

---

---

# **2011 National Youth Tobacco Survey**

## **Methodology Report**

---

---

PRINT ONLY

## **Introduction**

### *Overview of the National Youth Tobacco Survey (NYTS)*

In conjunction with the state Youth Tobacco Survey (YTS), the National Youth Tobacco Survey (NYTS) was developed to provide the data necessary to support the design, implementation, and evaluation of state and national tobacco prevention and control programs (TCPs).<sup>1,2</sup> In addition, NYTS data may supplement other existing data sources such as the Youth Risk Behavior Surveillance System (YRBSS), which provides prevalence estimates for selected tobacco use behaviors among high school students. The NYTS broadens our understanding by providing more comprehensive data for both middle school (grades 6-8) and high school (grades 9-12) students regarding tobacco use (bidis, cigarettes, cigars, kreteks\*, tobacco pipes, and smokeless tobacco); exposure to secondhand smoke; smoking cessation; school curriculum; minors' ability to purchase or obtain tobacco products; knowledge and attitudes about tobacco; and familiarity with pro-tobacco and anti-tobacco media messages. NYTS data also serve as essential benchmarks against which TCPs can assess the magnitude of youth tobacco use. The NYTS provides multiple measures and data for TU-18 of the Healthy People 2020 objectives (USDHHS, 2010). This includes two sub-objectives: TU-18.1 (internet advertising and promotion) and TU-18.2 (magazine and newspaper advertising and promotion).

First conducted during fall 1999<sup>4</sup> and again during spring 2000<sup>5</sup>, 2002<sup>6</sup>, 2004<sup>7</sup>, 2006<sup>8</sup>, 2009, and 2011, the NYTS surveys provide data that are representative of all middle school and high school students in the 50 states and the District of Columbia. The current NYTS was implemented in spring 2012.

### *Overview of the 2011 NYTS Methodology*

The 2011 NYTS employed a stratified three-stage cluster sample design to produce a nationally representative sample of middle school and high school students in the United States. Non-Hispanic black and Hispanic students were oversampled. Sampling procedures were probabilistic and conducted without replacement at all stages, and entailed selection of Primary Sampling Units (PSUs) (a county, or a group of small counties, or part of a very large county) within each created stratum, of schools within each selected PSU, and, lastly, of students within each selected school. Participating students completed the survey via a pencil and paper, self-administered, scannable questionnaire booklet.

Participation in the NYTS was voluntary at both the school and student level. At the student level, participation was also anonymous. Schools used either passive or active permission forms at their discretion to fulfill requirements of the No Child Left Behind Act whereby parents must be provided with a means to opt out of their child's participation.

The final sample consisted of 214 schools, of which 178 participated, for a school participation rate of 83.2%. The survey yielded 18,866 completed student questionnaires out of a sample of 21,584 students for a student participation rate of 88%. The overall participation rate, the product of the school-level and student-level participation rates, was 73%.

---

\* Kreteks (or "clove cigarettes") are flavored cigarettes containing tobacco and clove extract.

A weighting factor was applied to each student record to adjust for non-response and for varying probabilities of selection. Weights were adjusted to ensure that the weighted proportions of students in each grade matched national population proportions. Final adjusted weights were scaled to ensure that the weighted count of students was equal to the total sample size.

The remainder of this report provides detailed information on the methodology used in the 2011 NYTS sample selection, data collection, and weighting of student response data.

- Section I: NYTS Sampling Design
- Section II: NYTS Sampling Methods
- Section III: NYTS Data Collection
- Section IV: Weighting of NYTS Response Data

## SECTION I — NYTS SAMPLING DESIGN

### 1.1 Sampling Design Overview

The objective of the NYTS sampling design was to support estimation of tobacco-related knowledge, attitudes, and behaviors<sup>†</sup> in a national population of public, Catholic, and other private school students enrolled in grades 6 through 12 in the United States. More specifically, the study was designed to produce national estimates at a 95% confidence level by school level (middle school and high school), by grade (6, 7, 8, 9, 10, 11, and 12), by sex, and by race/ethnicity for non-Hispanic white, non-Hispanic black, and Hispanic students<sup>‡</sup>. Additional estimates, such as cross-tabulations of grade by sex and of race/ethnicity by school level, were also supported; however, precision levels will vary considerably according to differences in sub-population sizes.

