

Appendix 6

Monitoring the Future Survey methodology and definitions of terms

Note: The following information was excerpted from Jerald G. Bachman, Lloyd D. Johnston, and Patrick M. O'Malley, *Monitoring the Future 2000* (Ann Arbor, MI: Institute for Social Research, University of Michigan, 2001), pp. 2-11, 13, 14; Lloyd D. Johnston et al., *Monitoring the Future National Survey Results on Drug Use, 1975-2003*, Volumes I and II (Bethesda, MD: U.S. Department of Health and Human Services, 2004); and information provided by the Monitoring the Future Project. Non-substantive editorial adaptations have been made.

Survey methodology

The research design involves annual data collections from high school seniors during the spring of each year, beginning with the class of 1975. Each data collection takes place in approximately 120 to 146 public and private high schools selected to provide an accurate cross-section of high school seniors throughout the coterminous United States.

Since 1986, the results of a followup survey of those young adults 1 to 10 years beyond high school have been presented. These results should accurately characterize approximately 85% of the young adults in the class cohorts 1 to 10 years beyond high school who are high school graduates. The high school dropout segment, missing from the senior year surveys, also is missing from the followup segments.

Also, since 1980, the results of followup surveys of those high school students who have continued on to college have been presented. The college sample is limited to the most typical one for college attendance: 1 to 4 years past high school, which corresponds to the modal ages of 19 to 22 years old. This age category should encompass approximately 70% to 75% of all students enrolled in college full-time.

Sampling procedures

The procedure for securing a nationwide sample of high school seniors is a multistage one. Stage 1 is the selection of particular geographic areas, Stage 2 is the selection of one or more high schools in each area, and Stage 3 is the selection of seniors within each high school.

Stage 1: Geographic areas. The geographic areas used in this study are the

primary sampling units (PSUs) developed by the Sampling Section of the Survey Research Center (SRC) for use in the Center's nationwide interview studies. These consist of 74 primary areas throughout the coterminous United States—including the 12 largest metropolitan areas, which contain about 30% of the Nation's population. Of the 62 other primary areas, 10 are in the Northeast, 18 in the North Central area, 24 in the South, and 10 in the West. Because these same PSUs are used for personal interview studies by the SRC, local field representatives can be assigned to administer the data collections in practically all schools.

Stage 2: Schools. In the major metropolitan areas more than one high school is often included in the sampling design; in most other sampling areas a single high school is sampled. In all cases, the selections of high schools are made such that the probability of drawing a school is proportionate to the size of its senior class. The larger the senior class (according to recent records), the higher the selection probability assigned to the high school. When a sampled school is unwilling to participate, a replacement school as similar to it as possible is selected from the same geographic area.

Stage 3: Students. Within each selected school, up to 350 seniors may be included in the data collection. In schools with fewer than 350 seniors, the usual procedure is to include all of them in the data collection. In larger schools, a subset of seniors is selected either by randomly sampling classrooms or by some other random method that is convenient for the school and judged to be unbiased. Sample weights are assigned to each respondent so as to take account of variations in the sizes of samples from one school to another, as well as the (smaller) variations in selection probabilities occurring at the earlier stages of sampling.

The three-stage sampling procedure described above yielded the number of participating schools and students indicated in Table 1.

One limitation in the design is that it does not include in the target population those young men and women who drop out of high school before graduation (or before the last few months of the senior year, to be more precise). This excludes a relatively small proportion of each age cohort—between 15% and 20% of each age cohort nationally, according to the U.S. Census Bureau. This is not an unimportant segment, since certain behaviors such as illicit drug use and delinquency tend to be higher than average in this group. However, the addition of a representative sample of dropouts would increase the cost of the present research enormously, because of their dispersion and generally

higher level of resistance to being located and interviewed.

For the purposes of estimating characteristics of the entire age group, the omission of high school dropouts does introduce certain biases; however, their small proportion sets outer limits on the bias. For the purposes of estimating changes from one cohort of high school seniors to another, the omission of dropouts represents a problem only if different cohorts have considerably different proportions who drop out. The Source has no reason to expect dramatic changes in those rates for the foreseeable future, and recently published Census Bureau statistics indicate a great deal of stability in dropout rates since 1970.

