User's Guide for the Evaluation of the National Youth Anti-Drug Media Campaign

Public Use Files for the National Survey of Parents and Youth (NSPY), Rounds 1, 2 and 3

Draft #2

September 2004

Prepared for:

National Institute on
Drug Abuse
Bethesda, Maryland

Prepared by:

WESTAT
Rockville, Maryland

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

List of Appendixes

TABLE OF CONTENTS (continued)

List of Tables

TABLE OF CONTENTS (continued)

List of Tables (continued)

List of Figures

List of Exhibits

TABLE OF CONTENTS (continued)

List of Exhibits (continued)

1. INTRODUCTION

This User's Guide contains documentation for the Round 1, Round 2, and Round 3 Public Use Files for the National Survey of Parents and Youth (NSPY), a survey conducted to evaluate the National Youth Anti-Drug Media Campaign (Media Campaign). This chapter gives the background of the Media Campaign, describes the evaluation of the National Youth Anti-Drug Media Campaign (Evaluation) design, and outlines the contents of the later chapters.

The Evaluation focuses on the measurement of the outcomes and impact of Phase III of the Media Campaign on children and their parents concerning illegal drug use among youth. The Evaluation has been developed and implemented under contract to the National Institute on Drug Abuse (NIDA). Westat has conducted the Evaluation in collaboration with the Annenberg School for Communication at the University of Pennsylvania and for the first 2 years of the project with the National Development and Research Institutes, Inc. (NDRI).

1.1 Background of the Media Campaign

When Congress appropriated funding for the Media Campaign in fiscal year 1998, it required that the Office of National Drug Control Policy (ONDCP) conduct an evaluation to assess the Media Campaign's impact on reducing youth drug use. ONDCP understood this requirement to mean that a scientifically rigorous and comprehensive evaluation must be undertaken in order to enable credible statements to be made on whether the Media Campaign had an impact. ONDCP realized that in order to ensure the most scientifically rigorous, comprehensive, and independent evaluation of the fully integrated Campaign, the evaluation would have to be administered by a second party. ONDCP turned to NIDA to conduct the evaluation.

The Media Campaign is part of an effort by ONDCP to stop drug use before it starts. Other important Campaign goals are to convince youth who are occasional users of drugs to stop using them, to enhance adult perceptions of harm associated with the use of marijuana, and to emphasize to parents and influential adults that their actions can make a critical difference in preventing drug use.

The Media Campaign has progressed through three phases of increasing complexity and intensity. Phase I involved pilot testing the intervention in 12 metropolitan areas, using then existing Partnership for a Drug-Free America advertisements. During Phase I, ads were placed on television and radio, in newspapers, and on billboards. (ONDCP, 1998)

Phase II, these advertisements along with other new advertisements appeared nationwide, in addition to the test areas. The advertisements appeared not only on television, radio, billboards, and in newspapers, but also appeared on cable television, Channel One (an educational television in schools), the Internet, and in movie theaters as well as other venues (ONDCP, 1999).

Phase III marks the full implementation of the Media Campaign. It began in September 1999 and is planned to run through the spring of 2004. As in the past, an extensive range of media is used to disseminate Media Campaign messages to a national audience. Advertising space had been purchased on television, radio, newspapers, magazines, billboards, transit ads, bus shelters, movies theaters, video rentals, internet sites, Channel One, and other venues as appropriate. The television buys include spot (local), network, and cable television. One of the requirements in the Media Campaign appropriations language is that each paid advertising slot must be accompanied by a donation of equal value for public service messages from the media, known as the pro bono match. The pro bono match involves one-to-one matching time for public service advertisements or in-kind programming. In addition, Phase III features a significant interactive media component, involving content-based web sites and internet advertising.

Additional information on the Media Campaign can be found on the Internet at several sites. These include the Media Campaign website, www.mediacampaign.org; the youth website, www.freevibe.com; the parent website, www.theantidrug.com; and the Spanish language website; www.laantidroga.com.

The Evaluation relates only to Phase III of the Media Campaign. As described below, the Evaluation is based on a nationwide panel survey of youth and their parents.

1.2 Evaluation Design

The Media Campaign seeks to educate and enable America's youth to reject illegal drugs; to prevent youth from initiating use of drugs, especially marijuana and inhalants; and to convince occasional users of these drugs to stop using them.

The Evaluation has four objectives: (1) to measure changes in drug-related beliefs, attitudes, and behaviors in children and their parents; (2) to assess the relationship of the changes (in drug-related beliefs, attitudes, and behaviors), and their association with self-reported measures of media exposure; (3) to assess the association between parent's drug-related beliefs, attitudes, and behaviors with those of their children; and (4) to assess changes in the association between parent's drug-related beliefs, attitudes, and behaviors and those of their children that may be related to the Media Campaign.

The primary tool for the Phase III Evaluation is the NSPY, a national household-based survey of youth aged 9- to 18- years old and parents from the same household. The Evaluation includes the full range of youth, starting at ages 9 to 10, so that initial interviews can be conducted with children before drug use is likely to begin. They are then followed up to evaluate the impact of the Campaign as they enter the "tween" and teen years.

The Evaluation employs a panel survey design with four rounds of data collection for youth and parents over the evaluation period. Figure 1-1 is a graphic representation of the NSPY design. Round 1 of the NSPY is the recruitment phase of the study. It consists of three cross-sectional surveys lasting about 6 months each. About 81,000 dwelling units were selected for the sample in Round 1 and approximately 8,100 youth and 5,600 parents were interviewed. Rounds 2 through 4 are the followup phases of the study. Round 2 included about 6,400 households from which approximately 6,500 youth and 4,600 parents were interviewed. Round 3 included about 5,950 households from which approximately 5,850 youth and 4,250 parents were interviewed. The Round 4 data collection ended in June 2004. The data generated by this design can be analyzed both cross-sectionally and longitudinally.

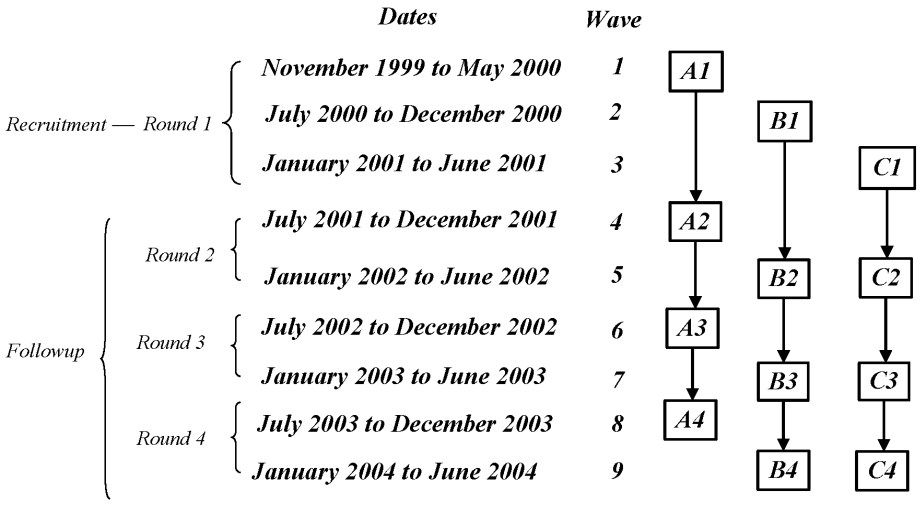

Figure 1-1. NSPY design

The Evaluation is based on presumed models of how the Media Campaign is expected to affect its target audiences. A detailed discussion of these models and factors that influence these models can be found in Chapter 2 of the 2003 Report of Findings (December 2003) prepared for the National Institute on Drug Abuse, by Westat and the Annenberg School for Communication. This report and the first five Evaluation reports are available on the ONDCP website: www.whitehousedrugpolicy.gov and on the NIDA website: www.drugabuse.gov.

1.3 Content of User's Guide

The User's Guide documents the design of the study. It provides documentation and guidance for users of three public-use data files of the Evaluation: the Round 1 cross-sectional youth file, the Round 2 cross-sectional youth file, and the Round 3 cross-sectional youth file. At this time, Round 4 data are not available. There are plans to release a Public Use File for Round 4 in the summer of 2005.

To reduce the possibility of disclosure risk, several techniques have been employed in the construction of the PUFs. First, the longitudinal data files and parent files are not included in the Public Use File data sets/documentation. (However, a few parent variables have been added to each individual youth record.) Data on the specific advertisements shown to each sampled youth have not been included in the youth record. In addition, separate random subsamples of about two-thirds of the interviewed youth have been selected for each round so that linkage of youth records is not possible across rounds and

identification of youth within a round is diminished. Each subsample was drawn in a manner designed to reduce the variation in weights to the extent possible. Each subsample was then reweighted. To further reduce disclosure risk, certain variables were suppressed or coarsened. Items suppressed for the PUFs include school enrollment; school attendance; frequency of attendance at religious services; media consumption language, habits, and choices; participation in extra-curricular activities; frequency of internet usage; extra-curricular drug-prevention services; and discussions about the interview. Items coarsened for the PUFs include race/ethnicity, age, geography, academic performance, history of cigarette and alcohol usage, youth reports of parent-child discussions, school-based drug education, importance of religion, family type, parents' education, household income, and parent reports of parent-child discussions. Readers should refer to the codebooks in Appendix A to review the coarsened categories and should refer to the annotated survey instruments in Appendix B to review all variables collected during the survey.

The User's Guide contains information about the sample design and implementation (Chapter 2), i.e., selection of the primary sampling units, area segments, dwelling units within the area segments, building permit sample, sampling and selection of respondents within households, and calculation, and information about use of sample weights (Chapter 3), i.e., weighting procedures, subsampling procedures for disclosure avoidance, and how to use the various weights provided. As mentioned above, when each round was subsampled, weights were specifically constructed for each subsample. As discussed in Chapter 3, these recalibrated (adjusted) weights should be used when analyzing the data. The replicate weights described in Chapter 3 should be used for variance estimation in order to take account of the survey's complex sample design and weighting procedures.

Data collection and methodology (Chapter 4) includes information on the data collection instruments, media activities, field staff training and data collection procedures, quality control procedures, coding and editing procedures, as well as the youth cross-sectional response rates for each round. Chapter 5 of the User's Guide shows how the exposure, outcome and risk score indices are constructed. File content and description of the variables (Chapter 6) documents the construction of the data files, the conventions used for variable naming, and the treatment of missing variables, and describes the format and content of the codebooks.

Appendices include the codebooks by round (Appendix A), the annotated survey instruments by round (Appendix B), and samples of common statistical software applications (Appendix C). The end of Appendix A contains an alphabetic index of variables (page A-204) with references to codebook page

numbers across rounds. When reviewing the annotated survey instruments, the reader can use this index to link the questions to the codebook entries. Appendix C provides information about using SAS programs, WesVar, and SUDAAN for variance calculation. These programs are set up so they can be run directly from the CD. A complete description of Appendix C is included in Chapter 6. The ReadMe.doc file on the CD provides a list of all data files, program files, and documentation files included on the CD.

REFERENCES

Office of National Drug Control Policy. (1998). Testing the Anti-Drug Message in 12 American Cities National Youth Anti-Drug Media Campaign Phase 1 (Report No. 1). Washington D.C.

Office of National Drug Control Policy. (1999). Investing in our Nation's Youth: National Youth Anti-Drug Media Campaign Phase II Final Report. Washington D.C.

Hornik, R., Maklan, D., Cadell, D., Barmeda, C.H., Jacobsohn, L., Henderson, V.R., Ramantam, A., Niederdeppe, J., Orwin, R., Sridharan, S., Chu, A., Morin, C., Taylor, K., and Steele, D. (2003). *Evaluation of the National Youth Anti-Drug Media Campaign: 2003 Report of Findings*. Report prepared for the National Institute on Drug Abuse (Contract No. N01DCA.8.5063). Washington, DC: Westat.

2. SAMPLE DESIGN AND IMPLEMENTATION

This chapter describes the sample design and its implementation for the initial data collection for NSPY (Round 1) and two followup data collection efforts (Rounds 2 and 3). Round 1 was conducted in three waves, covering the periods November 1999 through May 2000 (Wave 1), July 2000 through December 2000 (Wave 2), and January 2001 through June 2001 (Wave 3). Round 2, the first followup of the initial NSPY survey, was conducted in two waves covering the periods July 2001 through December 2001 (Wave 4) and January 2002 through June 2002 (Wave 5). Finally, Round 3, the second followup data collection, was conducted in two waves covering the periods July 2002 through December 2002 (Wave 6) and January 2003 through June 2003 (Wave 7).

Youth (and their parents) were selected for NSPY through a multistage, dual frame probability sample design. The sample was selected in a manner designed to provide an efficient and representative cross-section of America's youth. In general, youth living in all types of residential housing units were eligible for the study; however, youth living in institutions, group homes, or dormitories were excluded. Section 2.1 provides an overview of the sample design employed to select nationally representative samples of youth for the three rounds of NSPY. Sections 2.2 through 2.5 provide additional details about the sample selected for each of the three rounds. Note that to minimize disclosure risks, only a subsample of the NSPY respondents are included in the PUFs (see Section 3.2).

2.1 Overview of NSPY Sample Design

Respondents were found by door-to-door screening of a scientifically selected sample of 81,000 dwelling units within roughly 2,500 area segments and 400 building permit segments in 90 geographical areas called "primary sampling units" (PSUs). The screening sample was selected using two sampling frames. The primary frame was a list of housing units that was compiled by Westat for the sampled segments for use in a national adult literacy survey. This frame is referred to as the "area sampling frame." The selection of dwelling units from the area sampling frame was accomplished in three stages. The first stage consisted of the selection of a sample of primary sampling units (PSUs) defined to be metropolitan areas or groups of nonmetropolitan counties. The second stage consisted of the selection of area segments within the sampled PSUs, where each segment was defined to be a block or group of contiguous blocks having a minimum of 60 housing units according to the 1990 Census. For each of the

selected area segments, lists of dwelling units were compiled in a separate listing operation. Although all of the segments were originally selected in late 1991, the compilation of the listings was done at different times. The listings for segments selected for Wave 1, for example, were prepared in 1991, whereas the listings for all of the segments selected for Waves 2 and 3 were prepared in 1999. The third stage consisted of the selection of individual dwelling units from the listings compiled for each of the sampled segments.

The second frame used for NSPY consisted of building permits issued for new housing units constructed between January 1990 and December 1998 in the sampled PSUs. Thus, the building permit frame was designed to cover new construction completed between 1990 and 1998, even though the area segments for Waves 2 and 3 had been listed in 1999. This meant that dwelling units in the area sample that were determined to have been built in the period covered by the building permits were deleted from the NSPY sample. As discussed in Judkins, Cadell, and Sczerba, 2000, the use of the building permits improved sampling precision by distributing the sample of dwelling units built after 1990 more widely across the entire PSU, rather than restricting such dwelling units to the selected area segments. A question on the age of the housing unit was used to screen out housing units in the area sample that could have been selected from building permits.

Additional details about the selection of the area and building permit samples are given below.

Primary Sampling Units (PSUs)

The PSUs (defined to be metropolitan areas or rural counties) selected for NSPY consisted of a subsample of 90 of 100 PSUs that had been designed and selected for other national surveys. The PSUs in the underlying sampling frame were constructed using 1990 Decennial Census information and met the following general criteria:

- Each PSU consisted of a single county, a group of counties, or a metropolitan statistical area (MSA).
- No PSU had a 1990 population larger than 5,400,000 (the New York, Chicago, and Los Angeles PMSAs were divided into three, two, and two PSUs, respectively, in order to meet this criterion).

- The PSUs were geographically contiguous, mutually exclusive, and covered the United States.
- Nonmetropolitan PSUs did not cross state boundaries.
- Each PSU had a total population of at least 15,000 as of 1990.
- Each PSU was designed to be easily traversable by an interviewer or lister, given population density, minimum size constraints, and natural topography.

The PSU sampling frame comprised 1,404 PSUs, from which 100 PSUs were selected as follows. First, the 24 PSUs with populations greater than 2,100,000 were made "certainty" selections (i.e., selected with probability 1). The 24 certainties included three PSUs in New York, two PSUs in Chicago, and two PSUs in Los Angeles. The remaining 1,380 noncertainty PSUs were then assigned to 38 strata defined by region, metropolitan status, per capita income, percent minority population, and population size of PSU. Two PSUs were sampled from each stratum with probabilities proportionate to 1990 population using the Durbin-Brewer method (Durbin, 1967).

For NSPY, a subsample of 90 PSUs was selected from the original 100-PSU sample. To select the subsample, the noncertainty strata and two pairs of small certainty PSUs were grouped into 10 superstrata consisting of four original strata per superstratum. One stratum was then selected from each superstratum. Within the selected stratum, one of the two sample PSUs was randomly deselected. This left a total of 90 PSUs for NSPY. Additional details about the PSU selection process are given in Rizzo and Judkins (2004).

Area Segments

The second-stage sampling units for the area sample were segments defined to be groups of neighboring Census-defined blocks. Each area segment contained a minimum of 60 dwelling units according to the 1990 Census. More than 2,500 area segments were selected in Round 1 (including 1,180 in the first wave, and approximately 690 in each of the second and third waves). A fixed whole number of segments was allocated to each PSU based on the projected count of 9- to 18-year-olds in 1999 for the stratum that the PSU represented. The desired number of area segments was then selected with probabilities proportionate to the number of households in the segment. The use of a relatively large minimum size of 60 dwelling units per segment helped avoid the selection of neighboring households for the sample.

Dwelling Units in Area Segments

The area segments selected for screening during the first wave of Round 1 had been listed in late 1991 and early 1992 for a previous study. From the lists of dwelling units, systematic samples were drawn at rates designed to achieve a total sample size of 30,993 dwelling units. From the 689 segments selected for the second wave of Round 1, a total of 20,125 dwelling units were sampled. Similarly, from the 694 segments selected for the third wave of Round 1, a total of 20,248 dwelling units were sampled. Unlike the segments selected for Wave 1, the listings of dwelling units for all of the segments selected for Waves 2 and 3 were compiled in the fall of 1999 because these were not included in the sample for the prior study.

For a subsample of the dwelling units selected for Wave 1, there was a quality control check on the original 1991-1992 listings. For all single-family dwelling units, the interviewer checked for hidden apartments (such as converted basements, garages, and attics) that might have been missed by the original lister. Any detected hidden apartments were added to the sample. Also, in a subsample of multifamily housing structures, the interviewer checked for missed apartments. Using these procedures, 192 missed dwelling units were added to the sample. There was also a check for new mobile homes, which added 99 mobile homes to the sample. Thus the total sample from the Wave 1 area segments was 31,284 dwelling units. Because the Waves 2 and 3 segments were listed in the fall of 1999, these special procedures were not employed for these waves.

