

1999-2001 NATIONAL HOUSEHOLD SURVEY ON DRUG ABUSE

CHANGES IN RACE AND ETHNICITY QUESTIONS

Contract No. 283-98-9008
Project 7190 - 1999-2003 NHSDA

Authors:

Jeanne A. Snodgrass
Eric A. Grau
Rachel A. Caspar

Project Director:

Tom Virag

Prepared for:

Substance Abuse and Mental Health Services Administration
Rockville, MD 20857

Prepared by:

RTI
Research Triangle Park, NC 27709

October 18, 2002

1999-2001 NATIONAL HOUSEHOLD SURVEY ON DRUG ABUSE

CHANGES IN RACE AND ETHNICITY QUESTIONS

Contract No. 283-98-9008
Project 7190 - 1999-2003 NHSDA

Authors:

Jeanne A. Snodgrass
Eric A. Grau
Rachel A. Caspar

Project Director:

Tom Virag

Prepared for:

Substance Abuse and Mental Health Services Administration
Rockville, MD 20857

Prepared by:

RTI
Research Triangle Park, NC 27709

October 18, 2002

Acknowledgments

This report was developed for the Substance Abuse and Mental Health Services Administration (SAMHSA), Office of Applied Studies (OAS), by RTI under Contact No. 283-98-9008. Contributors at RTI listed alphabetically include Peter Frechtel, Amy Licata, Lisa Packer, and Tom Virag (Project Director). The report was edited by Richard Straw and word processed by Cheri Thomley. Pam Couch Prevatt, Teresa F. Gurley, and Kim Cone prepared the Web versions. The reviewers at SAMHSA included Art Hughes and Joseph C. Gfroerer.

Originating Office:
SAMHSA, OAS
5600 Fishers Lane, Room 16-105
Rockville, MD 20857

October 18, 2002

Table of Contents

List of Appendices	ii
List of Tables	ii
1.0 Introduction.....	1
2.0 Revision of the Race and Ethnicity Questions in the NHSDA.....	1
2.1 The 1999 NHSDA.....	3
2.2 The 2000 NHSDA.....	4
2.3 The 2001 NHSDA.....	5
3.0 Editing and Imputation of Race Variables.....	6
3.1 Introduction.....	6
3.2 PAPI Versus CAI.....	6
3.2.1 Race.....	6
3.2.2 Ethnicity (Hispanic Indicator).....	7
3.2.3 Race and Ethnicity	8
4.0 IRRACE.....	8
4.1 IRRACE and IIRACE: PAPI.....	9
4.1.1 Creation of Preliminary Variables	9
4.1.2 Creation of Base Variable for Imputation.....	10
4.1.3 Creation of IRRACE and IIRACE.....	10
4.2 IRRACE and IIRACE: CAI, 1999.....	11
4.2.1 Creation of Base Variable for Imputation.....	11
4.2.2 Creation of IRRACE and IIRACE.....	13
4.3 IRRACE: CAI, 2000 and 2001	14
4.3.1 Creation of Base Variable for Imputation.....	14
4.3.2 Creation of IRRACE, IIRACE, and II2RACE	15
4.4 Summary of Changes in IRRACE: 1998-2001.....	17
5.0 IRHOIND.....	17
5.1 IRHOIND, PAPI.....	18
5.2 IRHOIND and IIHOIND, CAI	18
5.2.1 Creation of Base Variable for Imputation.....	18
5.2.2 Creation of IRHOIND and IIHOIND	18
6.0 NEWRACE1 and IRNWRACE.....	19
6.1 NEWRACE1 and IINWRAC1, 1999	20
6.1.1 Background for Creating NEWRACE1	20
6.1.2 Creation of NEWRACE1.....	21

Table of Contents (continued)

6.2	NEWRACE1, IRNWRACE, and IINWRACE, 2000 and 2001	22
6.2.1	Creation of Base Variable for Imputation of IRNWRACE	22
6.2.2	Creation of IRNWRACE, IINWRACE, and NEWRACE1	24
7.0	Other-Specify Codes for Race and Hispanicity	25
7.1	Approach to Dictionary in PAPI	26
7.2	Approach to Dictionary in CAI	26
8.0	Recoded Variables	29
8.1	RACE and HISPRACE	29
8.2	EXPRACE, NEWRACE1, NEWRACE2, and RACE4	30
9.0	Analytic Tables Using Race Variables	32
9.1	Tables Involving Race and Ethnicity, 1998	32
9.1.1	Main Findings Report	32
9.1.2	Summary Tables	33
9.2	Tables Involving Race and Ethnicity, 1999	33
9.3	Tables Involving Race and Ethnicity, 2000 and 2001	35
References	36

List of Appendices

Appendix A	Race and Ethnicity Questions in 1971 to 2001 NHSDAs	A-1
Appendix B	Main Race Question Responses from the 2001 NHSDA	B-1

List of Tables

Table 1	2001 NHSDA - Main Race Question (QD06)	B-1
Table 2	Distribution of 2001 NHSDA Respondents with Multiple Race, by Main Race (Unweighted)	B-2
Table 3	Distribution of 2001 NHSDA Respondents with Multiple Race, by Main Race and Hispanic Origin (Unweighted)	B-3
Table 4	Distribution of 2001 NHSDA Respondents with Multiple Race, by Main Race and Age Group (Unweighted)	B-4
Table 5	Respondent Comments to the 2001 NHSDA Main Race Question	B-5

Changes in Race and Ethnicity Questions

1.0 Introduction

Since the inception of the National Household Survey on Drug Abuse (NHSDA), renamed the National Survey on Drug Use and Health (NSDUH) as of 2002, questions have been included to determine the race and ethnicity of each respondent. Race and ethnicity are routinely used as part of the demographic breakdowns in the analyses and the various reports generated from the survey. From 1971 to 1998, the race and ethnicity questions underwent few changes. (See Appendix A for the full list of race and ethnicity questions used for each NHSDA survey year from 1971 to 2001.) However, along with the switch from paper-and-pencil interviewing (PAPI) methods of questionnaire administration to computer-assisted interviewing (CAI) methods in 1999, the race and ethnicity categories were updated pursuant to new Office of Management and Budget (OMB) directives. This report details the revisions to the race and ethnicity questions. It includes the history of the change, how this change affected the editing and imputation procedures, and how it changed the derivation of the race and ethnicity variables used in NHSDA analyses.

2.0 Revision of the Race and Ethnicity Questions in the NHSDA

Prior to 1999, the instrument for the NHSDA included two questions to collect data on race and ethnicity in a manner consistent with OMB's 1977 Statistical Policy Directive No. 15 (OMB, 1977). Directive No. 15 was developed to provide a common language to promote uniformity and comparability across Federal surveys for data on race and ethnicity. Development of these data standards stemmed in large measure from the need to enforce civil rights laws.

In response to criticisms that the minimum categories set forth in Directive No. 15 did not reflect the diversity of the U.S. population, OMB announced in July 1993 that it would undertake a comprehensive review of the categories used to collect data on race and ethnicity. OMB established the Interagency Committee for the Review of the Racial and Ethnic Standards in 1994. This committee worked with Federal agencies to develop recommendations for enhancing the accuracy of the data on race and ethnicity collected by the Federal Government. The committee's work included a request for public comment on the Directive No. 15 standards and research and testing to assess the possible effects of implementing the suggested changes on the quality and utility of the resulting data.

In October 1997, OMB released a notice, "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity" (OMB, 1997), that summarizes the results of both the research and the public comment. It also provides the new standards for maintaining, collecting, and presenting Federal data on race and ethnicity. The standard race question now includes five categories: American Indian or Alaska Native, Asian, black or African American, Native Hawaiian or Other Pacific Islander, and white. There are two categories for data on ethnicity: Hispanic or Latino and Not Hispanic or Latino. The OMB notice provides the following definitions for all of these categories:

American Indian or Alaska Native. A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

Asian. A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

Black or African American. A person having origins in any of the black racial groups of Africa. Terms such as "Haitian" or "Negro" can be used in addition to "black or African American."

Hispanic or Latino. A person of Cuban, Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race. The term "Spanish origin" can be used in addition to "Hispanic or Latino."

Native Hawaiian or Other Pacific Islander. A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

White. A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

The notice states that respondents must be offered the option of selecting one or more racial designations. The notice also suggests that to provide flexibility and ensure data quality, separate questions should be used wherever feasible for reporting race and ethnicity. When race and ethnicity are collected separately, ethnicity should be collected first.

Finally, the notice states that when aggregate data are presented, data producers should provide the number of respondents who selected only one category, separately for each of the five racial categories. In addition, analysts are strongly encouraged to provide detailed distributions, including all possible combinations, of multiple responses for the race question. If data on multiple responses are collapsed, at a minimum the total number of respondents reporting more than one race should be provided as part of the data file.

OMB indicated that the new standard for collecting race and ethnicity superseded all previous requirements and for new or revised data collection activities, the standards were to be put into place by January 2003. The NHSDA instrument conversion from PAPI to CAI in 1999 afforded the opportunity to make this update. The new standards for collecting race and ethnicity were incorporated into the newly computerized 1999 NHSDA data collection instrument.

In October 2000, OMB approved the 2001 NHSDA with the conditional term that the Substance Abuse and Mental Health Services Administration (SAMHSA) include a main race question, and that they report to OMB on the effects of this question. In particular, SAMHSA was to report to OMB on the interaction between Hispanic status, age, and responses to the race question. OMB used this information for research purposes in support of the Tabulation Working Group of the Interagency Committee for the Review of Standards for Data on Race and Ethnicity. The requested data was sent to OMB in April 2002 (see Appendix B).

2.1 The 1999 NHSDA

Prior to 1999, the NHSDA PAPI instrument included the following ethnicity question (D-3):

Are you of Hispanic or Spanish origin or descent?

A race question (D-5) with only four categories was also included:

WHITE

BLACK

AMERICAN INDIAN OR ALASKAN NATIVE

ASIAN OR PACIFIC ISLANDER (INCLUDING ASIAN INDIAN)

In keeping with the new standard, the 1999 NHSDA instrument included the same two questions to collect race and ethnicity, only modified. The ethnicity item appeared first in the instrument as it did prior to 1999, but included the word “Latino.” It was read by the interviewer who then entered the respondent’s answer into the computer-assisted personal interviewing (CAPI) program. The question (QD03) asked:

Are you of Hispanic, Latino, or Spanish origin or descent?

1 YES

2 NO

It should be noted that OMB’s 1997 notice indicates that use of the phrase “Spanish origin” may also be included in the question text if desired.

In 1999, the race question (QD05) was also included in the CAPI module and was worded as follows:

HAND R SHOWCARD 2. Which of these groups describes you? Just give me the number or numbers from the card.

TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT.

1 WHITE

2 BLACK / AFRICAN AMERICAN

3 AMERICAN INDIAN OR ALASKA NATIVE

4 NATIVE HAWAIIAN

5 OTHER PACIFIC ISLANDER

6 CHINESE

7 FILIPINO

- 8 JAPANESE
- 9 ASIAN INDIAN
- 10 KOREAN
- 11 VIETNAMESE
- 12 OTHER ASIAN
- 13 OTHER (SPECIFY)

In administering this question, the interviewer handed the respondent a showcard that included all of the categories shown above with the exception of the “Other, specify” category. Although this question includes more categories than explicitly included in OMB’s 1997 notice, the categorization can be collapsed to provide the level of detail required by OMB. The 1997 notice explicitly states that additional detail is allowable so long as the listing allows for back coding to the original six categories defined in the notice.

A follow-up item was also included in the 1999 instrument for respondents who selected more than one race. This item was included to reduce the amount of imputation needed for the new race question and to ensure categories could be back coded to the original mutually exclusive race categories. The item (QD06) asked:

Which **one** of these groups, that is [LIST RACES SELECTED FROM QUESTION ABOVE], **best** describes you?

SELECT ONLY **ONE** ANSWER.

The same race categories included in QD05 above were used for this item. However, only the categories that had originally been entered by the interviewer were available to be selected. For example, if a respondent had originally selected white and Korean, then Japanese could not be selected as the best descriptor in the follow-up item.

2.2 The 2000 NHSDA

The ethnicity item remained unchanged for the 2000 NHSDA. However, a few revisions were made to QD05. Most notably, the categories were revised to align more directly with OMB’s 1997 notice. Additional information on specific Asian races that had been collected within the main race question in 1999 were moved to a followup question. Also, category 3 was revised to provide an explicit definition of “American Indian.” The exact wording of the race series of questions was as follows:

HAND R SHOWCARD 2. Which of these groups describes you? Just give me the number or numbers from the card.

TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT.

- 1 WHITE

- 2 BLACK / AFRICAN AMERICAN
- 3 AMERICAN INDIAN OR ALASKA NATIVE (AMERICAN INDIAN INCLUDES NORTH AMERICAN, CENTRAL AMERICAN, AND SOUTH AMERICAN INDIANS)
- 4 NATIVE HAWAIIAN
- 5 OTHER PACIFIC ISLANDER
- 6 ASIAN (FOR EXAMPLE: ASIAN INDIAN, CHINESE, FILIPINO, JAPANESE, KOREAN, AND VIETNAMESE)
- 7 OTHER (SPECIFY)

For respondents who selected category 6, an additional question (QD05ASIA) was included:

HAND R SHOWCARD 3. Which of these Asian groups best describes you? Just give me the number or numbers from the card.

- 1 ASIAN INDIAN
- 2 CHINESE
- 3 FILIPINO
- 4 JAPANESE
- 5 KOREAN
- 6 VIETNAMESE
- 7 OTHER (SPECIFY)

As in the 1999 NHSDA, any respondent who reported more than one race was asked to indicate the one race that best described him or her. In the 2000 NHSDA, this required a reprogramming of the CAI questionnaire so that multiple races reported in either of the two race items—the original question and the follow-up item to determine the specific Asian group—would trigger this main race question.

2.3 The 2001 NHSDA

Only one change was made to the race and ethnicity items in 2001. An additional category was added to the “single best race” question for respondents who reported that **none** of the races they had reported was the best descriptor of their race. This category was added based on discussions with interviewers who noted that some respondents who reported multiple races could not select the one best race descriptor for themselves. Indirectly related to this change, an interviewer instruction was also added to this screen. The instruction directed interviewers to enter a comment if a respondent expressed any concern about what data on “single best race” was being collected. Interviewers were instructed not to probe for such information but to record any spontaneous comments made by the respondents.

3.0 Editing and Imputation of Race Variables

3.1 Introduction

With the introduction of the CAI changes in the questions concerning race, and the decision to reduce the use of screener information when editing race and ethnicity variables, the subsequent processing of the race and ethnicity variables changed substantially. In this section, the methods used to process the data from the race and ethnicity questions are summarized and contrasted. In keeping with OMB guidelines, “Hispanic/Latino” is considered an ethnicity, not a race. However, many respondents identified their race as “Hispanic/Latino” in the other-specify response in both PAPI and CAI, resulting in a considerable amount of missing data for the race question. For this reason, both before and after the introduction of the CAI instrument, drug use patterns were tabulated with race and ethnicity considered together as cross-classifying variables (in effect, treating “Hispanic/Latino” as a race). This document summarizes the tables that were created using race and ethnicity. It does not discuss in detail the processing of the variables related to the specific Hispanic categories.

3.2 PAPI Versus CAI

3.2.1 Race

In the PAPI NHSDA questionnaire prior to 1999, and in the 1999 PAPI supplemental sample, a single question was asked about race (D-5). This question provided four categories for the respondent to choose from, plus a fifth category where respondents were given the opportunity to have the interviewer manually enter an alternative to the given four categories, denoted as an “other-specify” (or “alpha-specify”) response. These other-specify responses were incorporated into the other four race categories in the final imputation-revised race variable, IRRACE. If the other-specify response indicated more than one race, a priority rule was used to determine the single race category of IRRACE: black/African American, Asian/Pacific Islander, American Indian/Alaska Native, or white. This priority rule was also invoked in the rare instance where a race was entered in “other-specify” for the Hispanic groups question (D-4) instead of a Hispanic group, and that race disagreed with the one given in D-5.

In contrast to PAPI, several questions about race were asked in the 1999 CAI questionnaire, including 12 race categories. There was also the opportunity to manually enter a race if the other categories did not apply. Respondents could also have the interviewer enter multiple categories if needed. To maintain consistency with the PAPI results, the variable IRRACE was created using the CAI questions. (IRRACE was also used as a covariate in models for other variables requiring imputation.) However, no four-category question was asked in the CAI questionnaire, making it necessary to combine levels from the race questions.

If a single race was entered in the initial race question (QD05), this was the race used in the creation of IRRACE, where the categories white, black, and American Indian/Alaska Native mapped directly to the similarly named categories in the PAPI version of IRRACE. To create the “Asian/Pacific Islander” level of IRRACE, however, it was necessary to collapse the categories Native Hawaiian, other Pacific Islander, Chinese, Filipino, Japanese, Asian Indian, Korean, Vietnamese, and other Asian from the 1999 questionnaire, and Native Hawaiian, other Pacific Islander, and Asian from the 2000 and 2001 questionnaires.

If multiple races were entered in QD05, respondents were given the opportunity to enter the race with which they most closely identified themselves, and the levels of this variable associated with question QD06 were mapped to the levels of IRRACE in the same way as the initial race questions. The same priority rule was used as in PAPI if the respondent gave multiple races in QD05 and did not identify a single race in QD06. It should also be noted that, in PAPI, the screener information was used directly to fill in missing information in the final race variable, whereas in CAI, screener information was only indirectly used (in models).

3.2.2 Ethnicity (Hispanic indicator)

In general parlance, “ethnicity” is a general term that refers to a person’s ethnic background. However, in the NHSDA, “ethnicity” is a term that refers specifically to whether a respondent considers himself or herself Hispanic/Latino or not Hispanic/Latino. “Hispanicity” is also a term that is sometimes used in the NHSDA in place of ethnicity. Prior to the question about race in PAPI, respondents were first asked (in D-3) their “ethnicity” (whether they were Hispanic or not), then they were asked (in D-4) to identify the specific Hispanic group to which they belonged (where an alternative to the Hispanic categories provided could be entered as an “other-specify” response). Because this was a paper-and-pencil questionnaire, it was possible for a respondent to have a response for D-4, even though no affirmative response to D-3 was given. The final version of the binary indicator of a respondent’s ethnicity was given by the variable IRHOIND.

As with PAPI, respondents were first asked to identify their ethnicity in the 1999 CAI questionnaire (Hispanic/Latino vs. not Hispanic/Latino, in QD03). If they were Hispanic, they were asked to identify the Hispanic group to which they belonged in QD04, with an opportunity to enter an other-specify response if the given Hispanic categories did not apply to them. The only difference between the ethnicity questions in PAPI (D-3) and CAI (QD03) is the reference to Latino in CAI. Moreover, in PAPI, respondents had the opportunity to indicate a Hispanic group even if they did not respond affirmatively (or respond at all) to the Hispanicity question in D-3, which was not the case in CAI. However, the major difference between CAI and PAPI was the processing of the final Hispanic indicator variable, IRHOIND. In PAPI, screener information was used to fill in any gaps in information not provided by questions D-3 and D-4, so that no imputation was necessary. (The presence of the IR prefix in the PAPI version of the variable, which usually means “imputation-revised,” is therefore misleading.) On the other hand, missing values were imputed in the CAI version of IRHOIND, and no screener information was used.

3.2.3 Race and Ethnicity

As stated earlier, the variables used in tabulated results involved both race and ethnicity simultaneously, due to the large number of respondents who considered “Hispanic/Latino” to be a race. The 16-level variable for race and ethnicity, NEWRACE1, was created in 1999 and subsequent years to reflect this policy and to accommodate the 12 race categories in the CAI questionnaire: white, black/African American, American Indian/Alaska Native, Native Hawaiian, other Pacific Islander, Chinese, Filipino, Japanese, Asian Indian, Korean, Vietnamese, and other Asian. In addition, this variable also included a level for Hispanic. If the interviewer entered multiple responses for the race questions, the individual responses were not recorded. Provided the respondent was not Hispanic, these respondents were classified into one of three levels: “Native Hawaiian and other Pacific Islander,” “Asian multiple race,” and “more than one race.” A non-Hispanic respondent belonged to the “Asian multiple race” category if the multiple responses entered all mapped to the Asian category in IRRACE. The other-specify responses to the race and ethnicity questions were incorporated into these 16 levels.

Although the levels of NEWRACE1 did not change in the years 1999, 2000, and 2001, the method of its creation did change. In 2000 and 2001, an intermediate race variable was created and included on the analytic data file, IRNWRACE, which did not contain a level for Hispanic, so that Hispanics had values distributed across the 15 levels of the variable. Therefore, the race of a Hispanic respondent could be recorded. (For example, by looking at IRNWRACE and the Hispanic indicator, it would be possible to determine whether a Hispanic respondent was Japanese.)

The following sections of this document provide details on the creation of IRRACE (PAPI and CAI), IRHOIND (PAPI and CAI), IRNWRACE, NEWRACE1, and the recoded variables created from these variables. They also describe the handling of other-specify responses using the “dictionaries” for race and Hispanicity, and how the race and Hispanicity variables were utilized in the tables.

4.0 IRRACE

IRRACE has always had four levels: American Indian or Alaska Native, Asian or Pacific Islander, black, and white. In most cases, the race can be obtained directly from the race question (D-5 in PAPI and QD05 in the CAI questionnaire). As stated above, it was necessary in the CAI version to collapse the Asian subcategories into a single Asian category. Along with the creation of IRRACE, a concomitant indicator variable IIRACE was created, which indicated the source of information for IRRACE. The levels of IIRACE changed from PAPI to CAI. Details of the creation of IRRACE and IIRACE across the years are given below.

4.1 IRRACE and IRACE: PAPI

4.1.1 Creation of Preliminary Variables

In the processing of data from the PAPI questionnaire, it was necessary to create a number of preliminary variables before the levels of IRRACE could be defined. None of these preliminary variables was included on the analytic data file. The variable QD05 included responses from the race question D-5. In the rare instance where an interviewer entered a race in “other-specify” for D-4 (the Hispanicity question), and this race had higher priority than the race entered in D-5, then the race in QD05 was the race entered in D-4. In PAPI, interviewers also had the opportunity to record the race of the respondent through visual observation. This information, recorded in the intermediate variable QRACE, was only used in IRRACE if QD05 was missing and the race observed was black or white. The race of the respondent was also recorded on the screener, although the choices on the screener were limited to white, black, and other. The variable XINTRAC was created which incorporated information from QD05, QRACE, and the screener. In XINTRAC, the levels for American Indian/Alaska Native and Asian or Pacific Islander were collapsed into a single category called “other.” XINTRAC was formed in the following manner:

XINTRAC

- = QD05, provided QD05 nonmissing, else
- = QRACE, provided QRACE nonmissing, else
- = Response from screener, provided race for respondent on screener was nonmissing, else
- = Response from screener for majority of householders, provided a majority exists, else
- = Race of head of household, from screener, else
- = Blank

The levels of the variable INTRAC were identical to XINTRAC, except that missing values were imputed using an unweighted sequential hot deck. The details of this procedure are provided in the 2000 Methodological Resource Book, Section 4, Appendix A (RTI, 2002). For this procedure, the file was sorted by primary sampling unit (PSU) number, household type, segment identification number, household identification number, and roster number.

4.1.2 Creation of Base Variable for Imputation

The base variable used in the creation of IRRACE was called TEMPRACE and was created using the following rules:

TEMPRACE

- = 1 (American Indian/Alaska Native) if the questionnaire race variable(QD05) was 3 (American Indian/Alaska Native)
- = 2 (Asian or Pacific Islander) if the questionnaire race variable (QD05) was 4 (Asian or Pacific Islander)
- = 3 (Black/African American) if the questionnaire race variable (QD05) was 2 (black) or the questionnaire race variable was missing and INTRAC = 2 (black). (If INTRAC=2 and QD05 was missing, either the interviewer observed that the respondent was black, or no interviewer observation was available but the screener indicated that the respondent was black.)
- = 4 (White) if the questionnaire race variable (QD05) was 1 (white) or the questionnaire race variable was missing and INTRAC = 1 (white). (If INTRAC=1 and QD05 was missing, either the interviewer observed that the respondent was white, or no interviewer observation was available but the screener indicated that the respondent was white.)

The first two levels of TEMPRACE do not include information from INTRAC; therefore, they do not include information from interviewer observation or from the screener. If INTRAC = 3, and QD05 was missing, then TEMPRACE was blank.