Some may use the high school data to draw conclusions about changes for the entire age group. The Source does not encourage such extrapolation but suspects that the conclusions reached often would be valid, since over 80% of the age group is in the surveyed segment of the population and the Source expects that changes among those not in school are very likely to parallel the changes among those who are.

One other important feature of the base-year sampling procedures should be noted. All schools (except for half of the initial 1975 sample) are asked to participate in two data collections, thereby permitting replacement of half of the total sample of schools each year. One motivation for requesting that schools participate for 2 years is administrative efficiency; it is a costly and time-consuming procedure to secure the cooperation of schools, and a 2-year period of participation cuts down that effort substantially. Another important advantage is that whenever an appreciable shift in scores from one graduating class to the next is observed, it is possible to compare whether the shift might be attributable to some differences in the newly sampled schools. This is done simply by repeating the analysis using only the 60 or so schools that participated both years. Thus far, the half-sample approach has worked well; and examination of drug prevalence data from the "matched half-samples" shows that the half samples of repeat schools yielded drug prevalence trends that were virtually identical to trends based on all schools.

Questionnaire administration

Questionnaire administration in each school is carried out by the local SRC representatives and their assistants, following standardized procedures detailed in a project instruction manual. The questionnaires are administered in classrooms during normal class periods whenever possible, although circumstances in some schools require the

use of larger group administrations. Teachers are not asked to do anything more than introduce the SRC staff members and (in most cases) remain in the classroom to help guarantee an orderly atmosphere for the survey. Teachers are urged to avoid walking around the room, so that students may feel free to write their answers without fear of being observed.

The actual process of completing the questionnaires is straightforward. Respondents are given sharpened pencils and asked to use them because the questionnaires are designed for automatic scanning. Most respondents can finish within a 45-minute class period; for those who cannot, an effort is made to provide a few minutes of additional time.

Content areas and questionnaire design

Drug use and related attitudes are the topics that receive the most extensive coverage in the Monitoring the Future Project; but the questionnaires also deal with a wide range of other subject areas, including attitudes about government, social institutions, race relations, changing roles for women, educational aspirations, occupational aims, and marital and family plans, as well as a variety of background and demographic factors. The list below provides an outline of the 20 general subject areas into which all items are categorized. Given this breadth of content, the study is not presented to respondents as a "drug use study," nor do they tend to view it as such.

Measurement content areas

A. Drugs. Drug use and related attitudes and beliefs, drug availability and exposure, surrounding conditions and social meanings of drug use. Views of significant others regarding drugs.

B. Education. Educational lifestyle, values, experiences, and environments.

C. Work and leisure. Vocational values, meaning of work and leisure, work and leisure activities, preferences regarding occupational characteristics and type of work setting.

D. Sex roles and family. Values, attitudes, and expectations about marriage, family structure, sex roles, and sex discrimination.

E. Population concerns. Values and attitudes about overpopulation and birth control.

F. Conservation, materialism, equity, etc. Values, attitudes, and expectations related to conservation, pollution, materialism, equity, and the sharing of resources. Preferences regarding type of dwelling and urbanicity.

G. Religion. Religious affiliation, practices, and views.

H. Politics. Political affiliation, activities, and views.

I. Social change. Values, attitudes, and expectations about social change.

J. Social problems. Concern with various social problems facing the Nation and the world.

K. Major social institutions. Confidence in and commitment to various major social institutions (business, unions, branches of government, press, organized religion, military, etc.).

L. Military. Views about the armed services and the use of military force. Personal plans for military service.

M. Interpersonal relationships. Qualitative and quantitative characteristics of cross-age and peer relationships. Interpersonal conflict.

N. Race relations. Attitudes toward and experiences with other racial groups.

O. Concern for others. Concern for others; voluntary and charitable activities.

P. Happiness. Happiness and life satisfaction, overall and in specific life domains.

Q. Other personality variables. Attitudes about self (including self-esteem), locus of control, loneliness, risk-taking, trust in others, importance placed on various life goals, counter-culture orientation, hostility.

