 percent).	Same basic design as Phase 1 except extra DUs were selected that were not designated as reliability study cases. This was expected to yield an additional 12 initial interviews with no follow-up interview. Interviewers were not aware which cases were reliability cases and which were the extra nonreliability cases until the initial interview was completed. This provided some simulation of a design where the reliability study sample was embedded in the main study sample.	Decision was made to embed the full reliability study sample in the 2006 NSDUH main study sample where not every sample DU would be designated for the reliability study. Doing so during Phase 2 provided an opportunity to test field and reporting procedures using this design. An embedded design was selected mainly to reduce costs but also to reduce potential interviewer effects since interviewers will be blinded to which cases will be reliability study cases.

Table 5. NSDUH Reliability Study Pretest Design and Process Changes between Phase 1 and Phase 2: 2005 (continued)

Task	Phase 1		Phase 2	
	Description	**Issues/Results**	**Description**	**Reason for Change**
Field Staff	Chose experienced NSDUH FIs to carry out pretest.	Implemented as planned with no problems encountered.	All FIs who were selected to receive Phase 2 assignments must have worked in Phase 1.	No change.
Materials	The screening and initial interview materials were the same as those used for the main study except for minor wording changes due to the sample size difference. The follow-up interview materials were specially adapted for the pretest and approved by the RTI IRB and by OMB.	Implemented as planned with no problems encountered.	Same as Phase 1.	No change.
Training	Completed 4.5 hour in-person training.	Implemented as planned with no problems encountered.	Home study followed by group conference call with each State's interviewers was conducted.	This was as planned since all interviewers working in Phase 2 worked in Phase 1 and only refresher training was necessary.
Instrumentation	The screening interview, the CAI, and the case manager were set up as if for an embedded study by totally integrating these within the main study systems. Some extra interviewer debriefing questions were added to the initial interview and some additional ACASI respondent questions were added to the follow-up interview to capture respondents' reactions to the reinterview.	Implemented as planned with no problems encountered.	Same as Phase 1.	No change.

Table 5. NSDUH Reliability Study Pretest Design and Process Changes between Phase 1 and Phase 2: 2005 (continued)

Task	Phase 1		Phase 2	
	Description	**Issues/Results**	**Description**	**Reason for Change**
Schedule	Data collection conducted January 8, 2005 through February 28, 2005.	Implemented as planned with no problems encountered.	Data collection scheduled for April 1, 2005 through May 31, 2005.	Data collection changed from May 1, 2005 through June 30, 2005 since Phase 1 went very well and there were only minor changes needed for Phase 2. This change allowed more time to analyze the pretest data and develop the 2006 reliability study design.
Verification	Telephone verification of all pretest cases. The standard NSDUH verification interview with minor modifications was used and with special questions asked of respondents who refused the follow-up interview or were unable to be contacted at the follow-up interview.	Implemented as planned with no problems encountered. Conducted 34 screening verifications, 8 initial interview only verifications, and 103 verifications of cases where both the initial and follow-up interview was completed.	Same as Phase 1.	No change.

ACASI = audio computer-assisted self-interviewing, CAI = computer-assisted interviewing, DU = dwelling unit, FI = field interviewer, IRB = Institutional Review Board, OMB = Office of Management and Budget.

6. Data Processing

The reliability study design, reporting requirements, and analysis expectations provided new challenges to our data processing staff. This study required that the same person be interviewed twice within the same quarter and that the computer-assisted interviewing (CAI) instrument had to know if the second interview was to be conducted by the same field interviewer (FI) who completed the first interview or by a different FI. We considered three alternatives for accommodating the data processing requirements of the reliability study pretest:

- **Alternative 1**: Provide separate systems for main study and reliability study. FIs to carry two iPAQs and two laptops. All internal data processing operations, websites, and transmission systems are duplicated for the reliability study pretest.
- **Alternative 2**: Run two parallel systems based on separate databases. Duplicate all existing National Survey on Drug Use and Health (NSDUH) processing and customize it for the reliability study pretest. FIs carry only one iPAQ and one laptop, but RTI International-based[1] data processing operations are replicated.
- **Alternative 3**: Integrate data processing for reliability study pretest cases entirely within the existing production NSDUH data processing system.

After careful consideration, we chose Alternative 3. Although it offered less flexibility with respect to implementation details, it was deemed to be the most cost-effective and most rapidly deployable option. In order to implement Alternative 3, we designed, programmed, and tested modifications to almost every component and subsystem of the NSDUH data processing system.

To support the pretest protocol within the framework of the existing data processing system, pretest sample lines were embedded into the NSDUH control system using a combination of flags and lookup tables to identify them. The behavior of the iPAQ was controlled using flags that were bundled as part of the pretest sample line data structure prior to transmission. When a pretest sample line was screened on the iPAQ, these flags triggered a separate selection algorithm that allowed for only 0 or 1 respondent to be selected. Within the NSDUH control system, each pretest sample line was paired with an "R-line" (reinterview line). Server-side systems treated the R-line as a normal sample line. After transmission to an iPAQ, the R-line remained hidden until the corresponding pretest sample line was screened to select a single respondent and a CAI interview was subsequently completed on the laptop. An activation code was issued by the CAI software at the end of the initial interview, and the FI was prompted to enter this activation code into the iPAQ. In response to this activation sequence, the iPAQ revealed the R-line, prepopulated an event history and case status mirroring the corresponding pretest sample line, and allowed entry of events related to the execution of the reinterview case.