4.1.3 Creation of IRRACE and IIRACE

As with XINTRAC and INTRAC, the levels of the variable IRRACE were identical to TEMPRACE, except that missing values were imputed using an unweighted sequential hot deck (RTI, 2002). Therefore, the imputation of missing values in the race variables was accomplished in two stages. First, the imputation-revised variable INTRAC was created, with the levels white, black, and other. The four-level variable IRRACE was then created using the levels of INTRAC as imputation classes. Hence, at the second stage of imputation, only the levels American Indian/Alaska Native and Asian or Pacific Islander were imputed. The file was sorted within imputation classes by stratum, segment identification number, IRHOIND (the imputation-revised Hispanic indicator), household identification number descending, and a uniform random number.

To differentiate the final imputed values from nonmissing values, a concomitant indicator variable, IIRACE, gives the source of the information for IRRACE. The levels of IIRACE follow:

IIRACE

- = 1 race from questionnaire (either from D-4 or D-5, almost always D-5)
- = 2 race from interviewer observation
- = 3 race from screener roster
- = 4 race statistically imputed (screener roster is “other”—imputed to American Indian/Alaska Native or Asian/Pacific Islander)
- = 5 race statistically imputed (no information from screener roster)

In most cases (93.5 percent of the time in 1998), the race was obtained from the questionnaire. No information is available that states the number of times race was obtained from D-4 rather than D-5, nor is information available on the number of times race was obtained from the other-specify responses. For most of the remainder (6.3 percent in 1998), the race was obtained from interviewer observation. Only 0.2 percent of the race responses came from the screener roster in 1998, and only three cases were imputed (levels 4 and 5).

4.2 IRRACE and IIRACE: CAI, 1999

4.2.1 Creation of Base Variable for Imputation

As stated previously, there are major differences between the single PAPI question about race and the two questions in the CAI questionnaire concerning race. In short, two questions were asked about race in the CAI questionnaire, with 12 levels rather than just 4, and interviewers could mark multiple categories in QD05. To obtain the Asian/Pacific Islander level of IRRACE, several levels of the original race variables had to be combined. In addition, the other-specify responses were given greater scrutiny, with more informative codes that could be used to avoid unnecessary imputation. A simplifying aspect of the 1999 race processing, however, is that only a single stage of imputation was required to create IRRACE, with the only preliminary variable being the base variable, EDRACE. Because no interviewer-observed race was available in the CAI questionnaire, and screener information was not used, the base variable was created based entirely on the race questions in the questionnaire. EDRACE was created using the following rules under three possible scenarios:

EDRACE

Scenario 1: Given only one category identified in QD05:

- = The single race identified in QD05, provided that single race was not the “other” level with either multiple responses or a country of origin in the other-specify blank, else

- = Race assigned from the multiple responses in the other-specify blank of QD05, provided specific races were mentioned,¹ based on the same priority rule used in PAPI: black/African American, Asian/Pacific Islander, American Indian/Alaska Native, white, else
- = Race assigned based on the census of a multiracial country of origin as stated in the other-specify blank of QD05, provided the country of origin was not Hispanic, where a random number is used to allocate the race (provided the allocation was not to an “other” category), else
- = A code indicating that race is unknown, but the respondent belongs to a Hispanic group, provided the respondent is Hispanic, either as a result of his or her response to QD04 or as a result of his or her response in “other-specify” for QD05, else
- = Missing

Scenario 2: Given two or more categories identified in QD05:

- = The race response in QD06, provided that race was not the “other” level with either multiple responses or a country of origin in “other-specify” for QD05, else
- = Race assigned from the multiple responses in “other-specify” for QD05, provided specific races were mentioned and provided QD06 was “other”, based on the same priority rule used in PAPI: black/African American, Asian/Pacific Islander, American Indian/Alaska Native, white, else
- = Race assigned based on the census of a multiracial country of origin as stated in “other-specify” for QD05, provided QD06 was “other” and the country of origin was not Hispanic, where a random number is used to allocate the respondent to a race, (provided the allocation was not to an “other” category), else
- = Race assigned from the multiple responses in QD05, provided QD06 was not answered, based on the same priority rule used in PAPI: black/African American, Asian/Pacific Islander, American Indian/Alaska Native, white, else
- = A code indicating that race is unknown, but the respondent belongs to a Hispanic group, provided the respondent is Hispanic, either as a result of his or her response to QD04 or as a result of his or her response in “other-specify” for QD05, else
- = Missing

¹ If the response was “mixed,” “biracial,” or something of that nature, the final race had to be imputed. However, whites were excluded as donors, due to the priority rule indicating that persons of mixed race would be considered something other than white.

Scenario 3: Given QD05 left blank:

- = The race response in QD04 (the Hispanic group question), provided a race was given in QD04, else
- = A code indicating that race is unknown, but the respondent belongs to a Hispanic group, provided the respondent is Hispanic as a result of his or her response to QD04, else
- = Missing

It should be noted that in some cases a country of origin was given, where nearly 100 percent of the population was of one race. Under these circumstances (e.g., Nigeria), a code indicating that race was given for the respondent, and no census data were required. Otherwise, census data were obtained either from the latest *World Almanac* or from the Internet. In some multiethnic countries, a percentage of the population was described as “other.” In these cases, imputation would be required to determine the correct race (details are given in Section 6).

4.2.2 Creation of IRRACE and IIRACE

The levels of the variable IRRACE were identical to EDRACE, except that missing values were imputed using an unweighted sequential hot deck. Imputation classes were created based on the Hispanic-origin status and (if applicable) Hispanic-origin group of respondents with missing values of EDRACE. The imputation classes were Mexican, Cuban, Puerto Rican, Central or South American, and Hispanic (specific group unknown). For Hispanic respondents with missing EDRACE values, donors were restricted to respondents in the same Hispanic-origin group. For non-Hispanic respondents, or respondents whose Hispanic-origin status was unknown (because missing values of the Hispanic-origin indicator and group were imputed after the imputation of the race variable), all respondents with valid EDRACE values were eligible to be donors. However, if the ambiguous multiple race category was given in the other-specify response to QD05, donors could not be white due to the priority rule.

The file was serpentine sorted within imputation classes by design stratum, U.S. Census region, segment, household type, and a random number prior to imputation. In a serpentine sort, whether the direction is ascending or descending, the direction changes each time there is a change in the value of a sorting variable. For example, if two variables are to be serpentine sorted, when the value of the first variable changes in the sort, the direction of the sort for the second variable changes. This ensures that variables that are lower in the sort order do not change by a large amount when the value of a variable higher in the sort order changes.

To differentiate the final imputed values from nonmissing values, a concomitant indicator variable, IIRACE, gives the source of the information for IRRACE. The levels in this CAI version of IIRACE differ from those given by the PAPI version. The levels of the CAI version of IIRACE follow:

IIRACE

- = 1 Race from questionnaire (from QD05 or QD06)
- = 2 Race assigned from other-specify entry (from QD05)
- = 3 Race statistically imputed (no prior information)
- = 4 Race statistically imputed (restricted to Hispanic groups)
- = 5 Race randomly allocated using census data from country of origin

In most cases (96.2 percent of the time), the race was obtained from the questionnaire. For most of the remaining cases (2.9 percent), the race was imputed within Hispanic groups. Only 0.5 percent of the race responses came from the other-specify responses directly, 0.2 percent were randomly allocated using the census data from the country of origin, and 0.19 percent were statistically imputed with no prior information.

4.3 IRRACE: CAI, 2000 and 2001

4.3.1 Creation of Base Variable for Imputation

Unlike the 1999 NHSDA, in 2000 the race questions addressed the Asian races in a separate question. In 2000, as well as 2001, interviewers asked respondents if they described themselves according to one or more of the following categories: white, black/African American, American Indian/Alaska Native, Native Hawaiian, other Pacific Islander, or Asian. Respondents who reported “Asian” were then asked, in QD05ASIA, to describe themselves according to the following Asian categories: Asian Indian, Chinese, Filipino, Japanese, Korean, and Vietnamese. Interviewers could enter an other-specify response in either QD05 or in QD05ASIA.

For ease of processing, the levels of the single QD05 variable in 1999 were reproduced in 2000 and 2001 using the levels of QD05 and QD05ASIA. The base variable was then created in the same way as 1999. A special problem arose in 2000 that did not arise in 1999. When a respondent’s other-specify indicated “Indian” in 1999, this respondent was always considered American Indian/Alaska Native. In the processing of 2000 and 2001 data, this was only true if “Indian” was entered in the other-specify response for QD05. If “Indian” was the other-specify response for QD05ASIA, the respondent was considered an Asian Indian.

4.3.2 Creation of IRRACE, IIRACE, and IIR2RACE

As in 1999, the levels of IRRACE were identical to those of EDTRACE, except that missing values were imputed. However, in 2000, the unweighted sequential hot deck was dropped in favor of a new methodology developed specifically for the NHSDA, predictive mean neighborhoods (PMN), which is described in Section 4, Appendix C, of the 2000 Methodological Resource Book (RTI, 2002). PMN is a combination of a model-assisted procedure and nearest neighbor hot deck. As an additional change in 2000, missing values in the preliminary variable used in the creation of NEWTRACE1, IRNWRACE, were imputed at the same time as those in IRRACE.

In the modeling step, weights were initially adjusted for item nonresponse, using the item response propensity modeling features of the generalized exponential model (GEM). More details on GEM can be found in the 2000 Methodological Resource Book, Section 4, Appendix B (RTI, 2002). The covariates used in the item response propensity model were Census region, race of head of household (from screener), age of respondent, percent of segment population Hispanic, percent of segment population non-Hispanic black, and percent of owner-occupied households. Using the adjusted weights, all respondents with complete data were used to fit a multinomial logistic model, with separate models for the age groups 12 to 17, 18 to 25, and 26 or older. The separate age groups were used more for ease of processing rather than due to any strong correlation between age and race.

The model was fitted using PROC MULTLOG in SUDAAN (RTI, 2001), using the first three categories of IRRACE as response cells (American Indian/Alaska Native, Asian/Pacific Islander, black).² Covariates in the models included Census region, race of head of household (from screener), age of respondent, percent of segment population Hispanic, percent of segment population non-Hispanic black, and percent of owner-occupied households. Naturally, most of the strength of the race models comes from the screener variable, race of head of household.³

In the hot-deck step, a neighborhood of PMN donors is selected so that the three predicted probabilities of each candidate donor is “close” (in terms of Mahalanobis distance) to the three predicted probabilities of the recipient. PMN donors are limited according to constraints, which either cannot be loosened (logical constraints) or can be loosened if insufficient donors are available (likeness constraints). Some of the constraints given below are only relevant for the imputation of IRNWRACE. The logical constraints are as follows:

- If the recipient was of Hispanic-origin, the donor must also be of Hispanic origin.

² In a multinomial distribution with m cells, knowledge of the first $m-1$ cells will naturally determine the value for the m th cell. Therefore, the predicted probability of the m th cell is determined by taking 1-(the sum of the predicted probabilities for the first $m-1$ cells).

³ As a matter of policy, it had been decided not to use screener information if at all possible in the imputation of variables in the CAI questionnaire. However, excluding screener information completely meant that no available variable could provide the amount of information required for an adequate model.

- If the recipient was a member of a particular Hispanic origin group (e.g., Mexican, Puerto Rican, Central or South American, Cuban), the donor also must be a member of that group. If the recipient was a member of more than one Hispanic origin group, the donor must be a member of at least one of those specified by the respondent.
- If the recipient was known to be Asian, the donor must also have been Asian (pertains to IRNWRACE only).
- If the recipient was known to be of more than one race, but the specific races were unknown, the donor must not have been white, according to the donor's value for IRRACE. (This is because of the "priority rule": "white" has the lowest priority, so multiple-race respondents cannot be white in EDRACE.) The recipient's value for IRNWRACE is "more than one race," so no imputation of IRNWRACE would be necessary.

The likeness constraints are as follows:

- The donor must live in the same segment as the recipient. (This constraint is removed first.)
- Each of the donor's three predictive means must be within 5 percent of each of the recipient's three predictive means.