R. Background. Demographic and family background characteristics, living arrangements.

S. Deviant behavior and victimization. Delinquent behaviors, driving violations and accidents (including those under the influence of drugs), victimization experiences.

T. Health. Health habits, somatic symptoms, medical treatments.

Because many questions are needed to cover all of these topic areas, much of the questionnaire content was divided into five different questionnaire forms in 1976-88 and six different questionnaire forms for 1989 and beyond, which are distributed to participants in an ordered sequence that produces virtually identical subsamples. About one-third of each questionnaire form consists of key or "core" variables that are common to all forms. All demographic variables and some measures of drug use are included in this "core" set of measures. This use of the full sample for drug and demographic

measures provides a more accurate estimation on these dimensions and also makes it possible to link them statistically to all of the other measures that are included in a single form only.

Representativeness and validity

The samples for this study are intended to be representative of high school seniors throughout the 48 coterminous States. As previously mentioned, this definition of the sample excludes one important portion of the age cohort: those who have dropped out of high school before nearing the end of the senior year. But given the aim of representing high school seniors, it is useful to consider the extent to which the obtained samples of schools and students are likely to be representative of all seniors and the degree to which the data obtained are likely to be valid.

There are at least four ways in which survey data of this sort might fall short of being fully accurate. First, some sampled schools refuse to participate, which could introduce some bias. Second, the failure to obtain questionnaire data from 100% of the students sampled in participating schools also could introduce bias. Third, the answers provided by participating students are open to both conscious and unconscious distortions, which could reduce validity. Finally, limitations in sample size and/or design could place limits on the accuracy of estimates.

School participation

As noted in the description of the sampling design, schools are invited to participate in the study for a 2-year period. With very few exceptions, each school that has participated for one data collection has agreed to participate for a second. Thus far, approximately 65% of the schools initially invited to participate have agreed to do so each year; for each school refusal, a similar school (in terms of size, geographic area, urbanicity, etc.) was recruited as a replacement. However, because securing high school cooperation has become more difficult in recent years, payment of schools as a means of increasing their incentive to participate was implemented in the 2003 survey.

The selection of replacement schools almost entirely removes problems of bias in region, urbanicity, and the like that might result from certain schools refusing to participate. Other potential biases are more subtle, however. For example, if it turned out that most schools with "drug problems" refused to participate, that could seriously bias the drug estimates derived from the sample. And if any other single factor was dominant in most refusals, that also might suggest a source of serious bias. In fact, however, the reasons

for schools refusing to participate are varied and largely a function of happenstance events of the particular year. Thus, there is a fair amount of confidence that school refusals have not seriously biased the surveys.

Student participation

Completed questionnaires are obtained from approximately 80 to 85% of all students sampled. The single most important reason that students are missed is that they are absent from class at the time of data collection, and in most cases it is not workable to schedule a special followup data collection for them.

Students with high rates of absenteeism also report above-average drug use. Therefore there is some degree of bias introduced by missing the absentees. That bias could be largely corrected through the use of special weighting; however, this course was not chosen because the bias in estimates (for drug use, where the potential effect was hypothesized to be the largest) was determined to be quite small and because the necessary weighting procedures would have introduced undesirable complications.

In addition to absenteeism, student nonparticipation occurs because of schedule conflicts with school trips and other activities that tend to be more frequent than usual during the final months of the senior year. Of course, some students refuse to complete or turn in a questionnaire. However, the proportion of explicit refusals amounts to less than 1.5% of the target sample for each grade.

Research design for the surveys of lower grades

Beginning in 1991 the study was expanded to include nationally representative samples of eighth and tenth grade students. In general, the procedures used for the annual surveys of eighth and tenth grade students closely parallel those used for high school seniors, including the procedures for selecting schools and students, questionnaire administrations, and questionnaire formats. A major exception is that only two different questionnaire forms were used in 1991-96 and four forms were used beginning in 1997, rather than the six forms used with seniors. Identical forms are used for both eighth and tenth grades, and, for the most part, questionnaire content is drawn from the twelfth grade questionnaires. Thus, key demographic variables and measures of drug use and related attitudes and beliefs are generally identical for all three grades. Fewer questions about lifestyles and values are included in these forms than in the twelfth grade forms, in part because it is believed that many of these attitudes are more likely to be formed by twelfth grade, and therefore are best monitored there. For the national survey

of eighth graders, approximately 150 schools are sampled, and approximately 16,000 to 19,000 students are surveyed. For the tenth graders, approximately 130 schools are sampled, and approximately 14,000 to 17,000 students are surveyed. (See Table 2.)