Building Permit Sample

Separate building permit samples were drawn for each of the three waves constituting Round 1. Use of a separate building permit sample served two main purposes. First, it provided more efficient coverage of recently constructed housing units. Second, it helped control the inflation of between-segment variances that would have occurred had new construction been included in a current area listing process.

Permit sampling is possible because most localities require that a permit be obtained before a residential structure is built. (A permit may be issued for a single structure or for multiple structures.) Moreover, the U.S. Census Bureau conducts a regular census of building permit activity from which samples can be designed and drawn. This census is conducted every month for active building permit

offices and annually for less active offices. A benefit of the census is that it can be used to select specific building permit offices and months from which to draw samples of permits for national surveys.

The stages of building permit sampling for NSPY were generally similar to those used in the area sampling frame, but there were five stages of sampling instead of three. First, only permits issued within the 90 sample PSUs were selected. Next, a sample of building permit offices (BPOs) was selected from each PSU. These BPOs were local county and city offices that issue and maintain records of building permits. At the third stage, a sample of "segments" was selected from each BPO, where here a segment was defined to be a set of permits issued by an office within a specified time period. At the fourth stage, individual building permits were selected from the sampled segments. After selection of the permits, a lister visited the building site associated with each selected permit to compile a list of all dwelling units that were found there. From each list, a sample of (new construction) dwelling units was selected.

The total dwelling unit sample size from the building permit frame was set so that the proportion of the total sample selected through the permit frame would roughly equal the proportion of the total national housing stock that was built between April 1, 1990 and the end of 1998. Statistics from the U.S. Census Bureau indicated that about 10 percent of the housing stock at the end of 1998 was built after April 1, 1990. The dwelling unit sample size from the building permit frame for Wave 1 was 3,407, which was equal to about 10 percent of the total desired sample size. In Wave 2, the sample from the building permit frame was 2,875 dwelling units, compared to 20,125 dwelling units in the area sample for Wave 2. Because the building permit frame for Wave 2 covered dwelling units that were issued permits through the end of 1998, there was no coverage of dwelling units that were built in 1999 or 2000. The Wave 3 building permit sample was 3,052 dwelling units while the corresponding area sample for Wave 3 consisted of 20,248 dwelling units. For Wave 3, there was no coverage of housing units that were built in 1999, 2000, or the first half of 2001. Table 2-1 summarizes the numbers of dwelling units selected for NSPY, by wave and source.

The selection of area and building permit segments for the purpose of listing and sampling dwelling units was done only in Round 1. For subsequent followup data collection rounds (i.e., Rounds 2 and 3), there was no new selection of dwelling units. The selection of youth and parents within eligible households is described next.

Table 2-1. Number of dwelling units selected for NSPY by wave and source

Source of sample	Wave 1	Wave 2	Wave 3
Area segments	30,993	20,125	20,248
Missed DU Procedure	192	—	—
Missed mobile homes	99	—	—
Building permits	3,407	2,875	3,052
Total	34,691	23,000	23,300

2.2 Household Screening

In general, the sampling procedures employed for NSPY were designed to obtain adequate numbers of youth in each of three targeted age groups: 9 to 11, 12 to 13, and 14 to 18 years of age. These age ranges were judged to be important analytically for evaluating the impact of the Media Campaign. In households with multiple eligible youth, a maximum of two youth was selected for the study.

Household screening and subsampling were used to identify eligible households and to oversample those with specific compositions to satisfy precision requirements for the three youth age ranges. Only households containing youth who spent the majority of the year in the household were eligible for selection. In particular, youth with divorced or separated parents were associated with the household where the youth spent the majority of the year. In households with eligible youth, one youth was selected from each age range represented, but no more than a total of two youth were ever selected in one household. This had the effect of concentrating the youth interviews in a smaller number of households than would be expected if sampling were conducted independently for each age group.

To carry out this sampling efficiently, it was convenient to divide eligible households into three strata based on the combination of ages represented by the youth in the household. Because youth aged 12 to 13 were the rarest age domain, households containing such youth were always selected for the study. Households with youth aged 12 to 13 years were thus placed into a stratum by themselves. Youth aged 9 to 11 were the next rarest domain. Households that contained a youth aged 9 to 11 but no youth aged 12 or 13 were subsampled at Wave 1 and thus constituted a second stratum. (Unlike Wave 1, there was no subsampling of households with youth aged 9 to 11 but no youth aged 12 or 13 in either Waves 2 and 3; i.e., all such households were retained in the study, along with all households with youth aged 12

to 13.) Finally, households with youth aged 14 to 18 years (the most prevalent age group) were placed into a third stratum. Thus, the following strata were used for sampling purposes:

- Stratum A: households containing at least one youth aged 12 to 13;
- Stratum B: households containing at least one youth aged 9 to 11 but no youth aged 12 to 13; and
- Stratum C: households containing at least one youth aged 14 to 18 but no youth aged 9 to 13.

Table 2-2 summarizes the household retention rates employed in Round 1 by stratum and wave. The retention rates represent the percentage of the screened households of the given type that were included in the study. For example, in Wave 1, all of the screened households in stratum A were retained for NSPY, whereas only 70 percent and 45 percent of the households in stratum B and C, respectively, were retained. As mentioned previously, the retention rates for Waves 2 and 3 were modified in order to better meet the specified sample size goals. Thus, in these waves, all of the screened households in stratum A and B were retained for NSPY, while 45 percent of the households in stratum C were retained.

Table 2-2. Household sample retention rates for Round 1 by stratum and wave

Sampling stratum	Household composition	Retention rate (%)		
		Wave 1	Wave 2	Wave 3
A	At least one youth 12 to 13	100	100	100
B	At least one youth 9 to 11 but none 12 to 13	70	100	100
C	At least one youth 14 to 18 but none 9 to 13	45	45	45

2.3 Selection of Youth and Parents for Round 1

The procedure for Round 1 was to prepare a list of eligible youth in each eligible household and to randomly sample one youth within each nonempty age range, subject to a maximum of two sample youth per household. In a household with youth in all three of the age ranges, one youth from the 12-to-13 range was selected. A random decision was then made to select a second youth from either the 9-to-11

range or from the 14-to-18 range. Within an age range, all youth in the household had the same probability of selection.

Next, the interviewer determined the relationship of each adult in the household to the sampled youth. This information was then used to select a parent or caregiver for the study. For example, if one or two siblings were sampled, one adult from the set of parents or caregivers of either youth was randomly selected for the study. If two nonsiblings (such as cousins or children in multifamily households) were selected, one adult was selected from each set of associated parents and caregivers. For emancipated youth living separately from their parents, a caregiver was generally not required. Youth under age 19 who were serving in parental roles (e.g., an older sibling in a pair of orphans or a teenage stepmother) were considered ineligible for youth selection (but could be eligible for parent selection if their children were sampled for NSPY). All of the procedures used to select a parent/caregiver for the study were accomplished with the aid of computer software and a CAPI questionnaire.

Finally, it should be noted that youth residing in group quarters were not eligible for NSPY. In particular, youth living in boarding schools and college dormitories were excluded from the scope of the study. This exclusion was made because it was felt that dormitory residents could not be easily interviewed at their parents' homes. On the other hand, youth who lived at home or in private apartments while attending college were eligible for NSPY. The exclusion of dormitory residents required special weighting adjustments as described in Section 3.1.4.

The numbers of youth and parents who completed the NSPY interview during Round 1 are summarized in Table 2-3. Details about the response rate calculations are given in Section 4.9. Note that for purposes of disclosure avoidance, only a subsample of the youth indicated in Table 2-3 (along with selected variables from the parent data files) are included in the Round 1 PUF (see Section 3.2).

Table 2-3. Total number of completed youth and associated parent interviews by round

	Youth by age group					
Round	9-11	12-13	14-16	17-18	Total	Parents†
1	2,986	2,435	1,857	839	8,117	5,506
2	1,426	1,872	2,347	871	6,516	4,615
3	511	1,742	2,636	965	5,854	4,252
Total*	4,923	6,049	6,840	2,675	20,487	14,373

* The numbers of youth included in the PUFs are summarized in Section 3.2.

† Parents of youth who completed the youth interview.

2.4 Selection of Youth and Parents for Round 2

Round 2 was conducted in two waves covering the periods July 2001 through December 2001 (Wave 4) and January 2002 through June 2002 (Wave 5). Under the NSPY sample design, youth and parents selected for the initial recruitment waves (i.e., Waves 1 through 3) were retained for followup in subsequent data collection waves. No new samples of youth were selected for any of the followup waves (although it was possible to select a new parent for a previously responding youth if the youth no longer lived with the original parent).

For Wave 4, the first followup of Wave 1, all youth and parents in households that completed the screening enumeration in Wave 1 were included in the followup sample if the household contained at least one Wave 1 respondent (either youth or parent) and the youth was not older than 18 at the beginning of the followup wave. Note that under this selection criterion, a small number of youth and parents who were selected for Wave 1 but who did not complete a Wave 1 interview were refielded in Wave 4. For Wave 5, the first followup of Waves 2 and 3, a similar selection criterion was used.

The numbers of youth and parents who were interviewed during Round 2 are summarized in Table 2-3. Details about the response rates achieved in Round 2 are given in Section 4.9. For purposes of disclosure avoidance, only a subsample of the respondents indicated in Table 2-3 are included in the Round 2 PUF (see Section 3.2).

2.5 Selection of Youth and Parents for Round 3

Round 3 was conducted in two waves covering the periods July 2002 through December 2002 (Wave 6) and January 2003 through June 2003 (Wave 7). Wave 6 was the second followup of Wave 1, while Wave 7 was the second followup of Waves 2 and 3 combined. No new samples of youth were selected for any of the followup waves (although it was possible to select a new parent for a previously responding youth if the youth no longer lived with the original parent).

For Wave 6, only those youth who had not aged out of the sample (i.e., became 19 years old) and their parents who met the following criteria were refielded: (a) the youth/parent resided in a household in which at least one sampled youth completed either the Wave 1 or Wave 4 interview, and (b) the household was neither “not locatable” nor “out of area” in Wave 4. Under these selection criteria, a

small number of youth and parents who did not complete either the Wave 1 or Wave 4 interview were refielded in Wave 6. For Wave 7, a similar selection criterion was used.

The numbers of youth and parents who were interviewed during Round 3 are summarized in Table 2-3. Details about the response rates achieved in Round 3 are given in Section 4.9. For purposes of disclosure avoidance, only a subsample of the respondents indicated in Table 2-3 are included in the Round 3 PUF (see Section 3.2).

REFERENCES

Durbin, J. (1967). Design of multistage surveys for the estimation of sampling errors. *Applied Statistics*, 16, 152-164.

Judkins, D., Cadell, D.M., and Sczerba, K. (2000). *Costs and benefits of a permit sample late in the decade*. Presented at the Annual Meeting of the American Statistical Association.

Rizzo, L., and Judkins, D. (2004) *Replicate variance estimation for the national survey of parents and youth.* Expected to be published in Proceedings of the Survey Methods Research Section, American Statistical Association.

3. CALCULATION AND USE OF SAMPLE WEIGHTS

An analysis weight is provided for each responding youth included in each PUF. The weights are modified versions of existing cross-sectional weights that had been previously designed and constructed for analyses of individual waves of survey data. Since each PUF in the current data release incorporates results from two or more waves, the original weights had to be adjusted to reflect the expanded time period covered by the combined samples, and to compensate for the reduction in sample size resulting from the subsampling of the original respondents for purposes of disclosure avoidance. In addition, the nonresponse adjustments have been recomputed using the 2000 Census as the source of auxiliary data; the original adjustments were based on 1990 Census data, which was all that was available at the time. To improve sampling precision, the weights are calibrated to both independent and sample-based estimates of selected national totals. As background, an overview of the general NSPY weighting process is given in Section 3.1. The procedures used to subsample respondents for the PUFs are given in Section 3.2.1, and the procedures used to derive the corresponding sampling weights are described in Section 3.2.2. Finally, a discussion on the use of the sampling weights for analysis is given in Section 3.3.

3.1 NSPY Weighting Procedures

All of the previously calculated, wave-specific sampling weights were designed to compensate for differential selection probabilities and nonresponse. In addition, adjustments for undercoverage were made using a poststratification process called *ratio raking*, or simply *raking*. Under the raking process, the weights were adjusted in a manner that forced the weighted counts for important subgroups to agree with those from independent sources. The final weight for a respondent, after nonresponse and poststratification adjustments, may be viewed as the number of persons in the population that the respondent represents. Since the sampling weight reflects an individual's probability of selection into the sample, persons with small probabilities of selection generally have large weights and, conversely, persons with large probabilities have relatively small weights.

For multistage sample designs such as that employed for NSPY, the sampling weights reflect the probabilities of selection at each stage of sampling: PSU, segment within PSU, dwelling unit (DU) within segment, and person within DU. (In the case of the building permit sample, the DU within segment probability is the product of the probability of selecting the building permit within the building permit

office and the probability of selecting the dwelling unit within the selected permit.) Since the selection of persons could be carried out only in households where the screener was completed, the initial person-level weight included an adjustment for household nonresponse as described below in Section 3.1.1.

3.1.1 Dwelling Unit Weights

For each of the three waves composing Round 1, the base weight for a sampled dwelling unit was computed (within wave) as:

$$BW_{DU} = \frac{1}{\Pr\{\text{PSU}\}\ \Pr\{\text{segment} \mid \text{PSU}\}\ \Pr\{\text{DU} \mid \text{segment}\}}$$

where Pr{PSU} is the probability of selecting the PSU, Pr{segment|PSU} is the probability of selecting the segment within the PSU, and Pr{DU|segment} is the probability of selecting the dwelling unit within the segment.

Next, the dwelling unit-level base weights were adjusted for screener nonresponse in two phases. The first phase, referred to as the "doorstep nonresponse adjustment," was designed to compensate for those DUs for which it was not possible to determine whether eligible youth were present at the sampled address. For this adjustment, a dwelling unit was considered to be a respondent if information about the presence or absence of eligible children had been collected from either the occupants of the household or from their neighbors. The second phase adjustment, referred to as the "roster nonresponse adjustment," was designed to compensate for households known to contain eligible youth, but for which it was not possible to prepare a household enumeration and select a sample of youth (and parents). For this adjustment, a DU was considered to be a respondent if it contained any eligible children *and* an adult resident who was willing to provide an enumeration of the occupants of the household, their ages, and their relationships to the household members. Otherwise, the DU was classified as a nonrespondent.

The dwelling unit base weights were adjusted to compensate for unequal response rates across different groups in the sample. Special data mining software referred to as MART ("multiple additive regression trees") was used to form the groups required for the nonresponse adjustments (see Friedman, 1999, for a description of MART). The MART algorithm, which uses predictive data mining

techniques, was applied to identify subgroups that are homogeneous with respect to response propensity. About 60 block-group level variables describing characteristics of the neighborhood taken from the 2000 Population Census were specified as predictor variables in the MART software. Some examples of the variables used as predictors include percent of persons in various age groups, percent of persons of certain race and ethnic groups, percent of households that own (rather than rent) their homes, percent of persons residing in mobile homes, percent of persons who are U.S. citizens, and percent of households with incomes below the poverty level. Within the groups (referred to as "weighting classes") formed by the MART algorithm, the corresponding first-phase (doorstep) and second-phase (roster) weighted response rates were computed and used to obtain the nonresponse-adjusted dwelling unit weight. That is, the first-phase nonresponse-adjusted weight for the *i*th dwelling unit in first-phase weighting class *h* was computed as:

$$BW_{hi}^{Adj1} = F_h^{(1)} \; BW_{DU,i}$$

where $BW_{DU,i}$ is the base weight for dwelling unit *i* and $F_h^{(1)}$ is the reciprocal of the first-phase (doorstep) weighted response rate in class *h*. Next, using the adjusted weights calculated above, the final nonresponse-adjusted weight for the *i*th dwelling unit in second-phase weighting class *k* was computed as:

$$BW_{ki}^{Adj} = F_k^{(2)} \; BW_{ki}^{Adj1}$$

where BW_{ki}^{Adj1} is the first-phase nonresponse-adjusted weight for dwelling unit *i* in second-phase weighting class *k* and $F_k^{(2)}$ is the reciprocal of the second-phase (roster) weighted response rate in class *k*.

After adjustment for both phases of screening nonresponse, the weights were further adjusted for screening-based subsampling. As indicated in Section 2.2, dwelling units were randomly retained for the study at rates that depended on household composition. For example, dwelling units with at least one youth 12 to 13 years old were retained with certainty. On the other hand, dwelling units with youth 14 to 18 years old but no youth 9 to 13 years old were retained at a rate of 45 percent. Thus, in general, the final weight for DU *i* in stratum *C* (see Table 2-2 in Section 2.2) was computed as:

$$BW_{ci}^{Fin} = BW_{ci}^{Adj} \; / \; R_c$$

where BW_{ci}^{Adj} is the nonresponse-adjusted weight, for DU i in stratum C, and R_C is the corresponding retention rate.

3.1.2 Youth Weights

There were three age classes of special analytic interest for youth sampling purposes: 9 to 11, 12 to 13, and 14 to 18. If there were youth present in all three age classes, the first step in the youth subsampling was to select two out of the three age ranges. The 12-to-13 range was selected with certainty. One of the remaining two age ranges was selected with equal probability. Thus, the first component in the youth probability of selection for youth aged 9 to 11 or 14 to 18 in households with all three age classes was a factor of 0.5. Next, one youth was selected from each of the selected age ranges. For example, if there were four youth present in the given age range, the probability of selection within the range was 1/4 = 0.25. The two factors were multiplied together to create the probability of selecting the jth youth in dwelling unit i. $Pr\{\text{youth } j \mid \text{DU } i\}$. The initial youth weight for the jth youth in dwelling unit i was then calculated as the final dwelling unit weight divided by the within-dwelling unit probability of selecting the youth, i.e.,

$$IW_{ij}^{youth} = BW_i^{Fin} / Pr\{\text{youth } j \mid \text{DU } i\} .$$

Because no new youth were selected in Rounds 2 or 3, the initial weights calculated for the recruitment waves were also the initial weights for youth in the followup rounds. As described below, the wave-specific youth base weights were adjusted for nonresponse and then ratio-adjusted to independent population totals corresponding to each wave.