After a donor is selected using the PMN method, the donor supplies values to the recipient for IRRACE (if missing) and IRNWRACE (if missing). The levels of the concomitant variable IIRACE did not change from 1999. However, to increase the amount of information available concerning the source of data used for IRRACE, a new variable was created with more detailed categories, II2RACE. The levels of II2RACE follow:

II2RACE

- = 1 Race from single QD05 response
- = 2 More than one race given in QD05, race from single QD06 response
- = 3 Race assigned from other-specify entry (from QD05)
- = 4 Race randomly allocated using census data from country of origin
- = 5 More than one race given in QD05, no single race given in QD06, race determined from priority rule
- = 6 Race statistically imputed (no prior information)
- = 7 Race statistically imputed (restricted to Hispanic groups)

The levels of II2RACE map to those of IIRACE in the following manner:

IIRACE = 1 maps to II2RACE = 1, 2, 5

IIRACE = 2 maps to II2RACE = 3

IIRACE = 3 maps to II2RACE = 4

IIRACE = 4 maps to II2RACE = 6

IIRACE = 5 maps to II2RACE = 7

Hence, II2RACE provides more detailed information about how the race was determined if it was obtained from the questionnaire questions. In 2000 and 2001, 95.9 percent of the respondents had a value of IIRACE = 1. In both years, this percentage can be divided in the following manner (the percentages were the same in both years):

II2RACE = 1: 94.2 percent

II2RACE = 2: 1.6 percent

II2RACE = 5: 0.1 percent

4.4 Summary of Changes in IRRACE: 1998-2001

In summary, major changes in the creation of IRRACE were implemented between 1998 and 1999 in accordance with the changes in the race question(s) introduced with the new CAI questionnaire. However, other changes were also implemented that were unrelated to changes in the race question. These changes consist of the following items: (1) interviewer observation of race was no longer available, increasing the potential need for imputation; (2) race from the screener was no longer used directly, also increasing the potential need for imputation; (3) new imputation classes were implemented that incorporated information about the respondent's Hispanicity; (4) greater scrutiny was applied to the other-specify responses, with more informative codes applied to avoid unnecessary imputation; and (5) the levels of the imputation indicator variable IIRACE changed completely to reflect changes in the sources of information used to create IRRACE.

Between 1999 and 2000, further minor adjustments were required to account for the minor changes implemented in the questionnaire. However, other changes were also implemented that were unrelated to changes in the race questions. These changes consist of the following items: (1) the method of imputation was changed from unweighted sequential hot deck to PMN; (2) IRRACE and the preliminary variable for NEWRACE1, IRNWRACE, were imputed together; and (3) a new variable was created, II2RACE, giving more detailed information about the source of data for IRRACE.

5.0 IRHOIND

The final indicator of a respondent's ethnicity is called IRHOIND. This has always been a binary variable indicating whether a respondent considered himself or herself Hispanic or Latino (where a "1" indicates Hispanic/Latino, and "2" indicates otherwise). For most respondents, ethnicity was obtained directly from the questionnaire, either from D-3 or D-4 in PAPI or QD03 in CAI. No imputation was required in the PAPI version of IRHOIND because information was obtained from the screener where necessary and Hispanicity was known for all screened respondents. In CAI, however, imputation was required so that an accompanying imputation indicator, IIHOIND, was also created.

5.1 IRHOIND, PAPI

The Hispanic indicator variable was created as follows:

- = 1 (Hispanic) if the Hispanic indicator (D-3) was “Yes,” else
- = 1 (Hispanic) if the Hispanic group question (D-4) was nonmissing (indicating a valid Hispanic group) or Hispanicity was indicated by the other-specify response to the race variable, D-5, provided the “Hispanic” response in D-4 or D-5 did not disagree with all of following: the screener, D-3, the ethnicity of the head of household, and the ethnicity of the majority of the people in the household, else
- = 2 (non-Hispanic) if the Hispanic indicator (D-3) was “No,” else
- = Hispanic indicator from the screening questionnaire (XHISP), if it was nonmissing, else
- = Hispanicity of the majority of the people in the dwelling unit, provided a majority existed, else
- = Hispanicity of the head of household (FHHT), which is never missing.

Hence, IRHOIND was not imputed in its PAPI version.

5.2 IRHOIND and IHOIND, CAI

5.2.1 Creation of Base Variable for Imputation

The base variable used to create IRHOIND was called EDHOIND. It was not included on the analytic data file and was created using the following rules:

EDHOIND

- = 1 (Hispanic), if QD03 = 1 OR QD03=2 but the other-specify response to QD05 indicates that the respondent is Hispanic, else
- = 2 (Not Hispanic), if QD03 = 2 OR QD03 = 1, but the only response to QD04 (the Hispanic group question) is an other-specify response indicating the respondent is not Hispanic (and there is no indication from QD05 that the respondent is Hispanic), else
- = Missing

5.2.2 Creation of IRHOIND and IHOIND

IRHOIND was created after IRRACE, so that IRRACE could be used as a covariate in the imputation procedure. An unweighted sequential hot deck was used in 1999 to impute missing values in IRHOIND. The file was serpentine sorted by design stratum, Census region, IRRACE, and a random number.

The imputation procedure used in 2000 for IRHOIND was PMN, with weights adjusted using the response propensity modeling features of the GEM (RTI, 2002). As with race, imputation of missing values in EDHOIND was implemented separately within the age groups 12 to 17, 18 to 25, and 26 or older. Separating the imputations into age groups was necessary more for ease of processing rather than a strong correlation between age and Hispanicity. The covariates used in the item response propensity model were Census region, IRRACE, age, age squared, percent Hispanic population, percent non-Hispanic black population, and percent of owner-occupied households. Using the adjusted weights, all respondents with complete data were used to fit a (binomial) logistic model. The models were fitted using PROC LOGISTIC in SAS, using the “yes” response as the response cell. Covariates in the models included were Census region, IRRACE, age, age squared, percent Hispanic population, percent non-Hispanic black population, and percent of owner-occupied households.

IIHOIND follows the same pattern as most of the imputation indicators in other parts of the questionnaire, where

IIHOIND

- = 1 Hispanicity from questionnaire (QD03)
- = 2 Hispanicity logically assigned (from QD05)
- = 3 Hispanicity imputed

6.0 NEWRACE1 and IRNWRACE

NEWRACE1 was created for the first time in 1999 with the introduction of the CAI questionnaire. The amount of detail given in this variable would have been impossible with PAPI. NEWRACE1 has the following 16 levels:

- Non-Hispanic white
- Non-Hispanic black/African American
- Non-Hispanic American Indian/Alaska Native
- Non-Hispanic Native Hawaiian
- Non-Hispanic other Pacific Islander
- Non-Hispanic Chinese
- Non-Hispanic Filipino
- Non-Hispanic Japanese
- Non-Hispanic Asian Indian
- Non-Hispanic Korean
- Non-Hispanic Vietnamese
- Non-Hispanic other Asian

- Non-Hispanic Asian multiple categories
- Non-Hispanic multiple race
- Hispanic

Respondents in the non-Hispanic Asian multiple categories level gave two or more responses to the following Asian categories: non-Hispanic Native Hawaiian, non-Hispanic other Pacific Islander, non-Hispanic Chinese, non-Hispanic Filipino, non-Hispanic Japanese, non-Hispanic Asian Indian, non-Hispanic Korean, non-Hispanic Vietnamese, and non-Hispanic other Asian. Respondents with the non-Hispanic multiple race level gave at least one response that was not an Asian category. Any indication of Hispanicity meant that the respondent was coded as Hispanic, even if the respondent indicated a mixture of race and ethnicity in the other-specify response for QD05 (e.g., Hispanic and white). Because Hispanic respondents had their own level, the race of Hispanics is not recorded with this variable. Also, because Hispanicity is included as a level, NEWRACE1 cannot map directly to the levels in IRRACE. In 2000, an intermediate variable, IRNWRACE was created and included on the analytic data file, which contained 15 levels: the 12 levels from the race questions in CAI, plus “Native Hawaiian and other Pacific Islander,” “Asian multiple category,” and “more than one race.” Hence, the racial categories of Hispanics were preserved with this variable. The creation of this variable, and its inclusion on the analytic data file, is in congruence with OMB Directive No 15, which stipulates that Hispanicity should not be considered a racial category.

6.1 NEWRACE1 and IINWRAC1, 1999

6.1.1 Background for Creating NEWRACE1

In 1999, after processing had been completed for all demographic variables, including IRRACE, it became clear that a variable was needed that (1) accounted for the detailed levels given in the CAI questionnaire, and (2) accounted for a respondent’s ethnicity (i.e., his or her Hispanicity). This new variable, NEWRACE1, was defined with the 16 levels given in the introduction to this section. The level corresponding to Hispanic could be determined by the value of IRHOIND, which was already available. If the response to QD05 was nonmissing, and the respondent was not Hispanic, the value of NEWRACE1 could be determined using the levels of QD05. Otherwise, in order to maintain consistency with IRRACE, the same donors used for imputing IRRACE were used to determine the value of NEWRACE1. In particular, respondents with an imputed value of white, black/African American, or American Indian/Alaska Native were given the same value for NEWRACE1, provided they were not Hispanic and did not indicate more than one race. If a respondent was imputed to have a value of IRRACE indicating “Asian,” and the respondent was not Hispanic, then the donor for the IRRACE value was reselected for NEWRACE1 to determine the specific Asian category to which the respondent belonged. The only situation where a special imputation was required occurred if the interviewer indicated that the respondent was “Asian” in the “other-specify” response for QD05, but further information was unavailable. In this instance, the specific Asian category needed to be imputed.

6.1.2 Creation of NEWRACE1

The levels of NEWRACE1 are defined as follows:

NEWRACE1

- = QD05 (non-Hispanic race as indicated by a single response to QD05), if IRHOIND = 2, and either a single race was selected in QD05 besides "other," or "other," was the only race selected in QD05 and the other-specify response was interpreted as a single race corresponding to one of the categories in QD05 besides "other," else
- = Race based on other-specify response to QD04, if IRHOIND = 2 (and QD03 = 1), provided no valid response was given to QD05, and the other-specify response to QD04 mapped to a single race corresponding to one of the categories in QD05,⁴ else
- = Race assigned based on the census of a multiracial country of origin as stated in the other-specify blank of QD05, provided "other" was the only race selected in QD05 and the country of origin was not Hispanic, where a random number is used to allocate a race (provided the allocation was not to an "other" category), else
- = 6 (non-Hispanic Native Hawaiian/other Pacific Islander combination), if IRHOIND=2, both Native Hawaiian and other Pacific Islander levels were selected in QD05, and no other levels of QD05 were selected, else
- = 14 (non-Hispanic Asian multiple category), if IRHOIND = 2, and either more than one race was selected in QD05 where all those selected are considered "Asian," or "other" was the only race selected in QD05 and the other-specify response was interpreted as a combination of several Asian categories, else
- = 15 (non-Hispanic multiple race), if two or more races were selected in QD05 and (a) at least one was non-Asian, (b) at least one was something other than Native Hawaiian or other Pacific Islander, and (c) IRHOIND = 2, else
- = 16 (Hispanic) if IRHOIND = 1, else
- = Missing

⁴ This would occur if the respondent indicated "Hispanic" in QD03, then indicated that his or her Hispanic group was something non-Hispanic in the QD04 other-specify response, such as "Haitian," then did not respond to QD05.

As noted above, those respondents who indicated “Asian” in an other-specify response for race, but not one of the specific Asian groups included in NEWRACE1, were assigned a value of NEWRACE1 by statistically imputing a finer Asian category (12 respondents). This included respondents who indicated a country of origin and were also randomly allocated to “Asian.” (These respondents would be included under “Missing” above.) The imputation was performed using an unweighted sequential hot-deck procedure, sorting the file by design stratum, Census region, segment identification number, race of head of household (from the screener), and a random number. Donors were restricted to non-Hispanic Asian respondents with valid NEWRACE1 values.

To track the source of data used to create NEWRACE1, the variable IINWRAC1 was created. This imputation indicator had the following levels:

IINWRAC1

- = 1 From QD05 or IRHOIND
- = 2 Logically assigned from other-specify
- = 3 From IRRACE donor
- = 4 Statistically imputed

Note that if the respondent was imputed to be Hispanic, the level for this variable would still be 1. The imputation of Hispanicity was only recorded in IIHOIND. It was possible for a non-Hispanic respondent to have an imputed IRRACE and nonimputed NEWRACE1 if the interviewer marked “mixed,” “biracial,” or something of that nature in the other-specify response of QD05, but did not indicate any specific races. Moreover, it was also possible for a non-Hispanic respondent to have an imputed NEWRACE1 and nonimputed IRRACE if the interviewer marked “Asian” in the other-specify response of QD05, but did not specify a specific Asian group. For these reasons, it would not be possible to map the levels of IINWRAC1 to IIRACE.