Research design for the followup surveys after high school

Beginning with the graduating class of 1976, a sample of each class is followed and surveyed by mail after high school graduation. From the approximately 15,000 to 17,000 seniors originally participating in a given class, a representative sample of 2,400 individuals was chosen for followup. In order to ensure sufficient numbers of drug users in the followup surveys, those fitting certain criteria of current drug use (that is, those reporting 20 or more uses of marijuana or use of any of the other illicit drugs in the previous 30 days) were selected with higher probability (by a factor of 3.0) than the remaining seniors. Differential weighting is used in all followup analyses to compensate for the differential sampling probabilities.

The 2,400 selected respondents from each class were randomly assigned to one of two matching groups of 1,200 each; one group was surveyed on even-numbered calendar years, and the other group was surveyed on odd-numbered years. This biennial procedure is intended to reduce respondent burden.

Until 2002, each respondent was followed for up to seven times; at the seventh followup, which would occur either 13 or 14 years after graduation, the respondents had reached the modal ages of 31 or 32. Beginning in 2002, the seventh followup was discontinued, and each respondent was followed for up to six times, corresponding to the modal ages of 29 or 30. Additional followups occur at modal ages 35, 40, and 45.

Followup procedures

Using information provided by respondents at the time of the senior survey (name, address, phone number, and the name and address of someone who would always know how to reach them), students selected for the panels are contacted by mail. Newsletters are sent each year, and name and address corrections are requested. Questionnaires are sent by certified mail in the spring of each year. A check for \$5.00 made out to the respondent is attached to the front. Beginning with the class of 1992, the followup checks have been raised to \$10.00 to compensate for the effects of inflation over the life of the study. Reminder letters and post cards are sent at fixed intervals thereafter and finally, those not responding receive a

prompting phone call from the Survey Research Center's phone interviewing facility in Ann Arbor, MI. If requested, a second copy of the questionnaire is sent.

Panel retention rates

Retention rates in the biennial followups of all panel members ages 19 to 30 (corresponding to the first six followups) decline with the length of the followup interval. For the 5-year period from 1999 to 2003, the response rate in the first followup (corresponding to 1 to 2 years past high school) averaged 60%; for the second through the sixth followups (corresponding to 3 to 12 years past high school) response rates averaged 54%. Among the very long-term respondents--the 35- and 40-year-olds--the retention rates remain good. Among the 35-year-old respondents surveyed from 1999 to 2003 (corresponding to 17 years past high school), the average response rate was 52%. Among the 40-year-old respondents surveyed from 1999 (the first survey of this age group) to 2003, corresponding to a 22-year followup interval, the average retention rate was 59%. Among 45-year-olds surveyed in 2003, the retention rate was 59%.

Since attrition is to a modest degree associated with drug use, corrections to the prevalence estimates are presented for the followup panels. These raise the prevalence estimates from what they would be uncorrected, but only slightly. It is believed that the resulting estimates are the most accurate obtainable, but still low for the age group as a whole due to the omission of dropouts and absentees from the population covered by the original panels.

Validity of self-report data

Survey measures of delinquency and drug use depend upon respondents reporting what are, in many cases, illegal acts. Thus, a critical question is whether such self-reports are likely to be valid. Like most studies dealing with these areas, there is no direct, objective validation of the present measures; however, the considerable amount of inferential evidence that exists strongly suggests that the self-report questions produce largely valid data. A number of factors suggest a reasonable amount of confidence about the validity of the responses to what are presumably among the most sensitive questions in the study: a low nonresponse on the drug questions, a large proportion admitting to some illicit drug use, the consistency of findings across several years of the present study, strong evidence of construct validity (based on relationships observed between variables), a close match between these data and the findings from other studies using other methods, and the findings from

several methodological studies that have used objective validation methods.