3.1.3 Nonresponse Adjustments

Responding youth were defined to be those who answered the last question on general ad exposure (D13 in the youth questionnaire). Nonresponding youth were those who broke off the self-administered data collection prior to this point, or else failed to start the data collection. Nonrespondents included sampled youth whose parents refused consent or otherwise failed to provide consent, those who refused personal assent, and those who were never reached to do the interview for any reason. Among those who did not complete the questionnaire, a difference was drawn between those who were physically

or mentally incapable of completing the interview and those who simply chose not to. The first group was considered to be ineligible sample youth rather than nonresponding sample youth. The distinction matters because the weight of ineligible youth was not redistributed to responding youth through the nonresponse adjustment. Included in the category of ineligible youth were those who could not communicate in English or Spanish. Since the television and radio components of the Media Campaign (to be analyzed in the Evaluation) were broadcast in these languages, persons who could not communicate in either language were considered to be ineligible for the Evaluation. Also included in the ineligible youth category were young people who held parental roles for other youth aged 9 to 18. This might occur by reason of marrying an older person with such youth or by reason of caring for younger siblings.

The set of 60 block group-level variables used for doorstep and roster nonresponse adjustment, plus selected youth and parent characteristics (i.e., age, gender, and race/ethnicity) and characteristics on household composition, were used in conjunction with the MART software to develop a set of weighting cells for youth nonresponse adjustment. The household-level characteristics included items such as whether both parents lived in the household, whether the youth was an only child, the number of youth living in the household, and whether there was an unrelated person living in the household. All of these variables were obtained from the household enumeration. The set of nonresponse adjustment cells determined by the MART algorithm was then used to adjust the weights of the responding youth. That is, the nonresponse-adjusted weight for responding youth j in weighting class g was computed as:

$$W_{gj}^{NR} = G_g \, IW_{gj}^{youth}$$

where IW_{gj}^{youth} is the initial youth weight and G_g is the reciprocal of the weighted response rate in weighting class g.

3.1.4 Poststratification (Raking) Adjustments

Raking is a form of poststratification adjustment that is commonly used to calibrate survey estimates to known population totals. The goal of raking is to reduce biases due to undercoverage and nonresponse, and to reduce the sampling error of the estimates. Raking may be thought of as an iterative form of poststratification in which the weights are successively ratio-adjusted to multiple sets of marginal control totals (referred to as “raking dimensions”) until the resulting weighted sums equal the control

totals specified for each dimension. The sample sizes associated with the levels of each raking dimension are the important determinants of the stability of the raking procedure, not the cells formed by a complete cross-classification of the variables defining the raking dimensions. This permits the use of more auxiliary variables than would be feasible with traditional or "direct" poststratification. For this reason, raking was used to calibrate the youth weights rather than direct poststratification.

The following two dimensions were defined for the poststratification (raking) adjustments:

- Gender (male, female) crossed by three age groups (ages 9 to 11, 12 to 13, and 14 to 18); and
- Race/Ethnicity (non-Hispanic–non-Black, non-Hispanic-Black, Hispanic) crossed by three age groups (ages 9 to 11, 12 to 13, and 14 to 18).

Independent estimates of the total 9- to 18-year-old civilian population by gender, age group, and race/ethnicity were obtained from the Current Population Survey (CPS) for the relevant time periods covered by each wave. These population totals were then adjusted downward to exclude the civilian noninstitutional group quarters population (e.g., students in dormitories), using estimates derived from the 1990 Census Public Use Micro-data System (PUMS) files. The resulting adjusted population totals were used as control totals in the raking process.

After the youth nonresponse-adjusted weights were calculated as described earlier, the final ("raked") youth weights were computed as:

$$W_{kj}^{Fin} = H_k W_{kj}^{NR}$$

where W_{kj}^{NR} is the nonresponse-adjusted youth weight, and H_k is the final adjustment determined by the iterative raking process.

3.1.5 Calculation of Replicate Weights for Variance Estimation

The sampling errors of estimates derived from NSPY data have been calculated using a replication approach (e.g., see Hornik, et al., 2003). This replication method uses 100 replicates to measure the variance of the full sample estimates. The method developed for NSPY reflects the variance

due to each stage of sampling (i.e., selecting PSUs, segments within PSUs, dwelling units within segments, and persons within dwelling units), as well as finite population correction factors at both the PSU and segment levels. Full technical details of the method developed for NSPY are given in Rizzo and Judkins (2004).

To permit the calculation of sampling errors, a series of replicate weights were calculated and included in the data files. The first step was to construct 100 replicates as described in Rizzo and Judkins (2004). After the replicates were created, the full set of weight adjustment procedures described in the previous sections was applied to each replicate. This meant that each set of replicate weights was adjusted for nonresponse and raked to specified CPS-based control totals. As a result, the replicate weights reflect all of the adjustments used to create the full sample weights. The result of this process is that each youth record in the data file has 101 different weights, an overall weight, and 100 replicate weights.

3.2 Subsampling for Disclosure Avoidance

A random subsample of responding youth was selected for inclusion in each PUF. Subsampling was used to reduce disclosure risk by introducing uncertainty as to whether data for a particular youth can be found in the PUF. As described in Section 3.2.1, the subsample was drawn in a manner designed to reduce the variation in weights to the extent feasible. The subsample was then reweighted according to the procedures described in Section 3.2.2.

3.2.1 Subsampling Procedures

The complete sample of NSPY youth (respondents) was stratified by age category, and two-thirds of the sampled youth in each category were selected for the PUF. The age categories used for Round 1 were 9 to 11, 12 to 13, 14 to 16, and 17 to 18. The age categories for Rounds 2 and 3 are the same, except that youth aged 9 to 11 are excluded from the PUF for these rounds. Prior to subsampling, the youth were sorted by age, gender, and race category (white/non-white) within each age category. The subsamples were then selected systematically by round, ensuring appropriate representation by age, gender, and race.

Since the youth weights were originally created by wave (see Section 3.1), and not by round, the original weights were adjusted by dividing by the number of waves within each round. For example, the weights for all youth in Round 1 were divided by three since there were three waves of data collection in Round 1. In this way, the "round-adjusted" weights sum to the expected number of youth in the United States, instead of triple the number. These weights were then recalibrated to the corresponding CPS control totals by single years of age (within gender and broad race/ethnicity groups). This recalibration was designed to provide better representation of the sample for each single year of age.

The adjusted (recalibrated) weights were then used as a "measure of size," and youth were selected for the subsample with probabilities proportional to this size measure (PPS). Thus, although two-thirds of the responding NSPY youth were retained for the PUF at each round, the retention rates varied between youth. After drawing the subsample, the weights of the selected youth were recomputed to compensate for the subsampling. The initial weights for the youth included in the public use file were calculated as the original adjusted weight divided by the probability of selection for the subsample. A consequence of the use of PPS sampling was to reduce the variation in the weights. Table 3-1 shows how many youth were on the original files, and the corresponding numbers included in each PUF, by round and age group.

Table 3-1. Numbers of respondents in the original sample and numbers included in PUF

Round	Age category	Number in original sample	Number selected for PUF
1	9-11	2,986	1,991
	12-13	2,435	1,623
	14-16	1,857	1,238
	17-18	839	559
	Total	8,117	5,411
2	12-13	1,873	1,248
	14-16	2,346	1,565
	17-18	871	581
	Total	5,090	3,394
3	12-13	1,742	1,161
	14-16	2,636	1,757
	17-18	965	643
	Total	5,343	3,561

3.2.2 Weighting the PUF Subsample

To account for the subsampling described above, the weights of the selected respondents were recalibrated using the raking algorithm to the following seven dimensions:

- Gender by age;
- Gender by broad race/ethnicity (Hispanic persons, non-Hispanic black persons, non-Hispanic persons of all other races);
- Sensation seeking (high/low);
- Risk score (high/low);
- Cigarette use (yes/no);
- Alcohol use (yes/no); and
- Marijuana use (yes/no).

For the first two raking dimensions, the weights were adjusted (raked) to CPS-based control totals. For the remaining five dimensions, the weights were adjusted to sample-based control totals derived from NSPY. Thus, unlike the CPS-based control totals, the sample-based control totals for the last five raking dimensions varied by replicate to reflect the variance of the corresponding weighted counts derived from the full NSPY sample (i.e., prior to subsampling). Table 3-2 summarizes unweighted and weighted counts of youth included in the PUF subsamples by selected characteristics.

3.3 Use of Weights

The final (round-specific) weights to be used for analysis of data contained in each of the PUFs are named WEIGHT, REPLW1, REPLW2, ..., REPLW100. WEIGHT is the final overall raked weight described in Section 3.1.5. REPLW1-REPLW100 are the corresponding replicate weights to be used for variance estimation.

Table 3-2. Unweighted and weighted counts of respondents included in PUF by selected characteristics

Characteristic	Round 1		Round 2		Round 3	
	Number in PUF	Weighted estimate	Number in PUF	Weighted estimate	Number in PUF	Weighted estimate
TOTAL	5,411	40,221,300	3,394	27,884,705	3,561	28,476,700
Gender/age group						
Male 9-11	1,030	6,541,032	—	—	—	—
Male 12-13	827	4,091,265	636	4,199,367	591	4,340,070
Male 14-16	623	6,095,851	801	6,182,674	915	6,290,588
Male 17-18	277	3,873,483	297	3,879,947	316	3,936,413
Female 9-11	961	6,205,114	—	—	—	—
Female 12-13	796	3,908,713	612	4,067,373	570	4,124,654
Female 14-16	615	5,789,824	764	5,863,898	842	6,001,382
Female 17-18	282	3,716,018	284	3,691,445	327	3,783,592
Race/ethnicity						
Non-black, non-Hisp.	3,774	27,977,758	2,397	19,421,702	2,505	19,574,511
Black, non-Hispanic	833	6,271,642	490	4,341,104	510	4,433,678
Hispanic	804	5,971,900	507	4,121,899	546	4,468,511
Sensation seeking						
High	2,331	18,691,695	1,797	15,236,691	1,901	15,346,497
Low	2,881	20,128,623	1,528	12,068,606	1,605	12,708,844
Missing	199	1,400,982	69	579,408	55	421,359
Risk score						
High	978	10,177,328	998	10,113,873	1,067	9,701,884
Low	3,808	25,204,061	2,040	14,792,102	2,045	15,144,759
Missing	625	4,839,911	356	2,978,730	449	3,630,056
Cigarette use						
Yes	653	6,864,645	636	6,392,593	662	5,917,503
No	4,720	33,110,819	2,747	21,337,652	2,881	22,401,590
Missing	38	245,836	11	154,460	18	157,607
Alcohol use						
Yes	1,069	10,528,580	1,106	10,336,405	1,151	10,270,704
No	4,303	29,435,245	2,273	17,422,674	2,396	18,093,174
Missing	39	257,495	15	125,626	14	112,822
Marijuana use						
Yes	401	4,555,069	459	4,416,994	527	4,846,821
No	4,972	35,330,506	2,926	23,332,377	3,024	23,547,876
Missing	38	335,724	9	135,333	10	82,003

The final raked weight, WEIGHT, should be used to obtain round-specific national estimates. For example, if y_i is the reported value of a survey item for the ith respondent in the PUF, and w_i is the corresponding value of WEIGHT, then a weighted sum of the form:

$$\hat{y} = \sum_{i=1}^{n} w_i y_i$$

provides an estimate of the corresponding total value of the characteristic in the population. If y_i takes on values of 1 or 0 indicating the presence (1 = "yes") or absence (0 = "no") of a characteristic, then the weighted sum, $\hat{y}$, provides an estimate of the number of persons in the population that have the given characteristic.

The corresponding replicate weights, REPLW1, REPLW2, ..., REPLW100, are provided in the PUF to permit the calculation of sampling errors using replication techniques. (The replicate weights should not be used to make national estimates.) For example, software packages such as WesVar and SUDAAN can be used for this purpose.

Since the replicates used in NSPY were specifically designed to reflect certain features used in PSU selection, it is necessary to apply a set of replicate-specific factors in the calculation of the sampling variances. These factors (defined as h_k in the formula below) are: 2.57 for replicates 1-60, and 0.06 for replicates 61-100. In WesVar, these factors can be imported or entered using the "JKN" option. SUDAAN will also accept replicate weights and user-supplied factors for variance estimation.

As an example, suppose that $\hat{y}$ is the full-sample estimate, and let $\hat{y}_k$ be the estimate for replicate k using the corresponding replicate weights (k = 1, 2, ..., 100). Then the standard error of $\hat{y}$ is computed using the formula:

$$SE(\hat{y}) = \sqrt{\sum_{k=1}^{100} h_k (\hat{y}_k - \hat{y})^2}$$

where h_k is the specified factor for the kth replicate.

Table 3-3 summarizes weighted estimates and associated standard errors for selected variables included in this PUF.

Table 3-3. Selected estimates and standard errors

Characteristic	Round 1		Round 2		Round 3	
	Weighted estimate	Standard error	Weighted estimate	Standard error	Weighted estimate	Standard error
Sensation seeking						
High	18,691,695	293,642	15,236,691	207,275	15,346,497	215,856
Low	20,128,623	322,585	12,068,606	217,489	12,708,844	214,082
Missing	1,400,982	85,471	579,408	57,574	421,359	64,870
Risk score						
High	10,177,328	199,075	10,113,873	208,233	9,701,884	196,741
Low	25,204,061	261,619	14,792,102	187,976	15,144,759	223,917
Missing	4,839,911	217,493	2,978,730	151,015	3,630,056	209,241
Cigarette use						
Yes	6,864,645	221,652	6,392,593	184,860	5,917,503	193,467
No	33,110,819	222,992	21,337,652	181,507	22,401,590	194,908
Missing	245,836	41,001	154,460	44,499	157,607	34,690
Alcohol use						
Yes	10,528,580	263,483	10,336,405	221,313	10,270,704	273,860
No	29,435,245	275,186	17,422,674	224,511	18,093,174	274,895
Missing	257,495	51,624	125,626	30,972	112,822	25,377
Marijuana use						
Yes	4,555,069	191,598	4,416,994	204,862	4,846,821	182,633
No	35,330,506	196,761	23,332,377	202,148	23,547,876	188,913
Missing	335,724	60,193	135,333	36,958	82,003	26,723

NOTE: Standard errors were computed using WesVar. However, other software packages that accept replicate weights and related factors will produce identical results.

REFERENCES

Durbin, J. (1967). Design of multi-stage surveys for the estimation of sampling errors. *Applied Statistics*, 16, 152-164.

Friedman, J. (1999). *Tutorial: Getting Started with MART in Splus.* Stanford University.

Judkins, D., Cadell, D.M., and Sczerba, K. (2000). *Costs and benefits of a permit sample late in the decade*. Presented at the Annual Meeting of the American Statistical Association.

Korn, E.L., and Graubard, B.I. (1999). *Analysis of health surveys*. New York: Wiley.

Rizzo, L. (2000). Memorandum from Lou Rizzo to David Judkins. *NSPY variance estimation.*

Rizzo, L., and Judkins, D. (forthcoming) *Replicate variance estimation for the national survey of parents and youth. Expected to be published in Proceedings of the Survey Methods Research Section, American Statistical Association.*

Miller, D.R., and Hodges, K. (1994). *A population density approach to incorporating an urban-rural dimension into small area lifestyle clusters*. Presented at the Annual Meeting of the Population Association of America.

Weiss, M.J. (2000). *The clustered world: How we live, what we buy, and what it all means about who we are*. Boston: Little, Brown and Co.

4. DESCRIPTION OF DATA COLLECTION INSTRUMENTS AND METHODOLOGY

4.1 Overview

The primary tool for the Evaluation is the National Survey of Parents and Youth (NSPY). The NSPY is a longitudinal, nationally representative household-based survey of youth and their parents, being conducted in 90 locations across the United States. Figure 1-1 in Section 1.2 is a graphical depiction of the NSPY design. Round 1 data were collected between November 1999 and May 2000. During Round 1, youth and parents were recruited for the study and interviewed. During Round 2 and Round 3, data were collected between July 2001 and June 2002, and July 2002 and June 2003, respectively. Data for Round 4 (not covered by this report) will be available in late 2004. Respondents who were identified in Round 1 were tracked and recontacted for interview in the followup rounds.

This chapter describes the data collection methodology used during the survey. Topics include a discussion of the data collection instruments, media activities, field staff recruitment and training, procedures used during data collection, quality control procedures, data editing and cleaning procedures, as well as youth cross-sectional response rates for each round.

4.2 Data Collection Instruments

In preparation for the Phase III Evaluation, NIDA convened an expert panel to assist in the development of the data collection instruments. This group, which included specialists in adolescent drug use prevention and parenting behaviors, met and generated draft instruments for children (aged 9 to 11), teens (aged 12 to 18) and parents. NIDA shared these Phase III prototypes with Westat at the beginning of the contract period.

Westat formed an instrument development team whose members included evaluation experts from Westat, the Annenberg School for Communication at the University of Pennsylvania, and the National Development and Research Institutes (NDRI). This team reviewed the Phase III prototypes, as well as the survey instruments used in the Phase II Media Campaign Evaluation, and other surveys, including Monitoring the Future (MTF), Community Action for Successful Youth, National Household

Education Survey (NHES), and the National Household Survey on Drug Abuse now known as the National Survey on Drug Use and Health (NSDUH).

To facilitate the development of the instruments, the team developed a behavioral change model for the Evaluation and mapped each question back to this model, as well as to the communication objectives that had been established for the Media Campaign.

Question domains for parents include the following:

- Media consumption;
- Past discussions with child about drug attitudes and avoidance strategies;
- Past child monitoring behaviors;
- Self-efficacy of discussing drugs with child and of monitoring the child's actions;
- Belief that the child is at risk for drug use;
- Belief that drug use has bad consequences;
- Exposure to the Media Campaign's advertising, including brand recognition;
- Parent's own current and past use of tobacco, alcohol, and drugs; and
- Demographic information.

Youth question domains include the following:

- Exposure propensity to media;
- Youth's own current and past use of tobacco, alcohol, marijuana, inhalants, and Ecstasy;
- Past discussions with and communication of anti-drug messages from parents and friends;
- Expectations of others about respondent's drug use;
- Knowledge and beliefs about the positive and negative consequences of drug use;
- Exposure to the Media Campaign's advertising;
- Family and peer factors;

- Personal factors; and
- Demographic information.

During Round 1, virtually the same set of questions was asked of respondents. However, some new questions were added in later waves within the round. They included a question on Campaign brand recognition in the Teen and Parent instruments; questions about Ecstasy use in the Teen instrument (have used and when last used); questions about doing fun things with parents in the Teen and Child instruments; and a question about parents' perception of the efficacy of drug talk to the Parent instrument. To make room for these questions, other questions were deleted. They included questions about reading magazines or seeing television shows from the Teen and Parent instruments, questions about communicating rules for alcohol and smoking from the Teen and Child instruments, and a question about perceived consequences of inhalant use from the Child instrument.