6.2 NEWRACE1, IRNWRACE, and IINWRACE, 2000 and 2001

6.2.1 Creation of Base Variable for Imputation of IRNWRACE

In 2000, it was known at the beginning of processing that NEWRACE1 would need to be created. However, reproducing the method of creating NEWRACE1 that was used in 1999 (first imputing IRRACE, then backtracking using donors from IRRACE and doing a subsequent imputation where needed) seemed awkward and unnecessary. As an alternative, an intermediate variable was created that recorded all the levels of NEWRACE1 except Hispanicity. The imputation-revised version of this variable, IRNWRACE, was included on the analytic data file and combined with IRHOIND to create NEWRACE1 as a recoded variable.

No model was fitted that incorporated all 12 categories; only 4 categories were used in the model for race: white, black/African American, American Indian/Alaska Native, and Asian/Pacific Islander. However, in the hot-deck stage of PMN, donors for IRRACE were used to determine the values of IRNWRACE using the predicted mean vector from this model. The base variable for IRNWRACE, NEWRACE, was created in the following manner:

NEWRACE

- = QD05 (race indicated by a single response to QD05), if either a single race was selected in QD05 besides “other,” or “other” was the only race selected in QD05 and the other-specify response was interpreted as a single race corresponding to one of the categories in QD05 besides “other,” else
- = Race based on other-specify response to QD04, provided no valid response was given to QD05, and the other-specify response to QD04 mapped to a single race corresponding to one of the categories in QD05, else
- = Race assigned based on the census of a multiracial country of origin as stated in “other-specify” for QD05, provided “other” was the only race selected in QD05 and the country of origin was not Hispanic, where a random number is used to allocate a race (provided the allocation was not to an “other” category), else
- = 6 (non-Hispanic Native Hawaiian/other Pacific Islander combination), if both Native Hawaiian and other Pacific Islander levels were selected in QD05, and no other levels of QD05 were selected, else
- = 14 (Asian multiple category), if either more than one race was selected in QD05 where all those selected are considered “Asian,” or “other” was the only race selected in QD05 and the other-specify response was interpreted as a combination of several Asian categories, else
- = 15 (more than one race), if two or more races were selected in QD05 and (a) at least one was non-Asian and (b) at least one was something other than Native Hawaiian or other Pacific Islander, else
- = Missing

As with NEWRACE1, those respondents who indicated "Asian" in an other-specify response for race, but not one of the specific Asian groups, were assigned a code indicating that a finer Asian category needed to be imputed. This included respondents who indicated a country of origin and were randomly allocated to “Asian.” (These respondents would be included under “Missing” above.)

6.2.2 Creation of IRNWRACE, IINWRACE, and NEWRACE1

The levels of the newly created variable IRNWRACE were identical to those of NEWRACE, except that missing values were imputed. The imputation was performed simultaneously with that of IRRACE using PMN. Hence, if “Asian” was known, but finer categories needed to be imputed, the assignment of categories was based on a model for the broad categories given in IRRACE. (In other words, whether an Asian was, say, Chinese, depended upon the predicted means for white, black/African American, American Indian/Alaska Native, and Asia/Pacific Islander. It should be noted that this issue only affects a small handful of respondents.) Details of the PMN procedure as applied to the imputation of IRRACE and IRNWRACE are given in Section 4.3.2. NEWRACE1 was created as a recoded variable after the completion of processing of all demographic variables, incorporating the levels of IRHOIND and IRNWRACE. The levels of NEWRACE1 in 2000 were identical to those given in 1999.

As with IRRACE and the 1999 version of NEWRACE1, an imputation indicator, called IINWRACE, was created to track the source of data used for IRNWRACE. Because NEWRACE1 no longer had directly imputed values, but was now a recoded variable, the variable IINWRAC1 was dropped in 2000. The levels of IINWRACE are given below:

IINWRACE

- = 1 Race from single QD05 response, or more than one race as entered in QD05
- = 2 Race assigned from other-specify entry (from QD05), including possibly more than one race
- = 3 Race randomly allocated using census data from country of origin
- = 4 Statistical imputation of “Asian” into finer categories
- = 5 Race statistically imputed (no prior information)
- = 6 Race statistically imputed (restricted to Hispanic groups)

The presence of a separate category for more than one race complicates mapping directly to IIRACE or II2RACE. However, the following can be said:

IINWRACE = 1 maps to IIRACE = 1 and II2RACE = 1, 2, 5

IINWRACE = 2 almost maps to IIRACE = 2 and II2RACE = 3 (excludes people who indicated “Asian” in the other-specify for race, who were mapped to IINWRACE = 4)

IINWRACE = 3 almost maps to IIRACE = 5 and II2RACE = 4 (excludes people who were randomly allocated to “Asian” who were mapped to IINWRACE = 4)

IINWRACE = 4 takes people from IIRACE = 2, 3, 5 and II2RACE = 3, 4, 6

IINWRACE = 5 almost maps to IIRACE = 3 and II2RACE = 6 (excludes people who indicated “multiracial” or “biracial” in their other-specify who were mapped to IINWRACE = 4 (incorrectly).

IINWRACE = 6 maps to IIRACE = 4 and II2RACE = 7

7.0 Other-Specify Codes for Race and Hispanicity

It is not uncommon for a respondent to feel that the categories for race or Hispanicity given in the questionnaire do not apply to him or her. In these situations, interviewers are given the opportunity to manually enter (type) a category that the respondent feels best describes himself or herself. The manually entered responses are called “other-specify” or “alpha-specify” responses because they are typed in a part of the question that asks the interviewer to specify an “alphabetic” response. These other-specify responses are then matched to a code to describe the response, which are collected and maintained in a file known as a “dictionary.” Other-specify responses from each survey year are matched against this file, and any responses without a code are given a new code and added to the dictionary; therefore, the size of the dictionary file increases each year. (In most cases, new unmatched responses are just new misspellings of an already established category, such as a response of “Porto Rican” instead of “Puerto Rican.”)

In both PAPI and 1999 CAI, there were other-specify categories in both the Hispanic-origin group question (D-4 in PAPI, QD04 in CAI) and the race question (D-5 in PAPI, QD05 in CAI). All respondents were asked the race question, but the Hispanic-origin group question was only asked if the respondent indicated he or she was Hispanic in the question about Hispanicity (D-3 in PAPI, QD03 in CAI). Because race and Hispanicity are so closely related, interviewers occasionally entered a race in the other-specify category for Hispanicity and commonly entered a Hispanic group in the other-specify category for race. In fact, the vast majority of other-specify responses to the race question either indicated that the respondent was Hispanic, or that the respondent belonged to some Hispanic group. For these reasons, the codes for race and Hispanicity were placed in the same dictionary, so that an other-specify response to the Hispanic groups question could be cross-referenced against the codes for both race and Hispanicity. Likewise, the other-specify response to the race question could also be cross-referenced against the codes for both race and Hispanicity.

In CAI for 2000 and 2001, the race question was separated into two parts, where respondents who reported that they were Asian were given a separate opportunity to have the interviewers enter a category for them. These other-specify responses were considered simultaneously with the other-specify responses to QD05, the main race question. The only instance where separate codes were required was when the interviewer marked the Asian category, then manually entered “Indian” in the Asian other-specify category. Normally, “Indian” maps to a code for American Indian, but in this case the respondent was considered Asian Indian.

The other-specify responses were examined when (a) “other” was selected as a race in the race question(s), or (b) no race was given in response to the race question(s), but a race category was given as an other-specify response to the Hispanic-origin group question. In such cases, if a valid other-specify response was given, the code corresponding to that response was used in order to assign values to EDRACE and the base variables for imputing NEWRACE1 (in 1999) or IRNWRACE (in 2000 and 2001). Because identification of Hispanic groups is not the main focus of this document, this section concentrates on the dictionary codes for race.

7.1 Approach to Dictionary in PAPI

In the PAPI race-Hispanicity dictionary, five codes were available for race, which corresponded to the four categories in the race question, plus a code for bad data. Unless it was obvious that the race given in the other-specify response mapped to one of the four categories, the code was marked “bad data” and the level of IRRACE determined from some other source. Because most responses indicated “Hispanic” or some Hispanic group, most other-specify responses were usually ignored. Yet this was not a serious problem in PAPI in that the interviewer observation and screener race entries filled in most of the gaps. However, this strategy could not be maintained for CAI.

7.2 Approach to Dictionary in CAI

To reduce the amount of statistical imputation necessary to create the imputation-revised race variables, the race and Hispanic-origin group dictionary used in PAPI was expanded for CAI. Likewise, the procedures used to assign the large number of other-specify responses to the categories for these race variables were more comprehensive for CAI. In many cases, interviewers entered an other-specify response that could be mapped to 1 of the 12 categories in the race question. The responses used to do this are given below. In the lists of countries and ethnic groups corresponding to each code, it should be noted that these lists are not exhaustive. The countries and ethnic groups in the lists below appeared in the other-specify responses of 1999, 2000, and/or 2001. The categorization of countries and ethnic groups into racial categories follows, as closely as possible, those of the U.S. Census.

1. The following other-specify responses and their derivatives were classified as "black/African American": Afro American, brown, Haitian, Caribbean Creole, African or any country from sub-Saharan Africa except Namibia or South Africa (see #6), morena or moreno, negra or negro, triguena or trigueno, tan, St. Vincent. Also, any respondent who indicated that he or she was part Hispanic and part black was considered “black.”
2. The following responses and their derivatives would formerly have been considered within the “Asian/Pacific Islander” group, but were given separate codes in accordance with the race question in CAI: Native Hawaiian, Other Pacific Islander (which also included Micronesian, Polynesian, Samoan, Saipan, and Guamanian), Chinese (which also included Taiwanese, Cantonese, Guanma), Filipino, Japanese, Asian Indian (which also included Nepalese, Pakistani, Bengali [Bangladesh],

Hindu, Indian American, African Indian, Kashmirian, Punjabi, Sri Lankan, Sikh), Korean, Vietnamese, other Asian. The other Asian group included the following responses and their derivatives: Lao, Thai, Cambodian, Kampuchean, Malaysian, Burmese, Myanmar, Okinawan, Chaldean, East Indian, Indonesian, Eurasian, Iranian, Persian, Kurd, Afghan, Hmong, Kazakh, Mienh, Singaporean, Mongolian, Tibetan, Uzbek, Turkmenistan. A separate code was also given to cases indicating “Asian” with no specific group. Also, any respondent who indicated that he or she was part Hispanic and part Asian was classified as “Asian.”

3. The following other-specify responses and their derivatives were classified as "American Indian": American Indian or Alaska Native: Native American, Indian (except respondents who also indicated they were Asian), Indigenous, Mayan, Aztec, mestizo or mestiza, Yaqui, Zapotec, Apache, Blackfoot, Cherokee, Navajo, Tewa, Weott, Aleut, and Eskimo. Also, any respondent indicating that he or she was part Hispanic and part American Indian was classified as “American Indian.”
4. The following other-specify responses and their derivatives were classified as “white”: Caucasian, North African or any country from North Africa, Arabic, Turkish, Armenian, Jewish, Middle Eastern/Israeli, Canadian, Assyrian, any country from central, eastern, or southeastern Europe except Germany (see #6), blanco, Celtic, Anglo-Saxon, Armenian, Cajun, Caledonian, any combination of European nationalities, or part-Hispanic and part-white. (A separate code was available for Middle Eastern countries, but they were all finally classified as white. The same is true for Canada, although this country should also be treated in the same way as those in #6.)
5. If a respondent indicated a Hispanic-origin group in response to the race other-specify question, he or she was assigned to groups for restricted imputation of race. That is, race was statistically imputed for such respondents, using as donors only those respondents of the same Hispanic-origin group who gave a valid race response. The groups for restricted imputation were Hispanic nonspecific, Mexicans, Puerto Ricans, Cubans, Central or South Americans, Mexicans and Puerto Ricans combined, Mexicans and Central or South Americans combined, Mexicans and Cubans combined, Puerto Ricans and Central or South Americans combined, Puerto Ricans and Cubans combined, and Cubans and Central or South Americans combined.