Website functionality was added to achieve bundling of R-lines with pretest sample lines as part of the normal case assignment and transfer process. Several of the main study reports used for tracking case status, work progress, and verification were replicated, using reliability

[1] RTI International is a trade name of Research Triangle Institute.

study case data only, for purposes of managing the pretest fieldwork. In addition, some new reports specific to the reliability study pretest were created:

- Reinterview Activation Report (daily)
- Pending Reinterview Report (daily)
- Reinterview Time Window Discrepancy Report (daily)
- Reinterview Record of Calls (ROC) Time Discrepancy Report (weekly)

These reliability study pretest reports were posted to the NSDUH case management system (CMS) website, along with regular main study reports on a daily and weekly basis.

The behavior of the CAI on the laptop was controlled using predefined questionnaire ID ranges. Prior to sample selection and control system database preload, questionnaire ranges were defined to differentiate pretest cases from main study cases and to differentiate "same FI" versus "different FI" substudies within the reliability pretest sample. Upon initiating a CAI interview on the laptop, the interviewer entered a questionnaire ID provided by the iPAQ. In the pretest, this ID indicated one of several different possible scenarios:

- Main study interview
- Reliability study, initial interview, same FI substudy
- Reliability study, initial interview, different FI substudy
- Reliability study, follow-up interview, same or different FI substudy
- Main study interview (training case)
- Reliability study, initial interview, same FI substudy (training case)
- Reliability study, initial interview, different FI substudy (training case)
- Reliability study, follow-up interview, same or different FI substudy (training case)

The CAI used the information embedded in the questionnaire ID to present the appropriate interview script. For reliability study cases, this included new questions used to recruit the respondent for participation in the follow-up interview and a question to provide the interviewer with an activation code for use on the iPAQ.

Telephone verification systems were enhanced to reflect several different possible verification scenarios introduced by the reliability study pretest protocol:

- Initial interview and reinterview completed normally
- Initial interview completed, refused reinterview at time of initial interview
- Initial interview completed, refused reinterview at time of reinterview
- Initial interview completed, unable to contact the respondent for the reinterview

Scripts used by computer-assisted telephone interviewing (CATI) interviewers were altered to reflect the different scenarios, and new verification questions were added to verify

proper interviewer behavior with respect to recruitment of respondents for follow-up interviews, conduct of reinterviews, and payment of incentives. We modified algorithms used for random selection of verification cases accordingly, and we reprogrammed the verification control system to feed the appropriate reliability study parameters to the CATI systems. We also modified the software used to automatically generate letters for mail verification to support the scenarios listed above.

All pretest screening and interviewing data were transmitted to RTI using the existing data transmission systems, and those data were stored in the existing NSDUH production data structures. After post-transmission extraction of interview results, reliability study pretest interview results were separated from main study interview results, based on control system tables and using predefined questionnaire ID ranges. All subsequent processing steps applied to reliability pretest interview data (warehousing, documentation, editing and imputation) were performed in the usual manner after separation of pretest datasets from main study datasets.

7. Analysis

The purpose of the pretest analysis was to take an initial look at the data coming out of the two interviews and determine what should and could be analyzed. The main goal was to determine measures of the degree of agreement between the two administrations of the instrument, especially for the substance use and mental health items. We also wanted to look at questionnaire timing differences between the two administrations and determine if there were any unexpected results between those administrations conducted by the same interviewer compared with the administrations conducted by different interviewers. An analysis plan for the pretest was prepared and submitted to the Substance Abuse and Mental Health Services Administration (SAMHSA) for their review and input. SAMHSA requested some additions to the analysis, which were made, and the plan was finalized.

We conducted an analysis following completion of the Phase 1 data collection primarily as a means to inform us about issues that may have required changes to procedures for the Phase 2 data collection. A Phase 1 analysis was conducted and a report was prepared and submitted to SAMHSA prior to the start of the Phase 2 data collection. This analysis identified no alarming issues, and no major changes to the data collection procedures were recommended for Phase 2. Thus, the start of Phase 2 was accelerated so that it began just 1 month following Phase 1 instead of the planned 2 months.

Analysis of the full pretest was completed in August 2005, and the pretest analysis report was submitted to SAMHSA in September 2005. SAMHSA requested some additional analyses, which were completed, and the final reliability study pretest analysis report was included as Section 18 in the *2005 Methodological Resource Book* (Hunter et al., 2006).

Reference

Hunter, S., Feder, M., Granger, B., Piper, L., & Chromy, J. (2007, March). Reliability study pretest analysis [2005]. In *2005 National Survey on Drug Use and Health: Methodological resource book* (Section 18, prepared for the Substance Abuse and Mental Health Services Administration, Office of Applied Studies, under Contract No. 283-2004-00022, RTI/0209009). Research Triangle Park, NC: RTI International. [Available as a PDF at http://www.oas.samhsa.gov/nhsda/methods.cfm#2k5]