The universe for the study consisted of all public, Catholic, and other private school students enrolled in regular middle schools and high schools in grades 6 through 12 in the 50 states and the District of Columbia. Alternative schools, special education schools, Department of Defense operated schools, vocational schools that serve only pull-out populations, and students enrolled in regular schools unable to complete the questionnaire without special assistance were excluded.

The NYTS study is a continuation of the NYTS survey cycles that took place in 1999, 2000, 2002, 2004, 2006 and 2009. The NYTS survey system employs a repeat cross-sectional design to develop national estimates of tobacco use behaviors and exposure to pro- and anti-tobacco influences among students enrolled in grades 6-12. As in the 2009 survey cycle, the 2011 NYTS was coordinated with the national 2011 Youth Risk Behavior Survey (YRBS). As in the 2009 surveys, the samples for these two national surveys were drawn simultaneously for the 2011 survey cycle, and were designed to be non-overlapping samples of schools. The 2011 NYTS sampling design replicated key aspects of the 2009 NYTS.

The three-stage cluster sample was stratified by racial/ethnic composition and urban versus rural status at the first (primary) stage. Primary Sampling Units (PSUs) were defined as a county, a group of smaller counties, or a portion of a very large county. PSUs were classified as "urban" if they are in one of the 54 largest metropolitan statistical areas (MSAs) in the U.S.; otherwise, they were classified as "rural".\*\* Additional, implicit stratification was imposed by geography by sorting the PSU frame by state and by 5-digit Zip Code (within state). Within each stratum, a primary sampling unit (PSU) was randomly sampled without replacement at the first stage. In subsequent sampling stages, a probabilistic selection of schools and students was made from the sample PSUs. Black and Hispanic students were oversampled using a modified weighted measure of size (MOS) that increased the probability of selection of PSUs and schools with disproportionately high minority student enrollments. The MOS was adjusted to equalize the expected sample size by grade.

---

<sup>†</sup> See Section 3.1, survey instrument, for specific indicators measured.

<sup>‡</sup> Unless stated otherwise, subsequent use of the terms "white" and "black" pertain to non-Hispanic respondents only, and "Asian" is exclusive of Native Hawaiian/Other Pacific Islander (NH/IPI). Additional definitions are provided in Appendix A.<sup>9</sup>

\*\* Office of Management and Budget (OMB) Bulletin 09-01, November 20, 2008, "Metropolitan Statistical Areas."

The sampling stages may be summarized as follows.

*Stage I: Selection of PSUs*

At the first sampling stage, 82 PSUs were selected from sixteen strata with probability proportional to the total number of eligible students enrolled in all eligible schools located within a PSU.

*Stage II: Selection of Schools*

At the second sampling stage, two large schools were selected from each PSU. An additional 18 medium schools and 12 small schools were selected from subsample PSUs. Both were conducted with probability proportional to the total number of eligible students enrolled in a school. The three-size strata for schools are defined below:

- Small schools — those schools with 25 or fewer students in one or more of the eligible grades for the level.
- Medium schools — those schools with more than 25 students but fewer than 50 students in one or more of the eligible grades for the level.
- Large schools — those schools with at least 50 students at each grade.

*Stage III: Selection of Students*

At the third sampling stage, students were selected via whole classes whereby all students enrolled in any one selected class were by default chosen for participation. Classes were selected from course schedules provided by each school that agreed to participate. Schedules were constructed such that all eligible students were both represented and represented one time only. Classes were then randomly selected from each schedule using a fixed interval with a random start, with the start set within the initial interval. The interval used varied by school according to the total number of eligible students enrolled in a school, and thus how many students were anticipated to be enrolled on average in each class listed within a course schedule.