Accuracy of the sample

A sample survey never can provide the same level of accuracy as would be obtained if the entire target population were to participate in the survey--in the case of the present study, about 2.5 to 3.0 million seniors per year. But perfect accuracy of this sort would be extremely expensive and certainly not worthwhile considering that a high level of accuracy can be obtained by a carefully designed probability sample. The accuracy of the sample in this study is affected both by the size of the student sample and by the number of schools in which they are clustered. Virtually all estimates based on the total sample have confidence intervals of plus or minus 1.5 percentage points or smaller--sometimes considerably smaller.

Interpreting racial differences

Data are given for the two largest racial subgroups in the population--those who identify themselves as white or Caucasian and those who identify themselves as black or African-American. Data are not given for the other ethnic categories (American Indians, Asian Americans, Mexican Americans, Puerto Ricans, or other Latin Americans) since each of these groups comprises a small percentage of the sample in any given year, which means that their small Ns (in combination with their clustered groupings in a limited number of schools) would yield estimates that would be too unreliable. In fact, even black respondents--who constitute approximately 12% of each year's sample--are represented by only 269 to 425 respondents per year on any single questionnaire form. Further, because the sample is a stratified clustered sample, it yields less accuracy than would be yielded by a pure random sample of equal size. Therefore, because of the limited number of cases, the margin of sampling error around any statistic describing black respondents is larger than for most other subgroups described in this survey. There are factors in addition to unreliability, however, that could be misleading in the interpretation of racial differences. Given the importance that has been placed on various racial differences reported in the social science literature, the reader is cautioned to consider the various factors that could account for differences. These factors fall into three categories: differential representation in the sample, differential response tendencies, and the confounding of race with a number of other background and demographic characteristics.

Differential representation--A smaller segment of the black population than of the white population of high school age is

represented by the data contained here. Insofar as any characteristic is associated with being a school dropout or absentee, it is likely to be somewhat disproportionately underrepresented among blacks in the sample.

Differential response tendencies--In examining the full range of variables, certain racial differences in response tendencies were noted. First, the tendency to state agreement in response to agree-disagree questions is generally somewhat greater among blacks than among whites.

There also is a somewhat greater than average tendency for black respondents to select extreme answer categories on attitudinal scales. For example, even if the same proportion of blacks as whites felt positively (or negatively) about some subject, fewer whites are likely to say they feel very positively (or very negatively). In the process of interpreting racial differences, the reader should be aware that differences in responses to particular questions may be related to these more general tendencies.

A somewhat separate issue in response tendency is a respondent's willingness to answer particular questions. An exaggerated missing data rate for black males on the set of questions dealing with the respondent's own use of illicit drugs has been observed. Clearly, a respondent's willingness to be candid on such questions depends on his or her trust of the research process and of the researchers themselves. The reader is advised to consult the Source for exceptional levels of missing data when making comparisons on any variable in which candor is likely to be reduced by lower system trust. One bit of additional evidence related to trust in the research process is that higher proportions of blacks than whites indicated that if they had used marijuana or heroin they would not have been willing to report it in the survey.

Covariance with other factors--Some characteristics such as race are highly correlated with other variables--variables that may in fact explain some observed racial differences. Put another way, at the aggregate level one might observe a considerable racial difference on some characteristic, but once one controls for certain background characteristics such as socio-economic level or region of the country--that is, comparing the black respondents with whites who come from similar backgrounds--there may be no racial difference at all.

Definitions of terms

Drug types--Definitions or identifiers used in survey forms include:

Marijuana--pot, grass or hashish;
Other psychedelics--mescaline,

peyote, psilocybin, PCP. In 2001, the question text was changed from "other psychedelics" to "other hallucinogens" and "shrooms" was added to the list of examples;

Amphetamines--uppers, pep pills, bennies, speed;

Quaaludes--quads, methaqualone;
Barbiturates--downers, goofballs, reds, yellows;

Heroin--smack, horse;

Other narcotics--methadone, opium, codeine, paregoric. In 2002, the list of examples of narcotics other than heroin was updated by replacing Talwin, laudanum, and paregoric with Vicodin, OxyContin, and Percocet.