In Round 2, the instruments were essentially the same for parents, teens, and children, except for some additional questions on Ecstasy in the Teen instrument, including intention to use, perceived expectations of use by peers, and attitudes of use including approval/disapproval of use and perceived harm of use. A question about Campaign banner ads on the Internet was added to both the Teen and Parent instruments. In the Parent instrument, the branding question was rephrased to ask about the correct parent brand and one of the two "ringer" brands, mirroring the format of the Teen branding question. Other additions to the Parent instrument included a question about the presence and number of youth in the household in the age categories of interest; a question on parental perceptions of harm from trial use of marijuana, inhalants, and Ecstasy; and a question on the likelihood of youth use of inhalants and Ecstasy.

In Round 3, the instruments for teens and children were essentially the same, except for the questions pertaining to the Evaluation television ads. Prior to Round 3, teen and child respondents were asked followup evaluative questions about the first three television ads that they were shown, only if they answered that they had recalled having seen or heard them in the past. In Round 3, the Teen and Child instruments were changed and respondents were asked to evaluate the first three television ads they were shown, regardless of whether they recalled having seen the ads in the past. No changes were made to the Parent instrument.

Appendix B contains annotated instruments for the Child and Teen in Round 1 and the Teen in Rounds 2 and 3. In addition, a few questions that were asked of the parents are included in the appendix. The parent variables represented by these questions are included in the youth data records.

During Round 1, a brief, hard-copy household screening instrument was used to determine a sample household's eligibility. All other data were collected using a laptop computer and a combination of computer-assisted interview technologies. Computer-assisted personal interview (CAPI) involved having the interviewer read the questions to the respondent and record the answers in the computer. In Round 1, CAPI was used to enumerate the household and select a parent/caregiver and one or two youth. In Rounds 2 and 3, CAPI was used to determine respondent eligibility and to select a new parent, if appropriate. CAPI also was used for the nonsensitive questions in the extended interview (parent, teen, and child) instruments in all rounds. For collection of sensitive data in the extended interview, audio computer-assisted self-interview (ACASI) technology was employed. This allowed respondents to self-administer the survey in total privacy. They listened to the question on headphones and recorded their own responses by touching the computer screen. These technologies were selected based on the theory that providing respondents with a methodology that improved privacy and confidentiality would make reporting of potentially embarrassing, stigmatizing, or illegal behaviors (such as drug use) less threatening, and enhance response validity and response rates.

4.3 Media Activities

Because this is an evaluation of a media campaign, activities such as media buying, ad creation, and broadcast levels played key roles in the instruments as well. Because the Media Campaign is dynamic over time, the media-specific questions in the instruments also changed appropriately.

In the Child, Teen, and Parent instruments, some questions were asked about the respondent's media usage patterns, including television, radio, and magazines. All NSPY instruments contained a section of questions devoted to how the respondent received anti-drug messages. In these questions, selected television and radio Media Campaign ads that were broadcast during the prior 2 calendar months were played for the respondent. Questions were then asked about the respondent's recall of prior exposure (viewing or listening) to the ad, and his/her assessment of the ad's message and impact. The set of television and radio ads that were played for respondents changed monthly, with a set protocol being used to determine which ads were played during each month and for which respondents.

Each month Ogilvy, the Media Campaign media buy contractor, produced an updated copy rotation schedule. This schedule outlined, by month, each ad that was slated for broadcast, its target audience (parents or youth), and racial or ethnic group (general market, African American, or Hispanic). Included were each ad's planned broadcast dates and the Media Campaign behavioral platform that the ad addressed. As ads were produced, Ogilvy forwarded them to Westat for digitizing; a process that put the ads into an electronic format that was then incorporated into the computerized laptop questionnaires.

Using the copy rotation schedule, Westat determined the television and radio ads that had aired over the proceeding 2 months. A CD containing those ads was then produced and sent to the field interviewing staff. A look-up table was also developed for each interview month and transmitted to the field staff. It provided the specifications for ad selection and randomization for each respondent that month.

During Round 1, questions were asked about viewership of specific television shows and readership of specific magazines from which Ogilvy purchased advertising time or space. The specifics of these media buys were determined based on the Gross Rating Points (GRPs) that the television shows, radio programs, or magazines were expected to earn. Ogilvy sent updated information on those television shows and magazines, for which ad time or space had been purchased, to Westat every 3 months, and appropriate updates were transmitted to the field interviewers' laptop. (Questions on specific television shows and specific magazine readership were dropped from the instruments during Round 1.)

Ogilvy also provided data regarding the planned GRP levels for the previous quarter, by target audience (parents or youth), creative ad execution, media (television, radio, print, and out of home), and week/month. GRPs refer to the percentage of the target population who are estimated to be watching a particular TV show, listening to a specific radio program, or reading a certain magazine, and are therefore exposed to the advertising messages provided. These GRPs are based on data from that media's audience ratings company(s) (Nielsen Media Research for television, Arbitron Research and RADAR for radio, and MRI for print). Knowing the reach and frequency objectives for the Media Campaign's messages, the media buyers then purchase a mix of media whose GRPs, when aggregated, should achieve the desired intensity of the Media Campaign message exposure. This information was used by the Evaluation's analysts to look for correlation between recalled exposure to ads by respondents and the ads' reach and frequency levels.

4.4 Field Staff Recruiting and Training

The NSPY initial data collection design was based on hiring one primary interviewer in each of 90 primary sampling units (PSUs) and hiring additional interviewers to supplement efforts in larger PSUs, PSUs geographically clustered, and in PSUs where primary interviewers quit during the field period.

Over the rounds, interviewers were recruited from Westat's pool of experienced interviewers. Additional candidates were recruited through local organizations, and classified newspaper advertisements were placed in various PSUs as needed. These candidates were screened for communications skills and availability. Spanish language interviewer candidates were screened by bilingual project staff for their ability to communicate effectively in both Spanish and English. On average, 12 percent of the total interviewers hired were bilingual. Most English and bilingual candidates had prior experience relevant for data collection.

Over the rounds, all interviewers participated in an 8- to 10-day training session. The training program, which was staffed by qualified project staff and field supervisors, was designed to ensure consistency in data collection through the use of lectures, with a heavy focus on practice sessions. Trainees new to Westat attended an additional half-day training on general interview techniques. Bilingual trainees also attended an additional half-day training that concentrated on reviewing bilingual scripts and materials.

Supervisors received training and refresher training prior to the interviewer trainings. Supervisor trainings included an overview of field staff responsibilities and sessions on how to set up the field office; how to make assignments and monitor interviewer progress and work; quality control procedures, including editing materials, shipping materials to Westat, and validating special cases in the field; procedures for time and expense reporting; and how to use the management systems.

4.5 Round 1 Data Collection

The following sections discuss the procedures used in the Round 1 data collection phase of the NSPY, which took place from November 1999 through June 2001.

4.5.1 Advance Materials

About 2 weeks prior to interviewers receiving their assignments, advance letters were mailed to sampled addresses. Since these addresses were from listings, the envelopes and letters were not addressed to a person, but to "resident." The envelopes contained letters on NIDA stationary that provided an introduction to the study, information on confidentiality, and notice that an interviewer would contact the household to determine whether household members would be eligible to participate in an interview. Each person who completed an interview received a $20 incentive.

4.5.2 Determine Household/Respondent Eligibility

During Round 1, interviewers were required to make up to five in-person attempts to contact a household. A household was considered eligible if two criteria were met. First, the household must contain children of an age group specified for that household (age groups included households with children aged 9 through 13, 12 and 13, or 9 through 18). Second, the housing unit must have been built before April 1, 1990, be a mobile home, or be selected through the permit sample (see Section 2.1). All eligibility information was collected on hard copy and then entered into an electronic file on laptop computers.

During Round 1, interviewers were instructed to visit the sampled household three times to try to determine eligibility, prior to obtaining eligibility information from a neighbor. This procedure was changed for a short period of time during Round 1 to allow interviewers to determine eligibility information from neighbors after one attempt to contact the household. Because a neighbor might be less able to accurately know the exact ages of children, two questions about children were asked. First, the neighbor was asked whether any children aged 9 to 18 lived in the household. If yes, a followup question was asked to determine whether children of the specified age for the particular household (see categories above) lived in the household. In addition, the neighbor was asked if sampled housing units in area segments were built after April 1, 1990. Finally, the neighbor was asked what times members of the sampled household would be likely to be at home. If answers to both of the age questions were no, the household was considered ineligible. If the answer to either or both age questions was yes and if the housing unit was built before April 1, 1990, or if the housing unit was drawn from the permit sample, the interviewer continued to try to contact the sampled household. Remaining attempts were made to contact the sampled household to obtain an interview at times suggested by the neighbor.

4.5.3 Respondent Selection

During Round 1, the interviewer conducted a household enumeration with a household member 18 years of age or older, once a household was determined to be eligible. All members of the household, excluding children/students who were currently away from home or living at a boarding school or college, were enumerated, i.e., the first name, age at last birthday, and gender were collected. At this point, up to two eligible youth were randomly selected and the type of survey instrument, i.e., Teen and Child, to be administered to the youth was determined. Once the youth were selected, the relationship of every other person to the selected youth was obtained. One or two parents or primary caregivers were then selected based on a predetermined algorithm. (Two parents or primary caregivers were chosen only in the unusual situation where the selected youth were not siblings.) If two parents for a selected child resided in the household, the algorithm selected the male or female parent on a random basis. If one of the parents was a stepparent or foster parent, that parent must have lived with the child in the household for at least 6 months to be eligible for selection. If no parents lived in the household, the algorithm selected a primary caregiver. Once all respondents were selected, information on the race and ethnicity for each selected person was obtained.

4.5.4 Obtaining Parent Consent/Child Assent

The next step was to obtain parental consent and youth assent to conduct the youth interview. For each child aged 18 or younger, interviewers obtained written parental consent from the selected youth's parent or a legal guardian. Once done, written assent from each youth participant was obtained. If the parent refused to let the child be interviewed and refusal conversion efforts failed, the child could not be interviewed.

In addition, verbal consent was required for the parent interview. Once the parent answered yes to the verbal consent within the CAPI portion of the instrument, the extended interview continued.

Since the NSPY included sensitive questions, especially ones related to illegal behaviors (i.e., drug use), NIDA requested and obtained a Certificate of Confidentiality from the Department of Health and Human Services. The Certificate was discussed during the consent process.

4.5.5 Conducting the Interview

The extended interviews were administered in a CAPI and ACASI format. Sensitive questions were in ACASI format, which meant that respondents used the computer themselves to answer questions by touching the screen, and used headphones to hear the questions. The extended interview was programmed so that the interviewer was unable to go back into the interview and look at answers the respondent provided in the ACASI section.

Interviewers were instructed to, if possible, seat the respondent in a chair that was against the wall or a piece of furniture, so that no other person could stand or pass behind the respondent. This procedure hindered third parties from being able to observe the respondent's answers during the ACASI part of the interview. The interviewer also requested that parents not be present in the room while the interview was being conducted with the youth. If the parent insisted on being present in the room, the interviewer asked the parent not to stand directly behind the child during the ACASI portion of the interview.

4.6 Rounds 2 and 3 Data Collection

As discussed in Chapter 2, to be included in Round 2, a household must have had at least one selected person (parent or youth) complete his or her extended interview in Round 1. If no one who was selected in the household completed an interview in Round 1, then the household was not included in Round 2.

To be included in Round 3, a household must have had at least one selected youth complete his or her extended interview in Round 1 or in Round 2, and the household must not have been classified as "not locatable" or out of area in Round 2. If no selected youth completed an interview in Round 1 or Round 2, the household was not included in Round 3.

4.6.1 Tracking

Prior to Round 2, efforts were made to verify the location of Round 1 adult respondents. Address correction information was received from the U.S. Postal Service from Thank You Notes sent

out by Westat that were forwarded to new addresses by the Postal Service, and from calls received on the NSPY information line. Location information (i.e., address and telephone number) about respondents was sent to a national database company for tracking purposes. A high proportion of the new addresses provided by the database company had been also reported by the Postal Service or on the NSPY information line. Updated location information from these sources was sent to Westat's Telephone Research Center (TRC) and telephone interviewers placed calls to these households to verify the identity of respondents. Again, a high proportion of these households were contacted and respondents verified at their new addresses.

Prior to Round 3, efforts were made to update the location of Round 2 adult respondents. Again, address correction information was received from the U.S. Postal Service from Thank You Notes sent out by Westat that were forwarded to new addresses by the Postal Service, and from calls received on the NSPY information line. Address changes from these two sources were updated in the study's database and this information was sent to the TRC for followup. The proportion of reported address changes was very small in Round 3. Therefore, it was decided that it was not necessary to use the services of a database company for tracking purposes.

During Rounds 2 and 3 data collection, tracking of respondents continued. U.S. Postal Service address corrections from advance letters were forwarded to field interviewers to assist in locating respondents. The field interviewer attempted to locate a respondent using telephone tracing procedures, which included contacting directory assistance and calling the two tracing references. If telephone tracing was unsuccessful, the field interviewer then conducted in-person tracing, which included visiting the last known address of the parent respondent and two neighbors. When field tracing failed to locate respondents, the case was then returned to the home office for tracing by the TRC using directory assistance, criss-cross directory, and national internet database searches.

4.6.2 Advance Materials

As in Round 1, advance letters were mailed to respondents about 2 weeks prior to interviewers receiving their assignments. Envelopes and letters were addressed to the selected parent respondent, regardless of whether or not he or she completed the interview in the prior round. The letter was again on NIDA stationary, and the content of the letter was the same as in Round 1.

4.6.3 Confirming Respondent Eligibility

In Rounds 2 and 3, interviewers were allowed to screen households both by telephone and in person. Interviewers were required to make up to five telephone attempts to contact a household. If the telephone attempts were not successful, up to five in-person attempts were then made. Most first attempts were made by telephone; however, first attempts at contact were made in person if the selected parent had refused to complete his or her initial interview or if the interviewer did not have a telephone number to call.

A youth who had been selected at Round 1 and included in the Rounds 2 and 3 sample (see above) was considered eligible for Rounds 2 and 3 if he or she were 9 to 18 years old at the time of the Rounds 2 or 3 interview and not living in a group quarters situation (that is, not living away from home at school or in an institution). A parent or caregiver who had been selected at Round 1 was considered eligible for future rounds if he or she were still living with an eligible sampled youth at least 2 nights a week and not physically or mentally disabled. A new parent was chosen for interview in Rounds 2 or 3 if either of these two conditions were not met by the selected parent from the previous round.

4.6.4 Respondent Selection (New Parent)

If the parent who completed an interview in the previous round was not eligible, a new parent was selected. During the CAPI screening instrument, if it were determined from the screening answers that the parent was no longer eligible, a second screening module was initiated in order to select a new parent. The new parents selected during Rounds 2 and 3 were self-selected; that is, were defined as the adult living with the youth who knew the most about the child's daily life and activities, as opposed to the computer's random selection of the parent in Round 1.

4.6.5 Obtaining Parent Consent/Child Assent

Procedures for obtaining parent consent and youth assent to conduct the interviews are the same for Rounds 2 and 3, as for Round 1 (see Section 4.5.4).

4.6.6 Conducting the Interview

Procedures for conducting the interview are the same for Rounds 2 and 3, as for Round 1 (see Section 4.5.5).

4.7 Quality Control

Data collection quality control efforts began with the development and testing of the interview instruments and continued through validation of parent interviews and of ineligible households. This section describes the steps taken to ensure quality data, including testing of instruments; pilot testing of field procedures and data collection instruments; training procedures; and the validation of data collected.

4.7.1 Testing and Implementation of the Instruments

Using the detailed specifications as a baseline, automated systems and instrumentation were tested before field data collection started. A comprehensive and iterative testing approach was developed that included a plan for testing and tracking modifications of all systems and processes.

This plan used an iterative development process in which systems were developed and systematically tested to ensure that all identified errors were corrected and retested before use in the field. Software was run through multiple phases of testing for errors in instrument skip patterns, hard and soft error processing, and the transfer of completed interviews and related statuses. The testing process included module testing, instrument testing, subsystem testing, scenario testing, and integrated systems testing. Test scripts and scenarios were developed along with instrument specifications that were used as a reference for all levels of the testing process.

Programmers were primarily responsible for the initial module or unit testing of the applications. A unit is a defined function or an area of an interview in which complicated skip patterns or programming logic is implemented. These area testers tested several instrument sections or modules

which are related and are administered in conjunction with each other. This included testing one or more sections of the instrument as each section was developed and modified.

Instrument testing was done by developers, systems analysts, and project staff to test the entire questionnaire instrument from start to finish. This included checking all skip patterns, response categories, range checks, and both hard and soft edits; cross-checking items collected in multiple instruments; and reviewing instrument text for appropriate language, interviewer instructions, question displays, and scrolling issues.

Subsystem testing was completed by systems analysts, lead developers, and project staff to ensure that all interviewing and management functions worked appropriately together and the correct algorithms for selection, along with the correct ads to be played were applied. Testers reviewed application flow as well as population of response and status data in the database. In addition, this testing included testing the system security at the browser and server levels.

Scenario testing included the development and testing of the automated instruments for various age, race, gender, and other skip patterns and followed the interviewing process through predefined paths similar to what may be encountered in the field. Predefined scenarios were documented by field operation staff and provided to developers in advance as an aid to programmers in their initial unit and module tests. These documented plans were used in the testing of multiple scenarios.

Finally, all systems went through rounds of formal integrated systems testing. This included following the entire process through the cycle including loading and assigning cases using the management system, loading cases on the server, completing automated questionnaires, and updating completed work to the home office database. Any processes that failed the testing process or were modified went through the full testing life cycle again to ensure that modifications did not impact other related areas of code.

Problems identified in the testing process were documented and retested once a developer implemented the required correction.

4.7.2 Pilot Testing

Prior to Round 1, a pilot test was conducted in Baltimore, Maryland. Approximately 300 households were screened to obtain about 20 household interviews using the NSPY instruments. The purpose of the pilot was to test the adequacy of instrument skip patterns, question wording, and flow; to test the adequacy of the advance materials and interviewing procedures; and to test the application of the ACASI portion of the instrument, including the reaction of respondents to the content of the tutorial program in the ACASI that taught respondents how to complete the self-administered part of the interview. A debriefing was held at the end of the pilot data collection. From that, some questions were dropped from each of the extended interview instruments to keep within the OMB respondent burden estimates. Procedures and advance materials were updated as appropriate.

Westat conducted a second pilot test prior to Round 2 to test the followup screening instruments. The participating households in the first Baltimore pilot test were recontacted and screened for followup status. The purpose of the pilot was to test the adequacy of the screening instruments, skip patterns, question wording and flow, as well as advance materials and interviewing procedures. An interviewer debriefing was held at the end of this pilot data collection as well. From that debriefing, some minor changes were made in followup screening questions and procedures.