Exhibit 1-1 presents a summary of the sampling design features.

**Exhibit 1-1 Key Sampling Design Features**

Sampling Stage	Sampling Units	Sample Size	Stratification	Measure of Size
1	PSUs: Counties or groups of counties	82	Urban vs. non-urban (2 strata); Minority concentration (8 strata)	Aggregate school size in target grades
2	Schools	194 school selections: 164 large schools (2 per PSU), 18 medium schools and 12 small schools	Small, Medium and Large; High-school vs. middle-school	Weighted enrollment (increased for black, Hispanic groups)
3	Classes/ students	1 or 2 classes per grade (2 per grade in large high-minority schools)		

The design oversamples black and Hispanic students by using a modified MOS that increases the probability of selection of schools with disproportionately high minority student enrollments. In addition, more classes were selected in schools with high concentrations of students in these two minority groups. This refined method, described in the next section, was carefully calibrated to enhance the representation of minority students while also representing small and medium schools.

## 1.2 Designed Sample Sizes

The NYTS is designed to produce accurate estimation to within  $\pm 5\%$  at a 95% precision level for the following key subgroup estimates:

- Middle and high school estimates (school level)—middle school students in total (grades 6-8 combined) and high school students in total (grades 9-12 combined)
- Grade estimates - Individual grades—6, 7, 8, 9, 10, 11, and 12, separately
- Sex group estimates—males and females in total, by school level (male middle school students, female high school students, etc.), and by individual grade (6<sup>th</sup> grade males, 6<sup>th</sup> grade females, etc.)
- Racial group estimates (race/ethnicity)—in total and by school level (e.g., Hispanic middle school students)

The design was developed to meet the required precision levels for these subgroups assuming a design effect (DEFF) of 2.0. Under this scenario, the design was developed to generate at least 1,420 participating students by sex within grade, and at least that many by minority group at each school level. As shown below, the design was projected to yield at least 1,582 minority students at each level.

Exhibit 1-2 summarizes the sample sizes designed to generate approximately 2,840 participating students per grade; therefore, 1,420 by subgroup defined by sex and grade level. In this exhibit, the first two columns show the number of school selections and student selections per school while the last column shows the expected number of participating students. The numbers of large, medium and small schools were carefully calibrated to generate the required numbers of students overall, by school level and by grade.

Exhibit 1-3 below shows that the sample design was expected to yield 1,757 and 1,582 black students for high school and middle school, respectively. The exhibit also shows that higher numbers were expected for Hispanic students, 1,984 for high schools and 1,751 for middle schools than for black students.

For comparison purposes, Exhibit 1-4 presents the yields actually attained by race and ethnicity by school level in the 2011 NYTS and shows how the target projected sample sizes were exceeded for each subgroup. More importantly, all subgroup sample sizes exceeded the minimum subgroup sample size for the required precision levels (1,420).

**Exhibit 1-2 Sample Size Projections for Participating Students on the 2011 NYTS**

<b>School Type</b>	<b>Number of Schools Selected</b>	<b>Number of Students Selected per School</b>	<b>Expected Number of Participating Students Across Schools</b>
Large, High Minority High Schools	61	200	8,815
Large, Non-High Minority High Schools	21	100	1,517
Medium High Schools	9	100	620
Small High Schools	6	All available (assume 72)	312
Large, High Minority Middle Schools	61	150	6,611
Large, Non-High Minority Middle Schools	21	75	1,138
Medium Middle Schools	9	75	488
Small Middle Schools	6	All available (assume 54)	234
<b>TOTAL</b>	<b>194</b>		<b>19,765</b>