Inhalants--glue, aerosols, laughing gas;
Tranquilizers--Librium, Valium, Miltown.

In 2001, Miltown was replaced with Xanax.

Beginning with the 1979 survey, amyl and butyl nitrites were considered "other inhalants" for questions on one alternate survey form (N is one-fifth of total sample size in 1979-88 and N is one-sixth of total sample size in 1989-2003). This change was made because not all users of this subclass of inhalants were reporting themselves as inhalant users. Hallucinogen use had been similarly underestimated because some users of the hallucinogenic drug PCP do not report themselves as users of hallucinogens--even though PCP was included as an example of a hallucinogenic drug in earlier surveys and on other questions. The alternate questionnaire form contained a special set of questions about PCP that provided other street names for it (e.g., angel dust). As a result of these definition changes, since 1979 data for drug use in these two drug classes have been adjusted for underreporting. For more information, see the Source.

Four-year college plans--Percentage distributions are given separately for (1) respondents who indicate that they "definitely will" or "probably will" graduate from a 4-year college program and (2) those who say that they "definitely won't" or "probably won't" graduate from a 4-year college program. Respondents not responding are omitted from both columns. A number of those who do not expect to complete a 4-year college program do expect to get some post-secondary education.

Illicit drug use: Lifetime--Percentage distributions are given separately for five mutually exclusive subgroups differentiated by their degree of involvement with illicit drugs. Eligibility for each category is defined below.

None--Includes respondents who indicated that they had not used marijuana at any time and did not report use of any of the following illicit drugs in their lifetime: LSD, other psychedelics, cocaine, amphetamines, tranquilizers, methaqualone, barbiturates, heroin, or other narcotics.

Marijuana only--Includes other respondents who indicated that they had used marijuana (or hashish) but had never used any of the other illicit drugs listed above.

Few pills--Includes respondents who indicated having used one or more of the above listed drugs (other than marijuana) but who had not used any one class of them on three or more occasions and who had not used heroin at all.

More pills--Includes respondents who had used any of the above listed drugs (other than marijuana) on three or more occasions but who had never used heroin.

Any heroin--Includes respondents who indicated having used heroin on one or more occasions in their lifetime.

Race--Percentage distributions are given separately for those describing themselves as "white or Caucasian" and "black or African-American." Comparable data for the other racial or ethnic groups (Mexican Americans, Asian Americans, American Indians, etc.) are not shown because of the low number of cases in each group. For tables in section 3 presenting use of alcohol and illicit drugs, the category Hispanic is added, which includes respondents who in 1975-1990 describe themselves as Mexican American or Chicano, or Puerto Rican or other Latin American. After 1990, this group includes respondents who describe themselves as Mexican American or Chicano, Cuban American, Puerto Rican American, or other Latin American. After 1994, the term Puerto Rican American was shortened to Puerto Rican.

Region--Percentage distributions are given separately for respondents living in each of four mutually exclusive regions of the country. The regional classifications are based on U.S. Census Bureau categories and are defined as follows:

Northeast--Census classifications of New England and Middle Atlantic States; includes Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, and Pennsylvania.

North Central--Census classifications of East North Central and West North Central States; includes Ohio, Indiana, Illinois, Michigan, Wisconsin, Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, and Kansas.

South--Census classifications of South Atlantic, East South Central, and West South Central States; includes Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida, Kentucky, Tennessee, Alabama, Mississippi, Arkansas, Louisiana, Oklahoma, and Texas.

West--Census classifications of Mountain and Pacific States; includes Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, Nevada, Washington, Oregon, and California.

Sex--Percentage distributions are given separately for males and females. Respondents with missing data on the question asking the respondent's sex are omitted from both groupings.

Weighted number of cases (N)--The number of cases is stated in terms of the weighted number of respondents rather than the actual number, since all percentages have been calculated using weighted cases. The actual number of respondents is about 15% higher than the weighted number for data collected in 1975, 1976, and 1977. For data collected in 1978 or later, the actual number of respondents is roughly equal to the weighted number. Weighting is used to improve the accuracy of estimates by correcting for unequal probabilities of selection that arise in the multi-stage sampling procedures. Table 3 presents the number of weighted cases for each subgroup of the high school seniors samples.