4.7.3 Quality Control During Training

Interviewers participated in an 8- to 10-day in-person training session prior to their working on Round 1 (the recruitment phase) and Round 2 and/or Round 3 (the followup phase). Interviewers who continuously worked on NSPY throughout the data collection periodically participated in refresher trainings via telephone.

The in-person training program, which was staffed by qualified project staff and field supervisors, was designed to ensure consistency in data collection through the use of lectures, with a heavy focus on practice sessions. Each training session was scripted so that multiple room trainings and trainings for each round of data collection were consistent. During training, trainees were given many practice sessions using a training database. With each screening and interview instrument, a demonstration session was performed by trainers, followed by two group practices and, finally, trainees completed multiple role plays in pairs. Prior to the conclusion of training, trainees were paired and

completed a certification script where the trainee practiced contacting the respondent, conducting the screening and interview instruments and ending the contact. Trainers went around the room assessing performance on a standardized form and "certified" those trainees who conducted the screenings and interviews following procedures. A few interviewers did not pass certification and were released from the study at the end of training. Another small subset were asked to complete extra training over the telephone with the field supervisors before starting work.

Trainees new to Westat attended an additional half-day training on general interview techniques. Bilingual trainees attended an additional half-day training that concentrated on reviewing bilingual scripts and materials.

4.7.4 Validations

Validations procedures called for the validation of at least 10 percent of the parents interviewed and at least 2 percent of the screening respondents living in ineligible households. During Round 1, approximately 12 percent of parents interviewed and 3 percent of screening respondents in ineligible households were selected for validation. Approximately 70 percent were contacted by telephone and attempts to contact the remainder were made by mail. When interviewers were suspected of falsifying data, all of their completed work was redone by different interviewers. In three instances, interviewers were terminated for falsifying data.

During Round 2, approximately 11 percent of the parents interviewed and 19 percent of screening respondents in ineligible households were selected for validation. Approximately 86 percent were contacted by telephone and attempts to contact the remainder were made by mail. No invalid cases were found for interviewers completing Round 2 work; however, two interviews completed during Round 1 were identified as questionable during Round 2 when an interviewer revisited the households. Upon further review, the cases were determined to be valid. In addition, some cases were found where an interviewer did not follow proper screening procedures in the previous round. Additional information was obtained from these households and weighting factors were adjusted for the affected parent data.

During Round 3, approximately 10 percent of the parents interviewed and 76 percent of screening respondents in ineligible households were selected for validation. About 87 percent were contacted by telephone and attempts to contact the remainder were made by mail. No invalid cases were

found for interviewers completing Round 3 work; however, some of one interviewer's work appeared questionable from an earlier wave. Upon further review and verification of the interviewer's work, it was determined to be valid.

Purposive and Random Criteria

Selection of individual respondents to be validated was based on three goals: (1) to target a minimum number of cases for each interviewer; (2) to validate each interviewer's first few completed interviews; and (3) to ensure that at least one interview was validated in each PSU.

Criteria for respondents selected for validation were based on random and purposive selection. Random selection employed an algorithm developed by the statistical staff. These respondents were identified at receipt control when the final status code was recorded in the home office management system. Approximately 50 percent of respondents were selected at random. For purposive selection, field supervisors and field management staff selected the interviewer's first two or three completed interviews, or selected respondents when they wanted to check on an interviewer's work for some reason.

TRC and Mail Procedures

Attempts to contact respondents were first made by telephone. Trained telephone interviewers at the TRC administered the validation questionnaire. Interviewers asked questions such as interview contact, amount of incentive, and if the interview was completed, as well as verified information on the members in the household.

Respondents were mailed a validation questionnaire when telephone validation attempts were unsuccessful or when respondents did not have a telephone number. The data receipt supervisor filled out a questionnaire template and mailed the questionnaire to the knowledgeable adult at the address where the interview took place. Up to two mailings were made to respondents to complete the validation.

4.8 Data Management

4.8.1 Coding and Editing Specifications for Survey Instruments

Data collected for the NSPY were conducted using computer-assisted interviewing techniques. These types of techniques included audio computer-assisted self-administered interviewing (ACASI) and computer-assisted personal interviewing (CAPI). Editing specifications were built into the computer programs used by the respondents and interviewers. These specifications forced decisions to be made up front as the data were being entered. Acceptable ranges and consistency checks were programmed into the screening and parent, teen, and child interview modules.

Range Specifications

Across all types of instruments, respondent answers were subjected to both "hard" and "soft" range edits during the interviewing process. A "soft range" is one that represents the reasonable expected range of values but does not include all possible values. Responses outside the soft range were confirmed with the respondent and entered a second time. For example, the number of days a child was absent from school in the last month has a soft range of 1 to 9. A value outside this range could be entered and confirmed as correct by the interviewer as long as it was within the hard range of values (1 to 22).

"Hard ranges" are those that have a finite set of values for closed-ended questions that can be entered into the computer. Common examples of hard ranges are values denoting months of the year. Months can be defined and entered as 01 = January; 02 = February; 03 = March; and so on up to 12 = December. If a respondent were to enter 13, this out-of-range value would not be accepted when entering into the computer. A message would pop up alerting the interviewer or the respondent that "13" was not a legitimate value. If the respondent or interviewer insisted that a response outside the "hard range" was valid, the interviewer would enter the information in a comments field. Data management staff would review these comments.

Consistency Checks

Consistency checks or logical edits examined the relationship between responses to ensure that they did not conflict with one another or that the response to one item did not make the response to another item unlikely. For example, in the household enumeration, one could not be recorded as a mother and a male. A respondent could not answer "None" to how much television they watch during the week and then be asked "In what language are the television programs you usually watch?" Since this question was not applicable based on the previous response, the computer was programmed to skip this question for consistency. In some cases, just as with "hard ranges," if the verified response still resulted in a logical error, or the answer to the primary question was incorrectly entered, the interviewer recorded the problem either in a comment or on a problem report. Data management staff would review these comments. If the comments supported the response that could not be entered, or a change to primary value, the data management staff updated the response. All updates to the data were maintained in a transaction file that recorded the original value, the updated value, the user, the date, and the time.

Editing from Instrument Comments and Problem Sheets

Not all comments and problem sheets submitted by interviewers required the data management staff to update or review the value in the related variable to determine whether the correct information had been entered, but because some of them did, all comments and problem sheets were reviewed. Comments were displayed via the editing system in an Access spreadsheet that could be sorted on any of the column headings. Headings on the spreadsheet included the ID, type of instrument, the question number at which the comment was entered, the name of the variable collected at that question, the value that was entered, the date that the interview was conducted, and the initials of the interviewer. Problem sheets were sent to the home office in hard-copy form and were enclosed with the case household folder.

For comment review, data management staff considered not only the information entered into the comment but also the location in the interview at which the comment was included and the response that the interviewer entered into the data record at that place. Problem sheets also were reviewed for data updating. These situations occurred where the interviewer forgot to enter a comment and the case was closed. Some comments and problem sheets required no editor action other than just to read them.

In cases where the data management staff reviewing the comments or problem sheets determined a data problem existed, an update was made. As stated previously, all updates to the data were maintained in a transaction file that recorded the original value, the updated value, the user, the date and the time.

4.8.2 Reconciliation and Consistency Edits for the Home Management System

Using the Home Management System (HMS), the disposition for every sampled unit was entered at the household level. If a household were eligible, disposition codes were also entered at the person level. Consistency checks were programmed into the HMS between household-level codes and person-level codes. For instance, a household that had been assigned a code of "FSC" required that all persons in the household had a disposition code of "SFCO." If a household was assigned a code of "FXCE," then at least one person in the household must have had a nonresponse disposition code.

The HMS was updated weekly with information from the Blaise interviewing files and the Field Management System. Edits were built into the system that would not allow certain disposition codes to be entered if there were not an interview record in the Blaise file. For instance, if a sampled person had a record in the interview file, the only disposition code that would be valid is a "complete" code. If a staff member attempted to enter any other code, a message on the screen explained that there were data in the interview file. Likewise, if a staff member attempted to enter a code of complete and there was no record in the interview file, a message on the screen explained there was no record in the interview file. If problems arose that could not be resolved, the staff member would discuss the problem with the data manager.

4.8.3 Frequency and Cross-Tabulation Review

Frequencies and cross-tabulations of the data were run at a minimum of 3 times for each data collection period: preliminary, interim, and final. Each variable was reviewed and evaluated to ensure the accuracy of the Blaise programs and the integrity of the edited data.

4.9 Cross-Sectional Response Rates for Youth

Round 1

Across the three data collection waves constituting Round 1 (i.e., Waves 1 through 3), a total of 80,991 dwelling units was selected for NSPY. Of these, 9,326 (11.5%) were discovered to be either vacant or nonresidences (such as businesses or other institutions). That left 71,665 occupied dwelling units to be contacted and screened for study eligibility.

Of the 71,665 occupied dwelling units, answers to the screening questions were obtained for 68,353 (95.4%). Of these, 8,565 (12.5%) had children who were 9 to 18 years of age and were eligible to participate in NSPY. Among the 8,565 eligible households, data collection staff were able to complete screening enumeration for 6,398 (74.7%) households.

Within the 6,398 households that completed the screening, 8,920 youth aged 9 to 18 years were sampled. Of these, 8,117 (91.0%) completed the Round 1 interview. This corresponds to an overall (cross-sectional) Round 1 youth response rate of 64.8 percent (i.e., 95.4% x 74.7% x 91.0%).

Round 2

Round 2 consisted of two data collection waves, Wave 4 and Wave 5. Wave 4 was the first followup of the Wave 1 sample, and Wave 5 was the first followup of the combined samples from Waves 2 and 3. Under the NSPY sample design, youth and parents in households that completed a screening enumeration in Round 1 were included in the followup sample if the household also contained at least one sampled person (either youth or parent) who completed an interview in Round 1. As a result of this selection criterion, a small number of youth who were sampled at Round 1 but who did not complete a Round 1 interview were refielded in Round 2.

Both followup and cumulative cross-sectional response rates were computed for Round 2. The followup cross-sectional response rate at Round 2 represents the percentage of youth refielded in Round 2 who were successfully located and interviewed in Round 2. It is defined to be the product of (a) the percentage of households fielded in Round 2 for which eligibility screening was completed and (b) the percentage of eligible youth in the screened households who completed the Round 2 interview.

There were 6,398 households that completed the screening enumeration at Round 1. Of these, 5,975 (93.4%) contained at least one respondent from Round 1 (either a youth or a parent) and were eligible for refielding at Round 2. Of these, 277 households were ineligible (contained only youth who aged out of the study). Thus, of the 5,975 households with at least one respondent in Round 1, a total of 5,698 were refielded in Round 2. An additional 57 split households (i.e., households that contained two sampled youth in the prior round who were living at different addresses in Round 2), were later added to the sample in Round 2. Thus, the total number of households for which followup telephone or inperson eligibility screening was attempted in Round 2 was 5,755 households. Of these, eligibility was determined for 5,237 (91.0%) of the households. For the remaining 518 households, eligibility could not be determined for various reasons (e.g., the household moved out of the interviewing area or was not locatable, the household could not be contacted for some other reason, or the household refused to complete the eligibility screening). The 5,237 successfully screened households contained 7,166 youth who were sampled in Round 1. Of the 7,166 sampled youth, 201 (2.8%) were determined to be ineligible for the Round 2 survey (e.g., were 19 years or older, were institutionalized or living in group quarters, or were deceased). Of the 6,965 eligible youth in the screened households, 6,516 (93.6%) completed the Round 2 interview. Thus, the followup cross-sectional response rate for youth at Round 2 was 85.2 percent (91.0% x 93.6%).

The cumulative cross-sectional response rate at Round 2 reflects the total sample losses occurring at both Rounds 1 and 2. It is defined to be the product of the following four rates:

- The percentage of households at Round 1 for which eligibility was determined;
- The percentage of eligible households at Round 1 for which the household enumeration was completed;
- The percentage of Round 1 households that were refielded at Round 2; and
- The followup cross-sectional response rate at Round 2.

Thus, the cumulative cross-sectional response rate for youth at Round 2 was 56.7 percent (95.4% x 74.7% x 93.4% x 85.2%).

Round 3

Round 3 consisted of two data collection waves, Wave 6 and Wave 7. Wave 6 was the second followup of the Wave 1 sample. Wave 7 was the second followup of the combined samples from Waves 2 and 3. Both followup and cumulative cross-sectional response rates were computed for Round 3. Under the NSPY sample design, only those youth (and parents) who were expected to be eligible for NSPY at Round 3 *and* who met the following criteria were refielded: (a) the youth/parent resided in a household in which at least one sampled youth completed either the Round 1 or Round 2 interview, and (b) the household was neither "not locatable" nor "out of area" in Round 2. Under these selection criteria, a small number of youth and parents who did not complete either the Round 1 or Round 2 interviews were refielded in Round 3.

The (second) followup cross-sectional response rate represents the percentage of youth refielded in Round 3 who were successfully located and interviewed during Round 3. It is defined to be the product of (a) the percentage of households fielded in Round 3 for which eligibility screening was completed and (b) the percentage of eligible youth in the screened households who completed the Round 3 interview.

There were 6,398 households that completed the screening enumeration at Round 1. Of these, 5,945 (92.9%) contained at least one responding youth from either Round 1 or Round 2. Of these, 513 were ineligible (contained only youth who aged out of the study) and 292 were not located or had moved out of the PSU in Round 2. Thus, of the 5,945 households with responding youth in Round 1 or 2, 5,140 (94.6%) were refielded in Round 3. An additional 46 split households (i.e., households that contained two sampled youth in the prior round who were living at different addresses in Round 3) were later added to the sample in Round 3. Thus, the total number of households for which followup telephone or inperson eligibility screening was attempted in Round 3 was 5,186. Of these, eligibility was determined for 4,818 (92.9%) households. For the remaining 368 households, eligibility could not be determined for various reasons (e.g., the household moved out of the interviewing area or was not locatable, the household could not be contacted for some other reason, or the household refused to complete the eligibility screening).

The 4,818 successfully screened households contained 6,428 youth, of which 208 were determined to be ineligible (e.g., were 19 years or older, were institutionalized or living in group quarters, or were deceased). Of the 6,220 eligible youth in the screened households, 5,854 (94.1%) completed the

Round 3 interview. Thus, the (second) followup cross-sectional response rate for youth at Round 3 was 87.4 percent (92.9% x 94.1%).

The cumulative cross-sectional response rate at Round 3 reflects the total sample losses occurring at Rounds 1, 2, and 3. It is defined to be the product of the following five rates:

- The percentage of households at Round 1 for which eligibility was determined;
- The percentage of eligible households at Round 1 for which the household enumeration was completed;
- The percentage of Round 1 households that contained a responding youth in either Round 1 or Round 2);
- The percentage of households with responding youth in Round 1 or Round 2 that were refielded in Round 3; and
- The followup cross-sectional response rate at Round 3.

Thus, the cumulative cross-sectional response rate for youth at Round 3 was 54.8 percent (95.4% x 74.7% x 92.9% x 94.6% x 87.4%).

5. CONSTRUCTION OF EXPOSURE, OUTCOME, AND RISK SCORE INDICES

The PUF datasets contain two exposure indices, three outcome indices, and a risk score index, all for youth. The general exposure and specific exposure indices are described in Section 5.1. The three outcome indices are explained in Section 5.2. Finally, the construction of the risk score index is described in Section 5.3.

5.1 Exposure Indices

The general exposure index, described in Section 5.1.1, is based on questions that ask about exposure to anti-drug media messages through a variety of channels. The specific exposure index, described in Section 5.1.2, is based on the specific Media Campaign television ads being broadcast during the 60 days prior to the respondent's interview. Since each respondent was asked about recall of only a sample of the ads, responses to the nonsampled ads were imputed for the construction of the specific exposure index if more than four ads were eligible for inclusion. Section 5.1.3 presents evidence for the validity of the television ad recall measures for youth.

5.1.1 General Exposure Index

The general exposure index (variable name is *YGEIORD3* in the dataset) captures exposure to anti-drug ads, not limited to Media Campaign ads, through a wide variety of channels, including movies, TV, radio, and billboards (see Exhibit 5-1). Note that in each question, the reference period is "in recent months." This wording was used to maintain equivalence to the wording used by the Monitoring the Future surveys in questions about recall of anti-drug advertising.

The responses to these questions were combined in a way that is designed to reflect the total number of ad viewings experienced by the respondent. Each possible response was translated into a certain number of exposures over a 1-month period, as shown in Table 5-1, assuming that the average person would mostly refer to the last month in trying to interpret "recent months." The scores for the responses for the four questions in Exhibit 5-1 were then added together to create a variable running from 0 to a maximum of 180. This continuous scale was categorized into three classes as shown in Table 5-2.

Exhibit 5-1. Youth questions on general exposure

The next questions ask about anti-drug commercials or "ads" that are intended to discourage ***illicit drug*** use.

D10. In recent months, about how often have you seen such anti-drug ads on TV, or heard them on the radio?

Not at all	1
Less than one time a month	2
1 to 3 times a month	3
1 to 3 times a week	4
Daily or almost daily	5
More than 1 time a day	6

D11. In recent months, about how often have you seen such anti-drug ads in newspapers or magazines?

Not at all	1
Less than one time a month	2
1 to 3 times a month	3
1 to 3 times a week	4
Daily or almost daily	5
More than 1 time a day	6

D12. In recent months, about how often have you seen any anti-drug billboards or other public anti-drug ads such as on buses, in malls, or at sports events?

Not at all	1
Less than one time a month	2
1 to 3 times a month	3
1 to 3 times a week	4
Daily or almost daily	5
More than 1 time a day	6

D13. In recent months, about how often have you seen such anti-drug ads in the movie theaters or on rental videos?

Haven't gone to movies or rented videos in recent months	0
Not at all	1
Less than 1 time a month	2
1 to 3 times a month	3
1 to 3 times a week	4
Daily or almost daily	5
More than 1 time a day	6

Table 5-1. Coding of general exposure questions

Response Category	New Value
Not at all	0.0
Less than 1 time a month	0.5
1 to 3 times a month	2.0
1 to 3 times a week	8.0
Daily or almost daily	30.0
More than 1 time a day	45.0

Table 5-2. Categories for the General Exposure Index

Category	Label	New value for categorical version
Less than 4 times per month	Low	1
4 to less than 12 times per month	Medium	2
12 or more times per month	High	3

The categories in Table 5-2 were chosen for ease of communication and to produce a reasonable distribution of the sample.