6. For certain countries of origin given in the other-specify responses, race was randomly assigned using Census data for those countries. In many cases, a small percentage of respondents from a given country were left to be statistically imputed. The following is a list of the countries treated in this way and the percentages assigned to each race:⁵
- Dominican Republic: 84 percent black, 16 percent white, 0 percent statistically imputed
 - Caribbean and West Indies: 80 percent black, 14 percent Asian, 6 percent statistically imputed
 - Belize: 55 percent American Indian, 37 percent black, 8 percent statistically imputed
 - Guyana: 51 percent Asian, 43 percent black, 6 percent statistically imputed
 - Suriname: 52 percent Asian, 31 percent black, 17 percent statistically imputed
 - Trinidad and Tobago: 57 percent black, 40 percent Asian, 3 percent statistically imputed
 - Jamaica: 91 percent black, 9 percent statistically imputed
 - Bahamas and Virgin Islands: 85 percent black, 15 percent white, 0 percent statistically imputed
 - Western Europe, including Spain and Portugal: 95 percent white, 5 percent statistically imputed
 - New Zealand: 88 percent white, 9 percent black, 3 percent statistically imputed
 - South Africa: 84 percent black, 13 percent white, 3 percent Asian, 0 percent statistically imputed
 - Australia: 95 percent white, 4 percent Asian, 1 percent black, 0 percent statistically imputed
 - Barbados: 80 percent black, 16 percent mixed, 4 percent white

⁵ Note that these are the percentages used to randomly assign respondents to races although the distribution of assigned races in the sample does not match these exactly. Also note that if 0 percent are statistically imputed, no respondents are assigned to the races that are not listed.

7. If the respondent indicated a mixture of races in the other-specify responses, the particular mixture was recorded with a separate code. For example, a respondent who answered “black and white” would be given the code 201, while a “Korean and Chinese” respondent would be given the code 310. Respondents with these codes involving at least one non-Asian would be classified into the more than one race category, while respondents with more than one race codes involving all Asians would be classified as “Asian multiple categories.”

8.0 Recoded Variables

Recoded variables are variables that combine levels of variables derived from the questionnaire into new variables that are useful for analysis. New recoded variables have been created that reflect the changes due to the implementation of CAI.

8.1 RACE and HISPRACE

Two recoded variables have been created in both PAPI and in CAI. Both RACE and HISPRACE are recoded variables involving IRRACE and IRHOIND. The levels for each variable are given below:

RACE

- = 1 Non-Hispanic white (IRHOIND = 2, IRRACE = 4)
- = 2 Non-Hispanic black (IRHOIND = 2, IRRACE = 3)
- = 3 Hispanic (IRHOIND = 1)
- = 4 Non-Hispanic other (IRHOIND = 2, IRRACE = 1, 2)

HISPRACE

- = 1 Hispanic (IRHOIND = 1)
- = 2 Non-Hispanic black (IRHOIND = 2, IRRACE = 3)
- = 3 Non-Hispanic non-black (IRHOIND = 2, IRRACE = 1, 2, 4)

The levels of HISPRACE map to those of RACE in the following manner:

HISPRACE = 1 maps to RACE = 3

HISPRACE = 2 maps to RACE = 2

HISPRACE = 3 maps to RACE = 1, 4

8.2 EXPRACE, NEWRACE1, NEWRACE2, and RACE4

The recoded variables that relate specifically to the 12-category questions in CAI are EXPRACE, NEWRACE1, NEWRACE2, and RACE4. In 1999, NEWRACE1 was not a recoded variable, but was created directly using the donors from IRRACE, with a subsequent imputation used in the rare cases where it was necessary. In 2000 and 2001, NEWRACE1 was a recode involving IRNWRACE and IRHOIND, with the levels given below:

NEWRACE1

- = 1 Non-Hispanic white (IRNWRACE = 1, IRHOIND = 2)
- = 2 Non-Hispanic black (IRNWRACE = 2, IRHOIND = 2)
- = 3 Non-Hispanic American Indian/Alaska Native (IRNWRACE = 3, IRHOIND = 2)
- = 4 Non-Hispanic Native Hawaiian (IRNWRACE = 4, IRHOIND = 2)
- = 5 Non-Hispanic other Pacific Islander (IRNWRACE = 5, IRHOIND = 2)
- = 6 Non-Hispanic Native Hawaiian/other Pacific Islander combination (IRNWRACE = 6, IRHOIND = 2)
- = 7 Non-Hispanic Chinese (IRNWRACE = 7, IRHOIND = 2)
- = 8 Non-Hispanic Filipino (IRNWRACE = 8, IRHOIND = 2)
- = 9 Non-Hispanic Japanese (IRNWRACE = 9, IRHOIND = 2)
- = 10 Non-Hispanic Asian Indian (IRNWRACE = 10, IRHOIND = 2)
- = 11 Non-Hispanic Korean (IRNWRACE = 11, IRHOIND = 2)
- = 12 Non-Hispanic Vietnamese (IRNWRACE = 12, IRHOIND = 2)
- = 13 Non-Hispanic other Asian (IRNWRACE = 13, IRHOIND = 2)
- = 14 Non-Hispanic Asian multiple category (IRNWRACE = 14, IRHOIND = 2)
- = 15 Non-Hispanic more than one race (IRNWRACE = 15, IRHOIND = 2)
- = 16 Hispanic (IRHOIND = 1)

Regardless of whether NEWRACE1 was created using IRRACE donors and a supplemental imputation, or as a recode of IRNWRACE and IRHOIND, all other recoded variables newly created with the CAI questionnaire can be derived from NEWRACE1. EXPRACE was created by collapsing the categories that could contain respondents of different races (Hispanic, all multiple category levels, Hawaiian/other Pacific Islander, and other Asian). In particular,

EXPRACE

- = 1 Non-Hispanic white (NEWRACE1 = 1)

- = 2 Non-Hispanic black (NEWRACE1 = 2)
- = 3 Non-Hispanic American Indian/Alaska Native (NEWRACE1 = 3)
- = 4 Non-Hispanic Native Hawaiian (NEWRACE1 = 4)
- = 5 Non-Hispanic other Pacific Islander (NEWRACE1 = 5)
- = 6 Non-Hispanic Chinese (NEWRACE1 = 7)
- = 7 Non-Hispanic Filipino (NEWRACE1 = 8)
- = 8 Non-Hispanic Japanese (NEWRACE1 = 9)
- = 9 Non-Hispanic Asian Indian (NEWRACE1 = 10)
- = 10 Non-Hispanic Korean (NEWRACE1 = 11)
- = 11 Non-Hispanic Vietnamese (NEWRACE1 = 12)
- = 12 Other (NEWRACE1 = 6, 13, 14, 15, 16)

NEWRACE2 also collapsed levels of NEWRACE1, but this variable collapses all the Asian categories into one category, and the Native Hawaiian and other Pacific Islander categories are collapsed into a single category.

NEWRACE2

- = 1 Non-Hispanic white (NEWRACE1 = 1)
- = 2 Non-Hispanic black (NEWRACE1 = 2)
- = 3 Non-Hispanic American Indian/Alaska Native (NEWRACE1 = 3)
- = 4 Non-Hispanic Pacific Islander (NEWRACE1 = 4, 5, 6)
- = 5 Non-Hispanic Asian (NEWRACE1 = 7-14)
- = 6 Non-Hispanic more than one race (NEWRACE1 = 15)
- = 7 Hispanic (NEWRACE1 = 16)

Finally, RACE4 is very similar to the recoded variable RACE, except that it uses NEWRACE1 rather than IRRACE and IRHOIND. The only visible difference with RACE transpires when people of more than one race were allocated a race based upon their response to QD06 or when the priority rule was used. In RACE4, respondents of more than one race were placed in the “Other” category.

RACE4

- = 1 Non-Hispanic white, single race (NEWRACE1 = 1)
- = 2 Non-Hispanic black, single race (NEWRACE1 = 2)
- = 3 Hispanic (NEWRACE1 = 16)
- = 4 Non-Hispanic other or more than one race (all other values of NEWRACE1)

9.0 Analytic Tables Using Race Variables

The composition of tables that incorporated variables involving race and ethnicity changed from year to year, with the biggest changes occurring, naturally, between 1998 (with PAPI) and 1999 (with CAI). However, through the years, only recoded variables that had Hispanic as a separate level were used in the tables. In other words, the tables never used variables that measured race exclusively, but rather they included variables that simultaneously measured race and ethnicity. In 1998, the variable used was RACE, which included four levels: Hispanic, non-Hispanic white, non-Hispanic black, and non-Hispanic other. Three variables were used in 1999, 2000, and 2001: RACE4 (the same as RACE, except that multiple race respondents were classified into the “other” category), NEWRACE2 (a seven-level variable with separate levels for Hispanics and multiple race respondents, but with Asians collapsed into a single level), and EXPRACE (the “expanded” race/ethnicity variable, with separate levels for the Asian categories, but a single collapsed level that includes levels for Hispanic, other Asian, and the mixed categories).

The following sections summarize the tables involving race and ethnicity in the years from 1998 to 2001. These tables are included in contract deliverables. Across all the years from 1998 to 2001, a document called Summary Tables was produced that included a number of tables where race variables were incorporated as classifying variables. These tables were used to create the Summary of Findings report. In 1998, another report, called the Main Findings report, was also produced that included fewer tables but more detailed discussion on those tables. Other tables were also created. Although these other tables might also involve race and ethnicity, they are not discussed here.

9.1 Tables Involving Race and Ethnicity, 1998

In 1998, two deliverable documents included tables involving race: the Main Findings report (OAS, 2000) and the Summary Tables (RTI, 1999). Because all the tables included in the Summary of Findings report were also included in the Summary Tables report, no separate discussion of the tables in the Summary of Findings report is necessary.

9.1.1 Main Findings Report

In the 1998 Main Findings report, three levels of the variable RACE were used: non-Hispanic white, non-Hispanic black, and Hispanic. The level representing non-Hispanic others (including American Indians and Asians) were not shown or listed separately in the tables due to the low precision resulting from small sample sizes. However, these individuals were included in the calculations of prevalence estimates⁶ for the whole sample (as indicated in the “total” lines of

⁶“Prevalence” refers to the proportion of a given population subgroup that has used a given drug in the lifetime of an individual, in the past year, or in the past month. Whenever the term “prevalence” or “prevalence estimate” is used, it is referring to the proportion within each of these three periods.

the tables). The tables in the 1998 Main Findings report included prevalence estimates for all the drugs in the survey, and usage in the lifetime of drugs with a needle, by age group (12 to 17, 18 to 25, 26 to 34, and 35+) and RACE.⁷ In addition, tables were provided that gave, by RACE, the proportion who reported (1) ease of getting drugs; (2) engaging in risky behaviors; (3) receiving treatment in the past year for drug or alcohol use; (4) not receiving treatment though they felt they needed it; and (5) usage of drugs, tobacco, or alcohol by pregnancy status (among females aged 15 to 44).

9.1.2 Summary Tables

Similar information was provided in the 1998 Summary Tables, with more details, and with comparison numbers between 1997 and 1998. In particular, estimated numbers in the population, proportions, standard errors, p-values, and survey sample sizes were provided for each table where applicable. Unlike the 1998 Main Findings report, the non-Hispanic other level of RACE was provided in most of the tables in the 1998 Summary Tables, along with the other levels. The same prevalence estimate tables given in the 1998 Main Findings report were given in the Summary Tables, where the non-Hispanic other level was presented. In addition, prevalence estimate tables were provided that gave breakdowns by RACE and age group, gender, population density, region, adult education, and current employment. Additional tables using RACE gave breakdowns of prevalence by pregnancy status, and by the numbers of drug or alcohol users reporting physical or mental health problems, dependence, or requiring treatment. Two sets of long-term trend tables were included in the 1998 Summary Tables, each of which included RACE as a classifying variable. One set had trends from 1979, and the other from 1994. In addition, tables were provided that gave, by RACE, the proportion who reported (1) perceptions of great risk for various drugs and (2) ease of getting drugs. RACE was also used in the tables summarizing prevalence in California and Arizona.

9.2 Tables Involving Race and Ethnicity, 1999

The increased scope of questions in the CAI questionnaire, and the much larger sample size, dramatically increased the number of tables that could be produced. As in 1998, these tables were produced in the 1999 Summary Tables (RTI, 2000), which were then summarized in the Summary of Findings report. As in 1998, estimated numbers in the population, proportions, standard errors, p-values, and survey sample sizes were provided for each table where applicable. Three variables involving race and ethnicity were used in the 1999 tables: the 7-level variable NEWRACE2, the 4-level variable RACE4, and “expanded race” 12-level variable EXPRACE. The majority of tables used NEWRACE2. Hence, unlike in 1998, separate information was provided (in most cases) for American Indian/Alaska Natives, Native Hawaiian/Other Pacific Islanders, and respondents of more than one race.