In addition to the three-category general exposure measure (YGEIORD3), other recodes of the general exposure measure were also included in the data files. These included a four-category measure of general exposure (YGEIORD) and a continuous measure of general exposure (YGEI). The first category of the YGEIORD3 measure, "Less than 4 times per month" was divided into two categories: "Less than once a month" and "1 to less than 4 times per month" to create the YGEIORD measure. YGEI is a summed measure of the number of times per month, for the past few months, the respondent has seen or heard anti-drug ads on television, radio, magazines, newspapers, billboards, in movie theaters, or on rented movie videos.

5.1.2 Specific Exposure Index and Ad Imputation Procedures

The second exposure index is a recall-aided exposure index (variable name is *YRAEI* in the dataset) based on recall of the Media Campaign television ads. A sample of the Media Campaign ads was shown to each respondent. The sample was limited to ads that targeted the respondent. This means that

for youth, only youth ads[1] were sampled; for English speakers, only English ads were sampled; and for Spanish speakers, only Spanish ads were sampled unless they were bilingual, in which case, ads in both languages were sampled.

The sampled ads were played on the laptop computer for the respondents. The questions in Exhibit 5-2 were repeated for every television and radio ad the respondent was shown. Responses for ads that were eligible for selection but not actually selected for a particular respondent were imputed.

Exhibit 5-2. Specific ad questions

Television Ads:

D17a. Now we will show some ads that might or might not have been playing on television around here. Have you ever seen or heard this ad? (PLAY TV AD.)

Yes	1	
No	2	(D18a)
REFUSED		(D18a)
DON'T KNOW		(D18a)

D17b. In recent months, how many times have you seen or heard this ad?

Not at all	1	(D18a)
Once	2	
2 to 4 times	3	
5 to 10 times	4	
More than 10 times	5	

Radio Ads:

D23a. Now we will play you some ads that might or might not have been playing on the radio around here. Have you ever heard this ad? (PLAY RADIO AD.)

Yes	1	
No	2	(D24a)
REFUSED		(D24a)
DON'T KNOW		(D24a)

D23b. In recent months, how many times have you heard this ad?

Not at all	1	(D24a)
Once	2	
2 to 4 times	3	
5 to 10 times	4	
More than 10 times	5	

[1] Youth were shown youth television, youth radio and youth ringer ads (i.e., youth ad never aired on television) in Round 1. However, starting in Round 2, every month, youth had a random chance of being selected to see a spill ad, i.e., an ad geared to the parent audience, or a ringer ad. These spill and ringer ads were not used in the calculation of the Specific Exposure Index. Radio ads were not used in the calculation of exposure.

Generally, each respondent was shown up to four television ads and two radio ads, which had broadcast dates during the prior 2 months. In addition, youth were shown an ad that had not ever been broadcast to assess tendency to claim recall falsely. African Americans and bilingual Hispanics could be shown up to two additional television ads and two additional radio ads that were targeted to their racial/ethnic group. Radio ads were played only to teenagers and parents and not to children aged 9 to 11 years.

To accurately characterize each respondent's total exposure to all television ads on the air for the recall-aided exposure index, the respondent's viewing levels of the nonsampled ads was needed. To assess these viewing levels, if more than four ads were eligible, responses for the nonsampled ads were imputed. The exposure data were complicated by:

- NSPY procedures, which specified that different ad selection rules were used for minorities. Since the Media Campaign had ads targeted at different audience groups (parents or youth) and racial or ethnic groups (general market, African American, or Hispanic), each month a new ad lookup table was transmitted to interviewer laptops with algorithms of which ads to show each specific type of respondent.

- Variation in the number of gross rating points (GRPs) purchased for each ad. GRPs are the customary unit for measuring exposure to ads within the advertising industry. If 1 percent of the target population sees an ad one time, the ad earns one GRP (see Hornik et al., 2003, Chapter 3 for further discussion of GRPs).

As a result, developing a satisfactory index was complex. Simply summing the recall of the sampled ads would have made minorities appear to have been more heavily exposed because they were shown more ads on average during their interviews. Simply averaging the recall of the sampled ads would have made people who were shown ads with low GRPs appear to be less heavily exposed than those who were shown ads with high GRPs. Imputation for ads not shown to each respondent appeared to be the best approach. Two different imputation procedures were used depending on the total number of times that an ad was sampled during a wave, i.e., a 6-month data collection period. The two procedures used were single-cell hotdeck imputation and n-cell hotdeck imputation. A single-cell hotdeck imputation was generally used when the total number of respondents for which an ad was in-scope during a wave was 150 or less; when there was more information about the distribution of viewing of an ad (sample size more than 150), the more complex n-cell hotdeck imputation was used. These imputation techniques are described in Hornik et al., (2003).

After imputation, all respondents had answers (either actual or imputed) to the questions shown in Exhibit 5-2 for every television ad (1) that had been on the air in the 60 days preceding the day of interview and (2) that targeted them.

The responses were recoded as shown in Table 5-3. These recoded values were then summed across ads to get a total number of viewings. After summation, the resulting scales were broken into the categories shown in Table 5-4. In labeling the categories in this scale, the assumption was made that respondents would consider what they had seen in the 2-month period before the interview in response to the request to consider "recent months." Thus the total viewing frequencies were divided by 2 to generate per month frequencies. The choice of the "recent months" term reflected a decision to maintain the language of the MTF surveys in the absence of clear evidence that a more specific time period would generate better information. In fact, as shown in the validity analysis below, the estimated weekly frequencies based on the 2-month assumed period were quite similar to the purchased GRP estimates for weekly expected exposure, suggesting that this equation of recent months with 2 months has some empirical support. Four levels were chosen for this index instead of the three chosen for the general index because a sizable proportion of the respondents fell in the "None" category.

Table 5-3. Recoding of responses to exposure to specific ads

Question: Here is another TV ad. Have you ever seen or heard this ad?	[If yes,] In recent months, how many times have you seen or heard this ad?	Recoded response
No		0.0
Don't know		0.5
Yes	Not at all	0.0
Yes	Once	1.0
Yes	2 to 4 times	3.0
Yes	5 to 10 times	7.5
Yes	More than 10 times	12.5

Table 5-4. Categories for the recall-aided exposure index

Category	Label	New value for categorical version
None	None	0
One to less than 4 times per month	Low	1
4 to less than 12 times per month	Medium	2
12 or more times per month	High	3

In addition, a three-level, recall-aided exposure index (YRAEI3) was also included. The top two categories of YRAEI ("4 to less than 12 times per month" and "12 or more times per month") were combined to create the top category of YRAEI3 ("4 or more times per month").

5.1.3 Measurement Quality of the Specific Ad Recall Measure

The assessment of the validity of youth television ad recall was addressed in two ways. First, the recall of the ads actually shown on television was compared to the claimed recall of "ringer" ads, i.e., ads played for respondents although they had never appeared on television. Second, the average recall of each shown ad was compared with the total advertising time (GRPs) purchased for each ad. Positive correlation between the average recall of specific ads with the GRPs purchased for them would provide supportive evidence for the validity of the recall responses.[2]

The ringer ads that had never been aired were used to estimate the tendency to claim that an ad had been seen when it had not. This validation exercise was undertaken after the first wave of data collection. On average, the ringer ads were "recalled" 11 percent of the time. In contrast, of the 13 real general audience ads, seven had recall rates over 50 percent, four others were between 20 to 40 percent and only two were between 10 to 19 percent, not dissimilar to the ringer ads. These results give confidence that the ad recall responses do reflect true exposure to a large extent, as do the observed association between recall of ads and GRPs purchased for the ads.

Table 5-5 presents the data that were used to estimate the fit between youth recall of a specific ad and the GRP that was assigned to that ad. The first column of the table indicates, for instance, that the ad "Hockey" was shown to 1,145 youth respondents (out of 3,314). It had been on the air 16.35 days of the 60 days preceding the interview for the average respondent. Of the 1,145 respondents who were shown the ad, 51 percent recognized it. These respondents claimed to have seen it an average 2.59 times "in recent months." Estimated weekly exposure is computed from these data as the total number of times the ad was seen divided by the average number of days the ad was on the air, which gives an estimate of exposures per day, and then multiplied by 7 to estimate exposures per week. Thus, for the

[2] Note that the analysis presented in this section is discussed in the Evaluation of the National Youth Anti-Drug Media Campaign: Second Semi-Annual Report of Findings (2001). This analysis was completed on data collected from November 1999 to May 2000. An updated version of this analysis has been described in Southwell, et al. (2002). While the Southwell, et al. paper used more updated data, their substantive conclusions on the assessment of the validity of youth television ad recall are consistent with the analysis described in this section.

Table 5-5. Association between recall and GRP estimates

Ad name*	Number of eligible respondents	Number of days aired in 60 days before interview	Proportion who had seen the ad in recent months	Mean number of times ad seen in recent months	Estimated weekly recalled exposure	Gross Rating Points per week ad was on the air**
Mary J. Blige	1,636	35.42	0.60	3.27	0.65	94.80
Brothers	492	33.39	0.69	3.42	0.72	80.80
Andy McDonald	284	15.53	0.51	2.07	0.93	75.25
Howtosay	904	28.92	0.53	2.88	0.70	66.75
Hockey	1,145	16.35	0.51	2.59	1.11	66.33
No thanks	1,137	37.99	0.66	3.50	0.64	62.75
Drugs Kill (Williams sisters)	603	33.57	0.63	2.50	0.52	51.44
Dixie Chicks	560	29.02	0.40	1.59	0.38	43.71
Mother Daughter	492	26.39	0.21	0.60	0.16	21.00
Scatman	426	25.23	0.24	0.93	0.26	20.50
No skill	576	13.39	0.11	0.37	0.19	7.00
Michael Johnson	448	13.44	0.18	0.52	0.27	5.57

* Only English language ads are included in this analysis.

**GRPs are customarily reported on a scale equivalent to 100 times the scale on which the weekly exposure estimates are reported. Thus 94.80 GRPs is equivalent to an expected weekly exposure of 0.948.

Hockey ad, the estimated weekly exposure is $7 \times 2.59/16.35 = 1.11$. The final column is obtained from the advertising data reported by the media buy contractor by dividing the total GRPs purchased for the specific ad by the number of weeks the ad was on the air. In the case of the Hockey ad, there were a total of 398 GRPs purchased for a total of 6 weeks; therefore, the gross ratings points per week the ad was on the air was 66.33. This calculation was repeated for each of the general market ads. The data in the last two columns, one developed on the basis of the NSPY recall data and the other on the basis of reported GRPs, are remarkably consistent. With the ad as the unit of analysis, the correlation between the values in these two columns is $r=0.81$. Even more, they provide very similar estimates for weekly exposure for each ad, recalling that the GRPs, divided by 100, represent the exposures per person per week.

This is strong evidence for two inferences. First, GRPs matter; they largely define how well youth will recall the television ads. Second, the recall measures, at least when aggregated in this way, are strong measures. They would not be so highly related to weekly GRPs otherwise.

5.2 Outcome Indices

Three outcome indices that reflect the assumed theoretical model of campaign effects were developed. These three intermediate outcomes indices (the respective variable names are in parentheses) were "Self-efficacy to refuse marijuana" *(MJEFFIC),* "Attitudes and beliefs about marijuana" *(MJATTBEL),* and "Perceived social norms" *(MJNORM)*. The basic theoretical model underpinning the Evaluation argued that if the Media Campaign were to be successful, it would affect behavior through one or more of the paths depicted in Figure 5-1.

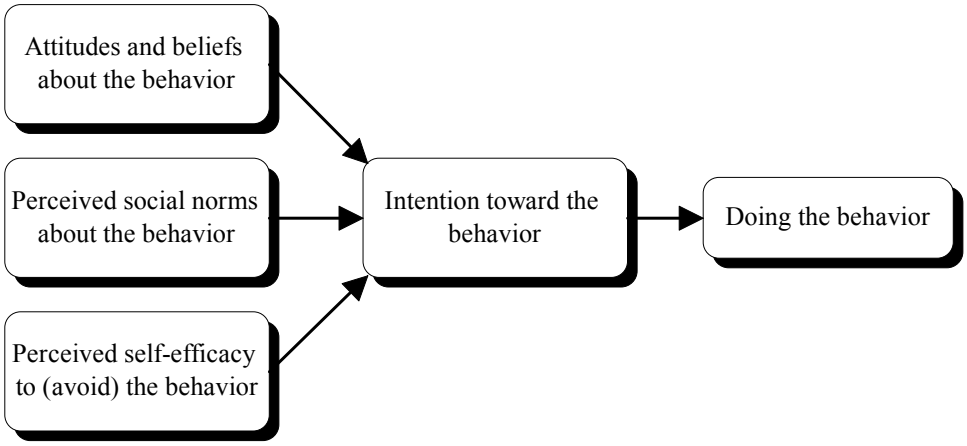

Figure 5-1. The expected relationships among cognitive outcomes

In the model in Figure 5-1, intentions are assumed to be influenced by (1) self-efficacy to avoid drug usage, (2) knowledge, beliefs, and attitudes, and (3) perceived social norms. The items used to measure these indices are as follows:

- **Self-efficacy to refuse marijuana:** All youth were presented with the same five items about their confidence that they could turn down the use of marijuana under various circumstances. ("How sure are you that you can say no to marijuana, if you really wanted to, if: You are at a party where most people are using it; A very close friend suggests you use it; You are home alone and feeling sad or bored; You are on school property and someone offers it; You are hanging out at a friend's house whose parents aren't home.")

- **Attitudes and beliefs about marijuana:** All youth were asked a question that assessed overall attitude toward "once or twice" (i.e., trial) use and a question that assessed overall attitude toward regular use. In addition to these two questions, youth were asked to respond to eight items about how likely it was that a series of specific consequences would result if "you" use marijuana. However, some youth were

presented with one set of items relating to the regular use of marijuana (every month or almost every month) and other youth were presented with a set of items about occasional or trial use (once or twice) over the next year. All of the youth who had used marijuana in the past year and a random half of the nonpast-year users were presented with a set of items about the consequences of regular use. The rest of the youth were presented with a set of eight items about the consequences of trial use. The consequences in the eight items relating to regular use were "Damage my brain," "Mess up my life," "Do worse in school," "Be acting against my moral beliefs," "Lose my ambition," "Lose my friends' respect," "Have a good time with friends," and "Be more creative and imaginative." The consequences in the eight items relating to trial use were "Upset my parents," "Get in trouble with the law," "Lose control of myself," "Start using stronger drugs," "Be more relaxed," "Have a good time with friends," "Feel better," and "Be like the coolest kids.

- **Perceived social norms:** All youth were asked to respond to five items that assessed social normative pressure. However, in the same way as for the attitude and belief questions, some youth were asked one set of items and the other youth were asked a different set. All of the youth who had used marijuana in the past year and a random half of the nonpast-year users were asked a set of items relating to regular use, while the remaining youth were asked a set of items relating to trial use. Parallel items were used in the two sets. They asked about the perception of friends' use of marijuana, other peers' use of marijuana, parents' disapproval of "your" marijuana use, friends disapproval of "your" marijuana use, and disapproval of "your" marijuana use by most people important to you.

For each of the three indices, the items were combined to form indices using the same basic approach. First, logistic regression equations were computed to predict the intention to definitely not use marijuana, even once or twice, in the next year. The predictor variables in the regression equations were the items in the index in question. The predicted values from the regressions were then rescaled to have means and standard deviations of 100 to form the index values. This process was straightforward for the self-efficacy index but more complex for the other two indices because of the use of different sets of items for different sets of youth.

Table 5-6 displays the results from the logistic regression for the self-efficacy items. The summed index was substantially associated with the intention to use marijuana in the next year (see Figure 5-2).The self-efficacy index predicted intentions, but less powerfully than the other two indices.

Table 5-6. Logistic regression model for intentions to avoid any marijuana use among 12- to 18-year-olds in terms of self-efficacy to refuse offers of marijuana

Item	Item wording	Values	Value Label	Coefficient	Standard Error
C9(a)	Certainty of refusing marijuana when at a party where most people are using it	1-3	Somewhat sure, slightly sure, or not at all sure	-0.1805	0.1421
		4	Mostly sure	0.2339	0.1130
		5	Completely sure	-0.0535	0.1166
C9(b)	Certainty of refusing marijuana when a very close friend suggests using it	1-3	Somewhat sure, slightly sure, or not at all sure	-0.0627	0.1530
		4	Mostly sure	-0.1604	0.1110
		5	Completely sure	0.2231	0.1197
C9(c)	Certainty of refusing marijuana when home alone and feeling sad or bored	1-3	Somewhat sure, slightly sure, or not at all sure	-0.6240	0.1402
		4	Mostly sure	-0.0458	0.1221
		5	Completely sure	0.6699	0.1051
C9(d)	Certainty of refusing marijuana when on school property	1-3	Somewhat sure, slightly sure, or not at all sure	0.6551	0.1892
		4	Mostly sure	-0.3183	0.1556
		5	Completely sure	-0.3367	0.1356
C9(e)	Certainty of refusing marijuana when hanging out at a friend's house whose parents aren't home	1-3	Somewhat sure, slightly sure, or not at all sure	-0.8485	0.1527
		4	Mostly sure	-0.1478	0.1118
		5	Completely sure	0.9963	0.1221

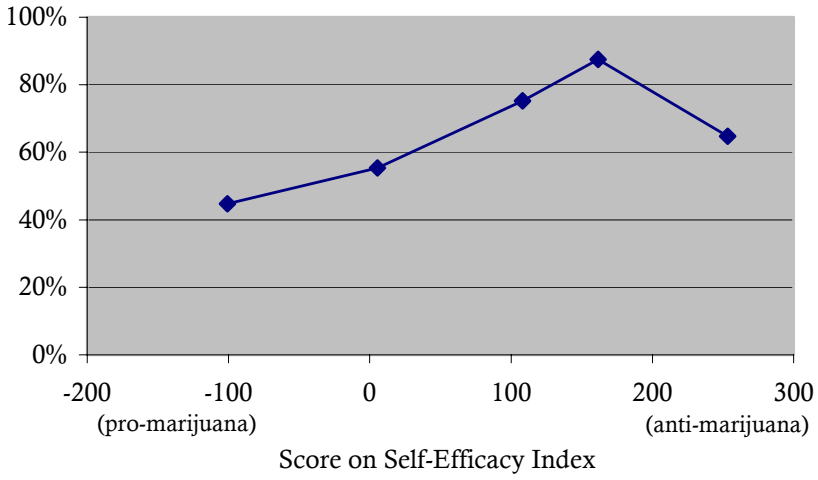

Figure 5-2. Marijuana nonuse intention by Self-Efficacy Index

The indices for beliefs/attitudes and for social norms were more difficult to construct. For the beliefs/attitudes index, the first step in the process was to model intentions to avoid future use for nonusers in terms of beliefs and attitudes about trial use. This logistic regression model is shown in Table 5-7. The second step was to model intentions to avoid future use for nonusers in terms of beliefs and attitudes about regular use. This model is shown in Table 5-8. Based on these two models, separate indices were formed for the two random subsets of nonusers. The third step was to equate the two indices by rescaling them so that they had a common mean and standard deviation on the population of nonusers. The rescaled regular use index was then applied to the items about regular use asked of users. The end result of this operation was to create an index for all youth in the dataset that reflects the influence on intentions for avoidance of future use of an amalgam of beliefs and attitudes about both marijuana trial and regular marijuana use. The summed Attitudes/Beliefs Index, as expected, was substantially associated with the intention to use marijuana in the next year. Figure 5-3 presents that relationship graphically. Only about 25 percent of those with the lowest scores on that index said "definitely not" to marijuana use in the next year, whereas 100 percent of those who were at the highest levels rejected such use.

A parallel process was used for social norms. Table 5-9 presents the parameter estimates for the logistic regression model for social norms about trial use and Table 5-10 presents the corresponding estimates for social norms about regular use. The perceived Social Norms Index was substantially correlated with intentions, although the relationship was not quite as strong as that between the Attitudes/Beliefs Index and intention (Figure 5-4).

In addition to the above indices, an additional index was created for youth to summarize personal beliefs about inhalants. (Variable name is INBELIEF in the data files; note that there were no questionnaire items on attitudes, social norms or self-efficacy with respect to inhalants.) As with marijuana, the importance of each component in the index was determined from the parametric model for intentions to avoid inhalant use in terms of the components. The fitted model is shown in Table 5-11.

Table 5-7. Logistic Regression Model for intentions to avoid any marijuana use among 12- to 18-year-old non-marijuana users in terms of personal beliefs and attitudes about trial marijuana use

Item	Item wording	Values	Value Label	Coefficient	Standard Error
C3a(a)	Trying marijuana would upset parents/caregivers	1-3	Very unlikely, unlikely, or neither likely nor unlikely	0.1524	0.2695
		4	Likely	-0.5901	0.3027
		5	Very likely	0.4377	0.2118
C3a(b)	Trying marijuana would cause legal trouble for youth	1-3	Very unlikely, unlikely, or neither likely nor unlikely	-0.3179	0.1949
		4	Likely	0.1289	0.2095
		5	Very likely	0.1891	0.2329
C3a(c)	Trying marijuana would cause youth to lose control	1-3	Very unlikely, unlikely, or neither likely nor unlikely	-0.1752	0.2224
		4	Likely	-0.2441	0.2164
		5	Very likely	0.4193	0.3087
C3a(d)	Trying marijuana would cause youth to use stronger drugs	1-3	Very unlikely, unlikely, or neither likely nor unlikely	-0.0221	0.2478
		4	Likely	0.3056	0.2823
		5	Very likely	-0.2835	0.3883
C3a(e)	Trying marijuana would cause youth to be more relaxed	1	Very unlikely	0.1361	0.2427
		2	Unlikely	0.0211	0.2468
		3-5	Neither likely nor unlikely, likely, or very likely	-0.1572	0.2036
C3a(f)	Trying marijuana would cause youth to have a good time with friends	1	Very unlikely	0.4546	0.2688
		2	Unlikely	-0.4197	0.2310
		3-5	Neither likely nor unlikely, likely or very likely	-0.0349	0.2180
C3a(g)	Trying marijuana would cause youth to feel better	1	Very unlikely	-0.1994	0.2331
		2	Unlikely	0.1629	0.2189
		3-5	Neither likely nor unlikely, likely, or very likely	0.0365	0.2327
C3a(h)	Trying marijuana would cause youth to be like the coolest kids	1	Very unlikely	0.3274	0.1942
		2	Unlikely	0.2613	0.2122
		3-5	Neither likely nor unlikely, likely, or very likely	-0.5886	0.2038
C4a	Youth perception of trying marijuana in the next year (7-point scale from "extremely bad" to "extremely good")	1		1.4258	0.2460
		2		-0.3259	0.2440
		3		-0.2839	0.3129
		4-7		-0.8160	0.2806
C5a	Youth perception of trying marijuana in the next year (7-point scale from "extremely unenjoyable" to "extremely enjoyable")	1		0.8747	0.2433
		2		0.2961	0.2593
		3		-0.6307	0.2843
		4-7		-0.5402	0.2846

Table 5-8. Logistic Regression Model for intentions to avoid any marijuana use among 12- to 18-year-old non-marijuana users in terms of personal beliefs and attitudes about regular marijuana use

Item	Item wording	Values	Value Label	Coefficient	Standard Error
C3b(a)	Regular marijuana use would damage youth's brain	1-3	Very unlikely, unlikely, or neither likely nor unlikely	-0.1549	0.2164
		4	Likely	-0.0435	0.1858
		5	Very likely	0.1984	0.2141
C3b(b)	Regular marijuana use would mess up youth's life	1-3	Very unlikely, unlikely, or neither likely nor unlikely	0.2318	0.2415
		4	Likely	-0.0884	0.1969
		5	Very likely	-0.1434	0.2395
C3b(c)	Regular marijuana use would make youth do worse in school	1-3	Very unlikely, unlikely, or neither likely nor unlikely	-0.3141	0.2464
		4	Likely	-0.0044	0.1933
		5	Very likely	0.3186	0.2318
C3b(d)	Regular marijuana use would be acting against youth's moral beliefs	1-3	Very unlikely, unlikely, or neither likely nor unlikely	-0.2912	0.1988
			Likely	0.1467	0.1973
		4	Very likely	0.1446	0.2104
		5			
C3b(e)	Regular marijuana use would cause youth to lose ambition	1-3	Very unlikely, unlikely, or neither likely nor unlikely	-0.0250	0.2259
		4	Likely	0.1443	0.1977
		5	Very likely	-0.1193	0.2447
C3b(f)	Regular marijuana use would cause youth to lose friends' respect	1-3	Very unlikely, unlikely, or neither likely nor unlikely	-0.5111	0.1967
		4	Likely	0.1517	0.1983
		5	Very likely	0.3594	0.2349
C3b(g)	Regular marijuana use would cause youth to have a good time with friends	1	Very unlikely, Unlikely	1.0099	0.2677
		2	Neither likely nor unlikely,	-0.6336	0.2172
		3-5	likely or very likely	-0.3762	0.1953
C3b(h)	Regular marijuana use would cause youth to be more creative and imaginative	1-3	Very unlikely, unlikely, or neither likely nor unlikely	-0.1549	0.2437
		4	Likely	0.1546	0.3294
		5	Very likely	0.0004	0.3749
C4b	Youth perception of regular marijuana use in the next year (7-point scale from "extremely bad" to "extremely good")	1		0.9698	0.2370
		2		-0.2337	0.2386
		3		-0.7086	0.2921
		4-7		-0.0275	0.3042
C5b	Youth perception of regular marijuana use in the next year (7-point scale from "extremely unenjoyable" to "extremely enjoyable")	1		0.7496	0.2271
		2		-0.1493	0.2414
		3		-0.2438	0.2936
		4-7		-0.3565	0.2451

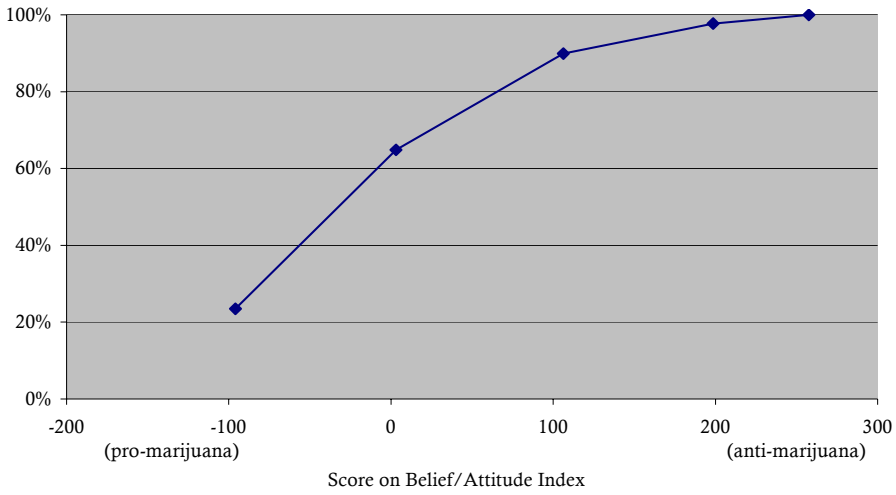

Figure 5-3. Marijuana nonuse intention by Attitudes/Beliefs Index

Table 5-9. Logistic Regression Model for intentions to avoid any marijuana use among 12- to 18-year-old non-marijuana users in terms of perceived social norms about trial marijuana use

Item	Item wording	Values	Value Label	Coefficient	Standard Error
C6a	Youth perception of most important people's reaction to youth trying marijuana	1	Strongly disapprove	0.3815	0.2229
		2	Disapprove	-0.4784	0.2455
		3-5	Neither approve nor disapprove, approve or strongly approve	0.0970	0.3381
C7a	Youth perception of close friends' reaction to youth trying marijuana	1	Strongly disapprove	1.0315	0.1786
		2	Disapprove	-0.0991	0.1618
		3-5	Neither approve nor disapprove, approve or strongly approve	-0.9324	0.1681
C8a	Youth perception of parents' reaction to youth trying marijuana	1	Strongly disapprove	0.5658	0.2729
		2	Disapprove	0.0545	0.3315
		3-5	Neither approve nor disapprove, approve or strongly approve	-0.6203	0.4227
C10a	Youth perception of how many friends have tried marijuana	1-2	None or a few	0.3854	0.1918
		3	Some	-0.1872	0.2012
		4-5	Most or all	-0.1982	0.2568
C11	Youth perception of how many kids in same grade or same age have tried marijuana	1-2	None or a few	0.3894	0.1764
		3	Some	-0.1868	0.1607
		4-5	Most or all	-0.2026	0.2039

Table 5-10. Logistic Regression Model for intentions to avoid any marijuana use among 12- to 18-year-old non-marijuana users in terms of perceived social norms about regular marijuana use

Item	Item wording	Values	Value Label	Coefficient	Standard Error
C6b	Youth perception of most important people's reaction to youth using marijuana regularly	1	Strongly disapprove	0.6495	0.2230
		2	Disapprove	-0.2729	0.2472
		3-5	Neither approve nor disapprove, approve or strongly approve	-0.3765	0.3476
C7b	Youth perception of close friends' reaction to youth using marijuana regularly	1	Strongly disapprove	0.9112	0.1844
		2	Disapprove	-0.0951	0.1722
		3-5	Neither approve nor disapprove, approve or strongly approve	-0.8160	0.1825
C8b	Youth perception of parents' reaction to youth using marijuana regularly	1	Strongly disapprove	-0.0445	0.2371
		2-5	Disapprove, neither approve or disapprove, approve or strongly approve	0.0445	0.2371
C10b	Youth perception of how many friends have used marijuana regularly	1-2	None or a few	0.2339	0.2050
		3	Some	0.0106	0.2192
		4-5	Most or all	-0.2445	0.2814
C12	Youth perception of how many kids in same grade or same age have used marijuana regularly	1-2	None or a few	0.3827	0.1874
		3	Some	-0.1066	0.1726
		4-5	Most or all	-0.2761	0.2353

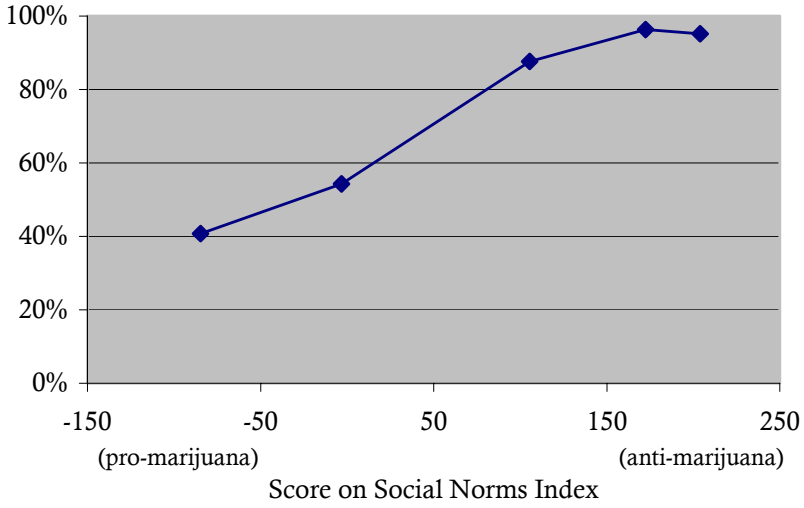

Figure 5-4. Marijuana nonuse intention by Social Norms Index

Table 5-11. Model for intentions to avoid any inhalant use among 12- to 18-year-olds in terms of personal anti-inhalant beliefs

Item	Item wording	Values	Value Label	Coefficient	Standard Error
C33a(c)	Youth perception of risk of harm when trying inhalants	1-2	No or slight risk	-0.3292	0.1177
		3	Moderate risk	0.0600	0.1066
		4	Great risk	0.2692	0.1249
C33a(d)	Youth perception of risk of harm when using inhalants regularly	1-2	No or slight risk	0.2185	0.1823
		3	Moderate risk	-0.3062	0.1339
		4	Great risk	0.0876	0.1328
C33(c)	Youth approval of others trying inhalants	1	Strongly disapprove	1.3941	0.1511
		2	Disapprove	-0.1367	0.1153
		3-5	Neither approve nor disapprove, approve, or strongly approve	-1.2574	0.1330
C33(d)	Youth approval of others using inhalants regularly	1	Strongly disapprove	0.2942	0.1249
		2	Disapprove	-0.1642	0.1162
		3-5	Neither approve nor disapprove, approve, or strongly approve	-0.1301	0.1412

5.3 Predictors of Marijuana Use and the Development of a Risk Model

In analyses published in the Evaluation reports, youth were stratified into lower and higher risk subgroups. This reflects the expectations of the Media Campaign implementers who have argued from the start of the Media Campaign that their target audience was those youth who were at risk of marijuana use.

Stratification into risk subgroups (variable name is *RISKSCR* in the dataset) was made on the basis of cross-predicted risk probabilities for marijuana use in the past year. This section briefly presents its underlying logic and the measures used. *The sample for the development of the "risk score" (the predicted probability of the undesired event) was the combination of the three waves in Round 1 of the NSPY data collection. (A respondent's risk score is calculated on the basis of the same equation across rounds).* The outcome variable was defined as marijuana use that began or continued in the past 12 months. Youth who had used in previous years but not in the past year were excluded from the analysis, as were youth who were under 12 years old at Round 1. The total number of cases in the analysis was 4,804. The list of youth and parent covariates for developing the risk model was gleaned from existing literature on risk factors for adolescent problem behavior in general and for substance use in particular.

However, the consideration of what variables were to be included was subject to an additional limitation. No variable that might have been affected by the Media Campaign directly or indirectly or that could be a consequence as well as a cause of marijuana use was eligible for inclusion. For example, a well known predictor of risk is the number of friends an individual has who use marijuana. However, there is a possibility that the friend's use may be an effect of the individual's use as well as a cause. Including such variables in the risk model would have created ambiguity in the interpretation of the risk measure, in its relationship to possible Media Campaign effects. Where it was possible, some variables that could have held such ambiguous relationships were constructed so that they would not. Thus, child cigarette and alcohol use as antecedent covariates are well established in the literature; the measures used here were constructed so as to avoid capturing reciprocal effects between them and marijuana use. Only cigarette or alcohol use that had occurred more than 1 year prior to the interview was included. Given the cross-sectional nature of the data, other promising risk covariates were excluded in order to avoid such causal ambiguity; for example, marijuana offers, association with deviant peers, and child-parent conflict, among others.

Table 5-12 presents the results for the final model.[3] The strongest predictors are: having started smoking prior to the past 12 months, sensation seeking, age, and having started drinking prior to the past 12 months, all of which are youth characteristics and behaviors. To ease interpretation, the last column presents the adjusted odds ratio estimates. Youth who had started using cigarettes prior to the past year had substantially increased odds (odds ratio = 4.43) to use marijuana in the past year compared to children who had not started smoking prior to the past 12 months. Each 1-point increase in the child's sensation-seeking tendencies was associated with an increase of 117 percent in the odds of marijuana use in the past 12 months. Each 1-year increase in age was associated with a 42 percent increase in the odds of marijuana use in the past 12 months. Youth who had started drinking prior to the past year had twice the odds of using marijuana in the past year, than did children who had not started alcohol use before that period. Youth living in large urban areas had 31 percent greater odds of having used marijuana in the past year than youth living in towns and rural areas.

[3] With regard to the analytical procedure, the data set was split into 10 random groups; one of these was randomly dropped, and a logistic regression model was fitted to the remaining 9 groups. The fitted model was then used to assign the risk scores of persons in the omitted group. The logistic regression model was run so that each of the 10 groups was dropped in turn, resulting in a cross-predicted risk score for every person in the sample. In a second step, all 10 models were rerun using only variables that had been found to be significant in any of the previous analyses. Coefficients were averaged across these latter 10 models, and they were the basis for the cross-predicted probability.

Table 5-12. Youth and parent covariates for youth past year marijuana use

	Estimate	Standard error	Wald χ^2	Significance level	Odds ratio*
Intercept	-9.9651	.5842	290.95	<.0001	
Youth covariates					
Age (12-18)	.3530	.0323	119.29	<.0001	1.42
Sensation seeking (high versus low)	.7730	.0692	124.83	<.0001	2.17
Started smoking 12+ months ago	1.4890	.1250	141.95	<.0001	4.43
Started drinking 12+ months ago	.7655	.1216	39.62	<.0001	2.15
Urbanicity 1 (urban vs. rural)	.2704	.0815	11.02	.0009	1.31
Urbanicity 2 (suburban versus rural)	-.0036	.0852	0.00	.9661	1.00
Parent covariates					
Marijuana use in past 5 years	.3361	.1678	4.01	.0451	1.40
Cigarette use in past month	.4127	.1233	11.19	.0008	1.51
Had no drink in past month	-.1727	.1180	2.14	.1433	0.84
Attendance at religious services	-.0943	.0656	2.07	.1502	0.91
Rating of importance of religion	-.0768	.0713	1.16	.2816	0.93
Shares parenting with other adult in household	-.4396	.1186	13.74	.0002	0.64

* Likelihood of a youth using marijuana in the past year.