⁷ Also provided in these tables were prevalence estimates by age group and other demographic characteristics (gender, population density, region, adult education, and current employment). As stated earlier, the “non-Hispanic other” level of RACE does not appear in these tables.

Tables involving NEWRACE summarized prevalence estimates for the drugs surveyed. Different age groupings were provided in the 1999 tables, where separate prevalence estimates were calculated, by the levels of NEWRACE2, for the age groups 12 to 17, 18 to 25, 26 or older, 18 or older, and 12 or older. Prevalence estimates were also compared for alcohol use and binge alcohol use⁸ by NEWRACE2 and the age groups 12 to 20 and 21 or older. Comparisons of prevalence estimates for various drugs, by NEWRACE2, were done between (1) college students and non-college students between the ages of 18 and 22; (2) pregnant and non-pregnant females aged 15 to 44; (3) those who perceived great risk of using a given drug and those who did not; (4) those who could easily obtain a given drug and those who could not; and (5) those who had various attributes of social environment variables and those who did not. In addition, tables were provided that summarized dependence on various drugs, treatment need, and whether treatment was received. In the 1999 tables, only the results from 1999 were presented. In subsequent years, however, 2 years' worth of data were presented for the sake of comparison.

The four-level race variable RACE4 was used in situations where more detailed race and ethnicity categories could not be justified due to small sample sizes. RACE4 appeared in tables summarizing dependence on illicit drugs by RACE4 and age first used marijuana or alcohol. RACE4 was also used in the table giving the survey sample size for respondents 18 or older by receipt of and perceived need of mental health treatment or counseling.

When only very general information about prevalence was given, it was possible to use more detailed race and ethnicity categories (given in the variable EXPRACE) and still have sufficient sample sizes. These prevalence estimates were only provided for the combined group of drugs "any illicit drugs," marijuana, cigarettes, alcohol, and binge alcohol, with estimates given over all age groups, and separately by the age categories listed above. Even though Hispanics were collapsed into a single level in EXPRACE with the other Asian and various multiple race categories, Hispanics were classified separately using a second variable listing Hispanic groups. Prevalence estimates were also compared for alcohol and binge alcohol by the EXPRACE and Hispanic group variables and the age groups 12 to 20 and 21 or older. Due to the lack of precision resulting from small sample sizes, non-Hispanics who belonged in the "other" category of EXPRACE could not be separately listed in the table. The non-Hispanic racial categories included in the "other" category are non-Hispanic Native Hawaiian/other Pacific Islander, non-Hispanic multiple Asian, non-Hispanic more than one race and the category non-Hispanic other Asian).

⁸ "Binge" alcohol use is defined as drinking five or more drinks on the same occasion on at least 1 day in the past 30 days. By "occasion" is meant at the same time or within a couple hours of each other.

9.3 Tables Involving Race and Ethnicity, 2000 and 2001

Although the order and makeup of some of the tables changed between 1999 and 2000, the use of race and ethnicity variables for various types of tables did not change. The summary provided in the previous section is also applicable for the 2000 and 2001 tables.

References

Office of Applied Studies. (2000, March). *National Household Survey on Drug Abuse: Main findings 1998* (DHHS Publication No. SMA 00-3381, NHSDA Series H-11). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Office of Management and Budget. (1977). *Race and ethnic standards for federal statistics and administrative reporting* (Statistical Policy Directive No. 15). Washington, DC: Author.

Office of Management and Budget. (1997). *Revisions to the standards for the classification of federal data on race and ethnicity*. Federal Register, 62 (210).

RTI. (1999, August). *1998 National Household Survey on Drug Abuse: Summary tables. Volume I: Prevalence estimates* (prepared for the Substance Abuse and Mental Health Services Administration, Office of Applied Studies, under Contract No. 283-97-9007, Deliverable No. 15, RTI 6908). Research Triangle Park, NC: Author.

RTI. (2000, August). *1999 National Household Survey on Drug Abuse: Summary tables* (prepared for the Substance Abuse and Mental Health Services Administration, Office of Applied Studies, under Contract No. 283-98-9008, Deliverable No. 15, RTI 7190). Research Triangle Park, NC: Author.

RTI. (2001). *SUDAAN user's manual: Release 8.0*. Research Triangle Park, NC: Author

RTI. (2002, July). *2000 National Household Survey on Drug Abuse: Methodological resource book* (prepared for the Substance Abuse and Mental Health Services Administration, Office of Applied Studies, under Contract No. 283-98-9008, Deliverable No. 28, RTI 7190). Research Triangle Park, NC: Author.

Appendix A

**Race and Ethnicity Questions
in 1971 to 2001 NHSDAs**

Race and Ethnicity Questions in 1971 to 2001 NHSDAs

I-V

Year	Mode	Ethnicity Question(s)		Race Question(s)
1971	PAPI	NONE		Which of the following describes you: <ol style="list-style-type: none"> 1. White 2. Negro, Black 3. Puerto Rican, Mexican, or other Latin 4. Other background
1972	PAPI	NONE		¹ Race <ol style="list-style-type: none"> 1. White 2. Negro 3. Puerto Rican or other Latin American group 4. Other (Specify)
1974-76	PAPI	¹ Is the respondent of Spanish origin? <ol style="list-style-type: none"> 1. YES→ 2. NO 3. CANNOT TELL 	¹ To which of the following groups does respondent belong? <ol style="list-style-type: none"> 1. PUERTO RICAN 2. MEXICAN AMERICAN 3. OTHER SPANISH AMERICAN 4. CANNOT TELL 	¹ What is the racial background of the respondent: <ol style="list-style-type: none"> 1. AMERICAN INDIAN 2. ORIENTAL 3. WHITE 4. BLACK 5. OTHER (SPECIFY) 6. CANNOT TELL
1977	PAPI	Are you of Spanish origin, that is, are you from a Spanish-American family? <ol style="list-style-type: none"> 1. YES→ 2. NO 	¹ Which of these types of Spanish-Americans best describes you: Puerto Rican, Mexican, or some other Spanish-American group <ol style="list-style-type: none"> 1. PUERTO RICAN 2. MEXICAN 3. SOME OTHER GROUP (SPECIFY) 	Which of the groups on this card best describes your family origin? <ol style="list-style-type: none"> 1. AMERICAN INDIAN 2. ALASKAN NATIVE 3. ASIAN 4. PACIFIC ISLANDER 5. WHITE 6. BLACK 7. OTHER (SPECIFY) 8. NO ANSWER

¹Question was not asked. Information from interviewer observation only.

Race and Ethnicity Questions in 1971 to 2001 NHSDAs

A-2

Year	Mode	Ethnicity Question(s)	Race Question(s)	
1979	PAPI	<p>Are you of Spanish origin; that is, are you from a Spanish-American family?</p> <p>1. YES→ 2. NO</p>	<p>Which of these types of Spanish-Americans best describe you: Puerto Rican, Mexican, or some other Spanish-American group?</p> <p>1. PUERTO RICAN 2. MEXICAN 3. SOME OTHER GROUP 8. NOT SURE</p>	<p>Which of the groups on this card best describes your family origin?</p> <p>1. AMERICAN INDIAN 2. ALASKAN NATIVE 3. ASIAN 4. PACIFIC ISLANDER 5. WHITE 6. BLACK 7. OTHER (SPECIFY) 8. NO ANSWER</p>
1982	PAPI	<p>Which of these types of Spanish-American groups best describes you: Puerto Rican, Mexican, Cuban, or some other Spanish-American group?</p> <p>1. PUERTO RICAN 2. MEXICAN 3. CUBAN 4. SOME OTHER GROUP (SPECIFY) 8. NOT SURE</p>	<p>²Which of the groups on this card best describes you?</p> <p>1. AMERICAN INDIAN OR ALASKAN NATIVE 2. ASIAN OR PACIFIC ISLANDER 3. BLACK 4. HISPANIC 5. WHITE 6. NOT SURE/NO ANSWER</p>	
1985	PAPI	<p>Which of these Spanish-American groups best describes you:</p> <p>1. Puerto Rican 2. Mexican 3. Cuban, or 4. Some other group</p>	<p>²Which of the groups on this card best describes you? Just give me the number.</p> <p>1. American Indian or Alaskan Native 2. Asian or Pacific Islander 3. Black, not of Hispanic Origin 4. White, not of Hispanic Origin 5. Hispanic 6. Other (Specify)</p>	

²Race question was asked first. Ethnicity question was asked if the response to the race question was “Hispanic.”

Race and Ethnicity Questions in 1971 to 2001 NHSDAs

Year	Mode	Ethnicity Question(s)		Race Question(s)
1988	PAPI	Are you of Hispanic or Spanish origin or descent? 1. YES→ 2. NO	Which of these Hispanic-origin groups best describes you? Are you ... 1. Puerto Rican 2. Mexican, 3. Cuban, or 4. Some other group? (SPECIFY)	Which of the groups on this card best describes you? Just give me the number. 1. AMERICAN INDIAN OR ALASKAN NATIVE 2. ASIAN OR PACIFIC ISLANDER 3. BLACK 4. WHITE 5. OTHER (SPECIFY)
1990-93	PAPI	Are you of Hispanic or Spanish origin or descent? 1. YES→ 2. NO	Which of these Hispanic-origin groups best describes you? Are you ... 1. Puerto Rican, 2. Mexican, 3. Cuban, or 4. Some other group? (SPECIFY)	Which of the groups on this card best describes you? Just give me the number. 1. WHITE 2. BLACK 3. INDIAN (American), ALEUT, ESKIMO 4. ASIAN OR PACIFIC ISLANDER 5. OTHER (SPECIFY)
1994-98	PAPI	Are you of Hispanic or Spanish origin or descent? 1. YES→ 2. NO	Which of these Hispanic-origin groups best describes you? Are you ... 1. Puerto Rican, 2. Mexican, 3. Cuban, or 4. Some other group? (SPECIFY)	Which of the groups on this card best describes you? Just give me the number. 1. WHITE 2. BLACK 3. INDIAN (AMERICAN), ALEUT, ESKIMO 4. ASIAN OR PACIFIC ISLANDER (INCLUDING ASIAN INDIAN) 5. OTHER (SPECIFY)

Race and Ethnicity Questions in 1971 to 2001 NHSDAs

Year	Mode	Ethnicity Question(s)	Race Question(s)
1999	CAI	<p>Are you of Hispanic, Latino, or Spanish origin or descent?</p> <ol style="list-style-type: none"> 1. YES→ 2. NO <p>Which of these groups best describes you? Just give me the number or numbers from the card.</p> <p>TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT.</p> <ol style="list-style-type: none"> 1. MEXICAN/MEXICAN AMERICAN/MEXICANO/CHICANO 2. PUERTO RICAN 3. CENTRAL OR SOUTH AMERICAN 4. CUBAN/CUBAN AMERICAN 5. OTHER (SPECIFY) 	<p>Which of these groups describes you? Just give me the number or numbers from the card.</p> <p>TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT.</p> <ol style="list-style-type: none"> 1. WHITE 2. BLACK/AFRICAN AMERICAN 3. AMERICAN INDIAN OR ALASKA NATIVE 4. NATIVE HAWAIIAN 5. OTHER PACIFIC ISLANDER 6. CHINESE 7. FILIPINO 8. JAPANESE 9. ASIAN INDIAN 10. KOREAN 11. VIETNAMESE 12. OTHER ASIAN 13. OTHER (SPECIFY) <p><i>IF MORE THAN ONE ANSWER TO THE R ACE QUESTION: Which one of these groups, that is [RACEFILL], best describes you? SELECT ONLY ONE ANSWER.</i></p> <ol style="list-style-type: none"> 1. WHITE 2. BLACK/AFRICAN AMERICAN 3. AMERICAN INDIAN OR ALASKA NATIVE 4. NATIVE HAWAIIAN 5. OTHER PACIFIC ISLANDER 6. CHINESE 7. FILIPINO 8. JAPANESE 9. ASIAN INDIAN 10. KOREAN 11. VIETNAMESE 12. OTHER ASIAN 13. SPECIFIED OTHER RACIAL GROUP