The strength of parental factors included in the model was, overall, of lower magnitude and some variables did not achieve statistical significance at the conventional level ($p<0.05$) in the final model. Youth from households in which parenting is shared have only 0.64 times the odds of using marijuana in the past year as youth living in single-parent households. Youth whose parent reported tobacco use in the past month had 1.5 times greater odds of using marijuana in the past year than youth whose parent had not smoked cigarettes in the preceding month. Likewise, parental marijuana use was associated with a 39 percent increase in the odds of child past-year marijuana use.[4]

Across Round 1 of NSPY data collection, the sample used to develop the risk model, only about 11.5 percent of youth reported marijuana use during the preceding year. Given such a low base rate, the risk probabilities for nonusers tend to be fairly low. The average 12- to 18-year-old had about a 12 percent predicted probability of annual marijuana use, with half of the youth having less than a 4 percent risk of use.

[4] Covariates that did not make it into the risk measure are wave, youth gender, youth race/ethnicity, parent binge drinking in past 30 days, age of parent, parental education, and annual household income.

About a third of the sample was classified as at higher risk (set at having a risk of use equal to or greater than 8 percent). While an 8 percent cutoff may seem low, this measure represents a relative risk and not an absolute risk, hence the use of the terms "higher" and "lower." There are, nevertheless, considerable differences in youth reports of marijuana use by risk group. Because child's age is an important covariate in the risk model, it is important to determine whether the differences by risk group disappear when controlling for age. Table 5-13 presents the results for different measures of marijuana use by age and risk subgroups for the first 2 years of the NSPY.

Table 5-13 shows that marijuana use reported by youth at higher risk tends to be about 10-fold that reported by youth at lower risk. This is true for lifetime, past year, and past month marijuana use, and across age groups. For example, among 12- to 13-year-olds, 0.6 percent of youth at lower risk and 6 percent of youth at higher risk reported past month marijuana use in 2001. Among the older group, a little more than 2 percent of youth at lower risk and nearly 19 percent of youth at higher risk reported past month use in the same period.

Table 5-13. Percent of youth reporting marijuana use by age and risk subgroup

Use measure	Youth characteristics		Year 2000 Average Waves 1 & 2	Year 2001 Average Waves 3 & 4
	Age group	Risk group		
Lifetime				
	12 to 13	Higher	34.1	22.4
		Lower	2.1	2.1
	14 to 18	Higher	49.3	52.0
		Lower	5.3	6.9
Past year				
	12 to 13	Higher	23.0	18.4
		Lower	1.5	1.3
	14 to 18	Higher	37.2	36.0
		Lower	3.4	4.9
Past month				
	12 to 13	Higher	11.3	6.2
		Lower	0.4	0.6
	14 to 18	Higher	17.8	18.9
		Lower	1.0	2.4
Regular				
	12 to 13	Higher	5.2	2.3
		Lower	0.0	0.1
	14 to 18	Higher	14.7	14.9
		Lower	0.4	1.8

REFERENCES

Hornik, R., Maklan, D., Cadell, D., Barmada, C., Jacobsohn, L., Henderson V., Romantan, A., Niederdeppe, J., Orwin, R., Sridharan, S., Chu, A., Morin, C., Taylor, K., and Steele, D. (2003). *Evaluation of the National Youth Anti-Drug Media Campaign: 2003 Report of Findings.* Report Prepared for the National Institute on Drug Abuse (Contract No. N01DA-8-5063), Washington, DC: Westat.

Hornik, R., Maklan, D., Judkins, D., Cadell, D., Yanovitzky, L., Zador, P., Southwell, B., Mak, K., Das, B., Prado, A., Barmada, C., Jacobsohn, L., Morin, C., Steele, D., Baskin, R., and Zanutto, E. (2001). *Evaluation of the National Youth Anti-Drug Media Campaign: Second semi-annual report of findings.* Report Prepared for the National Institute on Drug Abuse (Contract No. N01DA-8-5063), Washington, DC: Westat.

Southwell, B., Barmada, C., Hornik, R., and Maklan, D. (2002). Can we measure encoded exposure? Validation evidence from a national campaign. *Journal of Health Communication.* 7, 445-453.

6. FILE CONTENT AND DESCRIPTION OF VARIABLES

This chapter describes the content of the NSPY Public Use Files (PUFs). Section 6.1 covers how the PUFs were constructed, Section 6.2 describes the identifiers on the files, Section 6.3 describes the creation of recoded variables, Section 6.4 describes the use of the weights, Section 6.5 discusses missing values and Section 6.6 describes the format and content of the codebooks.

6.1 Construction of Data Files

Each PUF is a SAS Version 8 data set. There are three PUFs, one for each round of data collection. Within each file there is one record per youth included in the PUF subsample. The PUF for Round 1 contains a record for every responding youth in that round, covering the age range from 9 to 18 years old. The PUFs for Rounds 2 and 3 are restricted to youth in the age range from 12 to 18 years old. Each record contains responses to interview variables taken directly from the survey instrument, plus recoded variables and index variables (see Chapter 5). In addition to variables collected or derived from the youth interview, eight variables collected from the parent interview were added to the corresponding youth record. The annotated survey instruments in Appendix B can be used to determine which interview variables are and are not included in the PUFs. Variables with an asterisk following their names are not included in the PUFs, whereas those without an asterisk are included (perhaps in coarsened form). Furthermore, the annotated survey instruments give the name(s) of any recoded variable(s) associated with each interview item that is included in the PUFs. Also, the codebooks (see Appendix A) contain unweighted sample distributions, or ranges for continuous variables, for all variables found in the PUFs.

6.2 Identification Variables

In each PUF, every record has two identification variables. The first is ROUND, which has a value of 1, 2, or 3. The second is PUFID, which uniquely identifies each youth within a PUF. The identifier PUFID contains a random component. Hence, PUFID cannot be used to “link” youth records across the three PUFs, nor does it in any way disclose the true identity of the respondent.

As discussed in Section 3.2, random subsampling of youth was employed as a means of disclosure avoidance. To add further protection, variables that might be used to attempt to identity a youth were coarsened or suppressed from the PUFs. Basic demographic information, such as urbanicity, age group, gender, and race/ethnicity, was retained but some other data were suppressed if they were deemed to be identifying.

6.3 Recoded Variables

Recoded variables are derived from one or more interview variables. The algorithms that were used to define the recoded variables on the PUFs can be found in the codebooks (see Appendix A). Where a recoded variable is derived from only one interview variable, the recode represents a coarsening of the interview variable response categories. Such is the case with AGEGROUP (see Exhibit 6-1).

AGEGROUP A1(T,C) YOUTH AGE AT INTERVIEW [0013-0014 N]

1 = 09-11	1991
2 = 12-13	1623
3 = 14-16	1238
4 = 17-18	559

Exhibit 6-1. AGEGROUP

Recoded variables derived from two or more interview variables tend to be more complex in nature. Such recodes are generally designed either to account for "skip" patterns in the instrument, or to create a combined measure that summarizes a number of related interview variables. For example, MJREG (see Exhibit 6-2) accounts for a skip pattern and rescales the original value to create a dichotomous variable for analysis.

MJREG B11(T,C) - B14(T,C), B15(C) USED MJ REGULARLY LAST 12M
[0069-0070 N]

If YMJ12M = 4,5,6 then MJREG = 1
Else if YMJ12M = 1,2,3 or YMJEVER = 2 or YMJLAST = 3 or CMJHEAR = 2 then MJREG = 0

0 = No	5240
1 = Yes	130
.M = Not Applicable	41

Exhibit 6-2. MJREG

Alternatively, a recoded variable may have been created to deal with the situation where the child and teen instruments asked similar, but not identical, questions (see Exhibit 6-3). In this case the recoded variable represents a derived measure that is consistent for all youth.

YNONUSER B11(T,C), B12(C) MARIJUANA NON-USER [0063-0064 N]

```
If YTYPE = T then
   If YMJEVER = 2 then YNONUSER = 1
   Else if YMJEVER = 1 then YNONUSER = 0
Else if YTYPE = C then
   If CMJHEAR = 2 or YMJEVER = 2 then YNONUSER = 1
   Else if YMJEVER = 1 then YNONUSER = 0
```

0 = No	580
1 = Yes	4802
.M = Not Applicable	29

Exhibit 6-3. YNONUSER

6.4 Weight Variables

There are 101 weight variables for each respondent included in the PUF. The 101 weight variables are named WEIGHT, REPLW1, REPLW2, ..., REPLW100. WEIGHT is the overall (raked) weight described in Section 3.3. This weight should always be used to derive national estimates from the data contained in the PUF. For example, all population means and proportions should be calculated as weighted means and weighted proportions, using WEIGHT as the case (record) weight.

The corresponding replicate weights, REPLW1, REPLW2, ..., REPLW100, are provided in the PUF to permit the calculation of sampling errors using replication techniques. These replicate weights should not be used to make national estimates. For example, software packages such as WesVar and SUDAAN can be used for this purpose. Since the replicates used in NSPY were specifically designed to reflect certain features used in PSU selection, it is necessary to apply a set of replicate-specific factors in the calculation of the sampling variances. The required factors are: 2.57 for replicates 1-60, and 0.06 for replicates 61-100. In WesVar, these factors can be entered using the "JKN" option. SUDAAN will also accept replicate weights and user-supplied factors for variance estimation.

Some examples illustrating the use of WesVar and SUDAAN to analyze NSPY data are given in Appendix C. For additional information about the role of the weights in producing national

estimates and standard errors, see Westat (2000), WesVar 4.0 User's Guide, Rockville, MD, or Research Triangle Institute (2002), SUDAAN User's Manual, Research Triangle Park, NC.

6.5 Missing Values

A standard scheme for missing values was applied to all variables on the PUFs. Four SAS special missing values were used, three for interview variables and one for recoded variables, when constructing the files and tallying frequencies for the codebooks. These are:

Interview variables		Recoded variables	
ASC11	SAS	ASC11	SAS
-8	.D = Don't know	-6	.M = Missing value
-9	.R = Refused		
-7	.N = Not applicable		

The value .N only applies to interview variables. A missing value of ".N" may mean that the interview variable in question was part of a "skip" pattern in the survey instrument design or that the interview variable was unique to the Teen instrument (for a child record) or unique to the Child instrument (for a teen record).

The value .M only applies to recoded variables. Recoded variables have only one missing value format, ".M", due to the difficulty of propagating meaningful missing values when a recode is derived from a number of different interview variables.

6.6 Codebooks

Appendix A contains the codebooks for the three PUF data sets. Each codebook lists all the variables on the PUF in the order that they appear in the survey instruments. Each codebook is divided into four sections, with section headings printed in bold text, and the first section contains subheadings to distinguish the topical areas of the teen and child instruments:

1. Youth Interview
    a. Demographics And Time Use

b. Drug Experience

c. Expected Drug Experience

d. Exposure To Drug Information

2. Indices

3. Parent Interview000000

4. Youth Weights

All variables are described using the layout described in Exhibits 6-4 and 6-5.

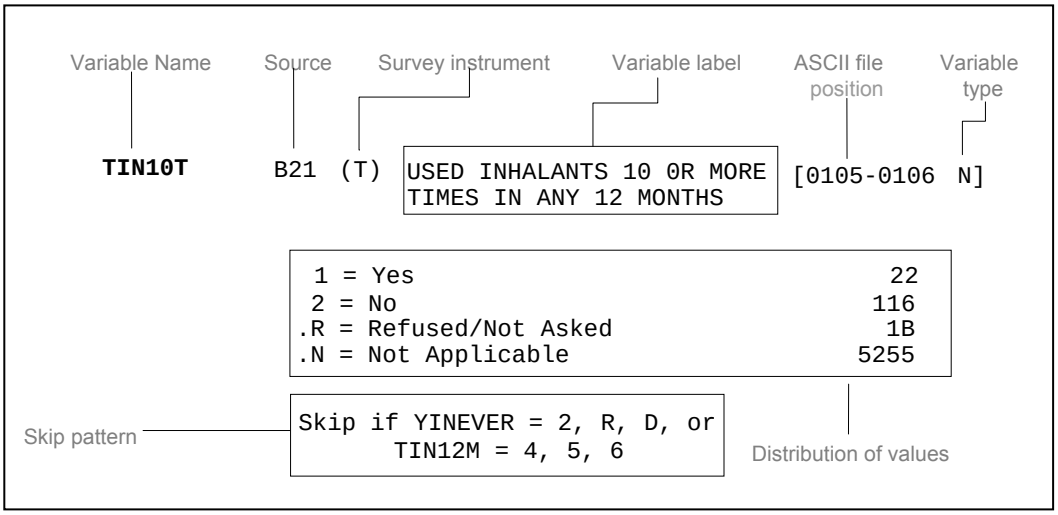

Exhibit 6-4. TIN10T (Categorical Variable)

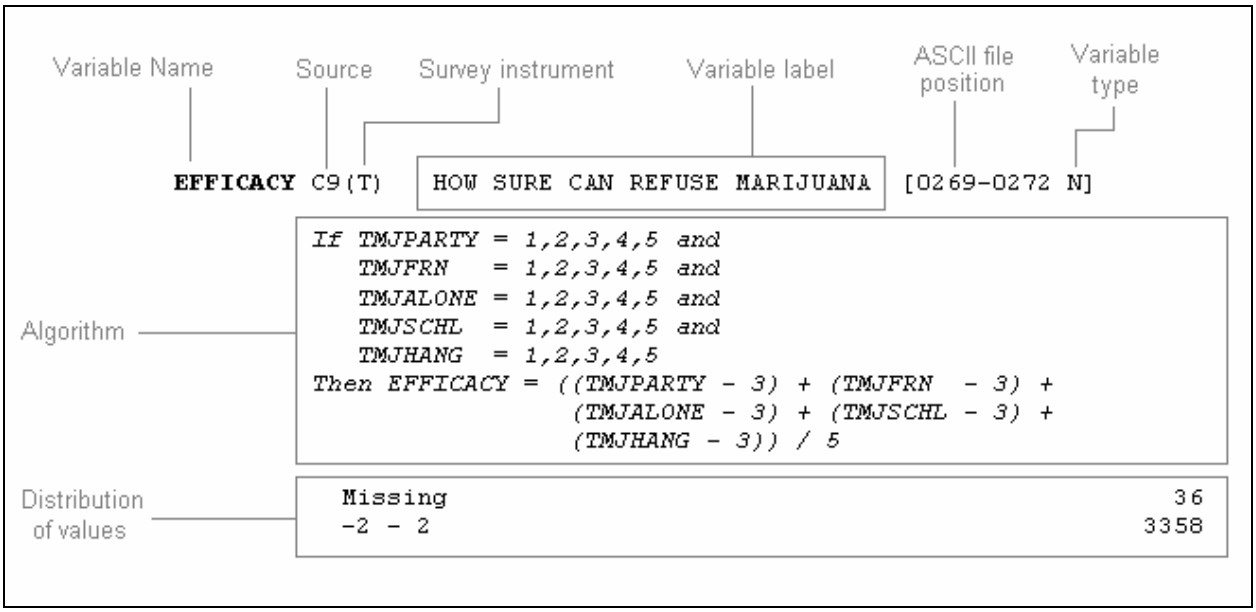

Exhibit 6-5. EFFICACY (Continuous Variable)

Key aspects of the layout include:

Variable Name. A maximum of eight characters, printed in bold type. (See Exhibit 6-4.) The variable name is useful as a cross-reference to the annotated survey instrument in Appendix B.

Source. The question number from the survey instrument. (See Exhibit 6-4.) (Source is not applicable to youth indices or youth weights.)

Survey Instrument. One letter designator, where T = Teen or C = Child or P = Parent. If the variable exists in more than one survey instrument, multiple values are listed, e.g., T, C. (See Exhibit 6-4.) (Survey instrument is not applicable to youth indices or youth weights.)

Variable Label. SAS data set variable label, a maximum of forty characters. (See Exhibit 6-4.)

ASCII File Position. Column position in ASCII data file. (See Exhibit 6-4.)

Variable Type. All variables are numeric ([N]) except PUFID and YOUTHTYP, which are character ([C]). (See Exhibit 6-4.)

Algorithm. Description of recode method or reference to User's Guide which discusses the process. (See Exhibit 6-5.) The algorithm is italicized and is applicable only to recodes and indices.

Distribution of Values. This area has two formats, dependent on whether the variable is categorical or continuous. For a categorical variable (see Exhibit 6-4) the layout is data value, value label and record count. For a continuous variable (see Exhibit 6-5) or a variable with more than twelve discrete values, the layout is minimum value, the dash symbol ('-') and maximum value and any missing value is labeled "missing."

Skip Pattern. Applicable only to interview variables, text is italicized. (See Exhibit 6-4.)

The codebooks are designed to be used in close conjunction with the annotated survey instruments that are found in Appendix B. The annotated survey instruments display the variable names associated with the various survey instrument items. When using the annotated survey instruments to explore the data collected for NSPY, the variable names can be used to provide a link to the information in the codebooks. In addition, the codebooks will reference sections of the annotated survey instruments that might be helpful in understanding a particular variable. For example, Section 5.2 in the User's Guide describes how the index variable MJATTBEL was constructed. (See Exhibit 6-6.)

MJATTBEL ANTI-MJ BELIEF/ATTITUDE INDEX [0480-0488 N]

See Chapter 5 of PUF user manual for description of how index is calculated.

Missing	1999
-228.706 - 264.432	3412

Exhibit 6-6. MJATTBEL

At the end of the Round 3 codebook (page A-204), there is an alphabetic index of variables with the codebook page numbers by round, where the variable is documented. This will be useful when looking for the codebook entry for a variable referenced in the annotated survey instruments or elsewhere in the User's Guide.

As mentioned in Section 4.5.3, the type of survey instrument (Child or Teen) to be administered to the youth was determined at the time of household screening based on the reported age of the youth. However, in a small number of cases, children aged 9 to 11 years completed the Teen instrument or teens aged 12 to 18 completed the Child instrument. Sometimes this occurred because the youth turned 12 between the time of household screening and the extended interview process. More commonly, the wrong instrument was administered because the age of the youth reported by a household member age 18 or older at household screening was incorrect. Many items are identical between the Child and Teen instruments and therefore deleting these cases from the PUFs would have resulted in unnecessary information loss. Instead, only child- or teen-specific question items that the youth should not have been asked were set to missing, and any recoded variables based on these items subsequently recomputed. The user can identify these cases in the PUFs by reconciling type of instrument (variable YTYPE) with youth age group (recode variable AGEGROUP), should it be deemed necessary.

REFERENCES

Research Triangle Institute (2001). SUDAAN User's Manual, Release 8.0, Research Triangle Park, NC: Research Triangle Institute.

WesVar (2003). WesVar 4.2 User's Guide, Westat: Rockville, MD (http://www.westat.com).