Race and Ethnicity Questions in 1971 to 2001 NHSDAs

Year	Mode	Ethnicity Question(s)	Race Question(s)
2000	CAI	<p>Are you of Hispanic, Latino, or Spanish origin or descent?</p> <p>1. YES 2. NO</p> <p><i>IF YES TO HISPANIC QUESTION:</i> Which of these Hispanic, Latino, or Spanish groups best describes you? Just give me the number or numbers from the card.</p> <p>TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT</p> <p>1. MEXICAN/MEXICAN AMERICAN/MEXICANO/CHICANO 2. PUERTO RICAN 3. CENTRAL OR SOUTH AMERICAN 4. CUBAN/CUBAN AMERICAN 5. OTHER (SPECIFY)</p>	<p>Which of these groups describes you? Just give me the number of numbers from the card.</p> <p>TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT.</p> <p>1. WHITE 2. BLACK/AFRICAN AMERICAN 3. AMERICAN INDIAN OR ALASKA NATIVE (AMERICAN INDIAN INCLUDES NORTH AMERICAN, CENTRAL AMERICAN, AND SOUTH AMERICAN INDIANS) 4. NATIVE HAWAIIAN 5. OTHER PACIFIC ISLANDER 6. ASIAN (FOR EXAMPLE: ASIAN INDIAN, CHINESE, FILIPINO, JAPANESE, KOREAN, AND VIETNAMESE) 7. OTHER (SPECIFY)</p> <p><i>IF ASIAN:</i> Which of these Asian groups best describes you? Just give me the number or numbers from the card.</p> <p>TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT.</p> <p>1. ASIAN INDIAN 2. CHINESE 3. FILIPINO 4. JAPANESE 5. KOREAN 6. VIETNAMESE 7. OTHER (SPECIFY)</p> <p><i>IF MORE THAN ONE ANSWER TO THE RACE QUESTION:</i> Which one of these groups, that is [RACEFILL], best describes you? SELECT ONLY ONE ANSWER.</p> <p>1. WHITE 2. BLACK/AFRICAN AMERICAN 3. AMERICAN INDIAN OR ALASKAN NATIVE (AMERICAN INDIAN INCLUDES NORTH AMERICAN, CENTRAL AMERICAN, AND SOUTH AMERICAN INDIANS) 4. NATIVE HAWAIIAN 5. OTHER PACIFIC ISLANDER 6. ASIAN INDIAN 7. CHINESE 8. FILIPINO 9. JAPANESE 10. KOREAN 11. VIETNAMESE 12. OTHER ASIAN 13. SPECIFIED OTHER RACIAL GROUP 14. SPECIFIED OTHER ASIAN GROUP/NOT APPLICABLE</p>

Race and Ethnicity Questions in 1971 to 2001 NHSDAs

Year	Mode	Ethnicity Question(s)	Race Question(s)
2001	CAI	<p>Are you of Hispanic, Latino, or Spanish origin or decent?</p> <ol style="list-style-type: none"> 1. YES 2. NO <p><i>IF YES TO HISPANIC QUESTION:</i> Which of these Hispanic, Latino, or Spanish groups best describes you? Just give me the number or numbers from the card.</p> <p>TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT.</p> <ol style="list-style-type: none"> 1. MEXICAN/MEXICAN AMERICAN/MEXICANO /CHICANO 2. PUERTO RICAN 3. CENTRAL OR SOUTH AMERICAN 4. CUBAN/CUBAN AMERICAN 5. OTHER (SPECIFY) 	<p>Which of these groups best describes you? Just give me the number of numbers from the card.</p> <p>TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT.</p> <ol style="list-style-type: none"> 1. WHITE 2. BLACK/AFRICAN AMERICAN 3. AMERICAN INDIAN OR ALASKA NATIVE (AMERICAN INDIAN INCLUDES NORTH AMERICAN, CENTRAL AMERICAN, AND SOUTH AMERICAN INDIANS) 4. NATIVE HAWAIIAN 5. OTHER PACIFIC ISLANDER 6. ASIAN (FOR EXAMPLE, ASIAN INDIAN, CHINESE, FILIPINO, JAPANESE, KOREAN, AND VIETNAMESE) 7. OTHER (SPECIFY) <p><i>IF ASIAN:</i> Which of these Asian groups best describes you? Just give me the number of numbers from the card.</p> <p>TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT.</p> <ol style="list-style-type: none"> 1. ASIAN INDIAN 2. CHINESE 3. FILIPINO 4. JAPANESE 5. KOREAN 6. VIETNAMESE 7. OTHER (SPECIFY) <p><i>IF MORE THAN ONE ANSWER TO THE RACE QUESTION:</i> Which one of these groups, that is [ARCEFILL], best describes you? SELECT ONLY ONE ANSWER FROM THE CA TEGORIES SHOWN IN BLUE BELOW.</p> <ol style="list-style-type: none"> 1. WHITE 2. BLACK/AFRICAN AMERICAN 3. AMERICAN INDIAN OR ALASKA NATIVE (AMERICAN INDIAN INCLUDES NORTH AMERICAN, CENTRAL AMERICAN, AND SOUTH AMERICAN INDIANS) 4. NATIVE HAWAIIAN 5. OTHER PACIFIC ISLANDER 6. ASIAN INDIAN 7. CHINESE 8. FILIPINO 9. JAPANESE 10. KOREAN 11. VIETNAMESE 12. SPECIFIED OTHER RACIAL GROUP 13. SPECIFIED OTHER ASIAN GROUP/NOT APPLICABLE 14. NONE OF THESE

Appendix B

**Main Race Question Responses
from the 2001 NHSDA**

Table 1. 2001 NHSDA - Main Race Question (QD06) (Note: shaded sections are changes made to the 2000 instrument)

QD06	[IF MORE THAN ONE RESPONSE SELECTED IN QD05] Which one of these groups, that is [RACEFILL], best describes you?	
SELECT ONLY ONE ANSWER FROM THE CATEGORIES SHOWN IN BLUE BELOW.		
	1	WHITE
	2	BLACK / AFRICAN AMERICAN
	3	AMERICAN INDIAN OR ALASKA NATIVE (AMERICAN INDIAN INCLUDES NORTH AMERICAN, CENTRAL AMERICAN, AND SOUTH AMERICAN INDIANS)
	4	NATIVE HAWAIIAN
	5	OTHER PACIFIC ISLANDER
	6	ASIAN INDIAN
	7	CHINESE
	8	JAPANESE
	9	FILIPINO
	10	KOREAN
	11	VIETNAMESE
	12	IF QD05 = 7, FILL TEXT FROM QD05OTHR
		IF QD05 NE 7 FILL WITH "OTHER (SPECIFY)"
	13	IF QD05ASIA = 7, FILL TEXT FROM QD05OHTA
		IF QD05ASIA = BLANK, FILL WITH "NOT APPLICABLE"
	14	NONE OF THESE
	DK/REF	
	INTERVIEWER NOTE:	
	If the respondent expresses any concerns about this question or asks why it has been included, please press F8 and record what they have said in the Comment Box. Do not probe for this information! Only record a comment if the respondent spontaneously brings up the topic.	
[NOTE: ONLY CODES FOR RESPONSE CATEGORIES ENTERED IN QD05 OR QD05OTHR OR QD05ASIA OR QD05OHTA WILL BE ACTIVE FOR THIS QUESTION. IF THE INTERVIEW ENTERS AN INACTIVE RESPONSE CATEGORY, THE RANGE ERROR BOX WILL APPEAR. CATEGORY 14 WILL ALWAYS SHOW IN BLUE AND WILL ALWAYS BE A SELECTABLE RESPONSE.]		

Table 2. Distribution of 2001 NHSDA Respondents with Multiple Race, by Main Race (Unweighted)

Responses to Question QD06 (Main Race Question)	Number of Respondents Selecting More Than One Race in Question QD05	Percent
Provided a Single Race	1,082	89.27
None of These	106	8.75
Don't Know	17	1.40
Refused	7	0.58
Total	1,212	100.00

Note: There were 68,929 respondents completing the questionnaire in 2001. Of these, 1,212 (1.8 percent) selected more than one race. This percentage is consistent with what was found in the 1999 survey (1.9 percent) and 2000 survey (1.7 percent).

Table 3. Distribution of 2001 NHSDA Respondents with Multiple Race, by Main Race and Hispanic Origin (Unweighted)

Main Race	Hispanic	Non-Hispanic	Don't Know/ Refused	Total
WHITE	45	437	2	484
BLACK	28	162	0	190
AMERICAN INDIAN/ALASKA NATIVE	31	108	0	139
NATIVE HAWAIIAN	11	70	0	81
OTHER PACIFIC ISLANDER	1	13	0	14
ASIAN INDIAN	5	0	0	5
CHINESE	3	9	0	12
JAPANESE	13	25	0	38
FILIPINO	4	35	0	39
KOREAN	0	5	0	5
VIETNAMESE	0	1	0	1
OTHER	46	28	0	74
NONE OF THESE	27	79	0	106
DON'T KNOW	2	15	0	17
REFUSED	0	7	0	7
<i>BLANK</i> ¹	<i>8,592</i>	<i>59,077</i>	<i>48</i>	<i>67,717</i>
Total (excluding Blank)	216	994	2	1,212

¹The row labeled "BLANK" corresponds to respondents who did not report more than one race in question QD05. Adding total BLANK (67,717) to total excluding BLANK (1,212) equals 68,929, which is the total number of respondents in the 2001 survey.

Table 4. Distribution of 2001 NHSDA Respondents with Multiple Race, by Main Race and Age Group (Unweighted)

Main Race	Age in Years			TOTAL
	12 to 17	18 to 25	26+	
WHITE	205	160	119	484
BLACK	96	61	33	190
AM INDIAN/ALASKA NATIVE	63	44	32	139
NAT HAWAIIAN	32	29	20	81
OTHER PACIFIC ISLANDER	5	5	4	14
ASIAN INDIAN	4	1	0	5
CHINESE	4	5	3	12
JAPANESE	20	13	5	38
FILIPINO	16	16	7	39
KOREAN	2	3	0	5
VIETNAMESE	1	0	0	1
OTHER	29	30	15	74
NONE OF THESE	48	40	18	106
DON'T KNOW	5	10	2	17
REFUSED	3	3	1	7
<i>BLANK</i> ¹	22,599	22,238	22,880	67,717
TOTAL (excluding BLANK)	533	420	259	1,212

¹The row labeled BLANK corresponds to respondents who did not report more than one race in question QD05. Adding total BLANK (67,717) to total excluding BLANK (1,212) equals 68,929, which is the total number of respondents in the 2001 survey.

Table 5. Respondent Comments to the 2001 NHSDA Main Race Question

Respondent Number	Main Race Response	Respondent Comment
1	NONE OF THESE	Half and half
2	NONE OF THESE	Half and half
3	REFUSED	Respondent says that they are both part of her heritage, so she will not change her answer!!!!
4	NONE OF THESE	Cannot differentiate between the two
5	NONE OF THESE	Both are equal
6	NONE OF THESE	Not more than the other, just both
7	NONE OF THESE	Not one more than the other, black and white
8	BLACK	Biracial
9	NONE OF THESE	Respondent says both
10	NONE OF THESE	Both, not just one
11	REFUSED	Respondent says both describe him—father is Chinese and mother is white
12	JAPANESE	1 (white), 8 (Japanese)
13	NONE OF THESE	I can't consider myself more one than the other/either combination of two of these if not all of them
14	NONE OF THESE	Chooses not to distinguish himself
15	NONE OF THESE	Respondent is 1 (white) and 2 (black)
16	NONE OF THESE	Respondent said that when she is in the states, which is most of the time, she feels more Thai, but when she is in Thailand, she feels more white
17	AMERICAN INDIAN	She is both white and South American/Latino; she feels that she shouldn't have to chose which nationality she is— "she's both equally"
18	NONE OF THESE	Respondent is half Japanese and half Mexican
19	JAPANESE	50/50

Table 5 (continued)

Respondent Number	Main Race Response	Respondent Comment
20	NONE OF THESE	She said that genetically more Asian, but was raised more western mind-set
21	NONE OF THESE	Half white, half Vietnamese
22	NONE OF THESE	I would prefer to be both and not one or the other
23	NONE OF THESE	½ of each, not picking one or the other
24	NONE OF THESE	I'm biracial
25	WHITE	Does not choose one over the other
26	NONE OF THESE	She states neither describes her
27	NONE OF THESE	Equally each
28	NONE OF THESE	Respondent does not choose between races, feels equal between white and American Indian
29	NONE OF THESE	I'm 50/50

Note: As requested, SAMHSA recorded whether the respondents in any way questioned why the interviewer was asking to select only one race. Of the 1,212 respondents providing more than one race, comments were recorded by field interviewers from 29 respondents.