**Exhibit 1-3 Expected Minority Yields**

**Part I: High Schools**

<b>School Strata</b>	<b>Number of Schools</b>	<b>Expected Hispanics</b>	<b>Expected Blacks</b>
Large HS, High-Minority	61	1587	1410
Large HS, Non-High Minority	21	243	212
Medium HS	9	104	91
Small HS	6	50	44
<b>TOTAL</b>	<b>97</b>	<b>1984</b>	<b>1757</b>

**Part I: Middle Schools**

<b>School Strata</b>	<b>Number of Schools</b>	<b>Expected Hispanics</b>	<b>Expected Blacks</b>
Large MS, High-Minority	610	1454	1322
Large MS, Non-High Minority	21	182	159
Medium MS	9	78	68
Small MS	6	37	33
<b>TOTAL</b>	<b>97</b>	<b>1751</b>	<b>1582</b>

### Exhibit 1-4 Sample Yields for Black and Hispanic Students by School Level

Subgroup	Projected Participants	Actual Participants
Middle School Blacks	1,582	2448
Middle School Hispanics	1,751	2844
High School Blacks	1,757	2094
High School Hispanics	1,984	3217

Across the six previous cycles of the NYTS, the school participation has averaged 89%, with a low of 83%. Student participation has averaged 90% with a low of 88%. To be conservative, we assumed slightly lower values than the averages in developing the sample design for the 2011 NYTS: 85% for schools and 85% for students.

The large projected sample size permits analysis by individual grade and by sex without any special considerations in the sampling plan. Additionally, grade and sex subgroups both typically cut across schools; and sex subgroups typically cut across grades with males and females each constituting about half of the students selected. For example, design effects will be relatively small for subgroups that cut across schools; therefore, sex group estimates will have better precision than other groups than are less evenly dispersed across schools (e.g., racial/ethnic groups). Thus, confidence intervals will be within  $\pm 3\%$ . Comparatively note, though, that because the design expects to yield a greater number of completed surveys from high school students than from middle school students, overall estimates will be more precise at the high school level than those at the middle school level. Moreover, because within grade estimates by sex have slightly larger standard errors than those for estimates by grade alone, estimates of sex will be within  $\pm 5\%$ .

The next paragraphs discuss how the design is especially balanced to achieve precise estimates for subgroups defined by school level, grade, sex and race/ethnicity.

#### *Middle School and High School Estimates*

Estimates by school level are required to support separate analysis of students across middle school grades (6, 7, and 8) and high school grades (9, 10, 11, and 12). However, schools tend to vary in their grade structures, an inconsistency that compromises the ability to easily and efficiently cluster schools for sampling purposes in a manner that also uniformly divides students by grade. For example, 9<sup>th</sup> grade students are served by both grades 7-9 junior high schools and by grades 9-12 high schools. As a result, approaches that require artificially slicing and melding schools (e.g., removing 9<sup>th</sup> grade enrollment counts from a junior high school and deeming them eligible for selection at the high school level) were not considered viable options. Likewise, any approach that would require the creation of a rule classifying all schools as either “Middle School” or “High School” and then selecting one of each and every PSU would also result in some grades being found in both groupings. To counter such challenges, a simple sampling approach utilizing an implicit stratification scheme by which the sampling frame was sorted by grade was ultimately selected, and this allowed the sample to be drawn in a concentrated time period.

### *Grade Estimates*

Note first that NYTS estimates are typically not reported by grade level but rather by school level. Still, the design balances the sample sizes for grade level with at least 3,000 students targeted per grade, and thus also ensures that estimates at the grade level achieve the required precision levels. It is worth noting that this entails a larger student allocation to the high school stratum than to the middle school stratum as high schools have four grades versus three grades for middle schools.

### *Sex Group Estimates*

The large sample size permitted analysis by sex without any special considerations in the sampling plan. During the class selection process, frames of eligible classes from co-educational schools in which classrooms were segregated by sex (i.e., an all male or all female class) were avoided if at all possible.

### *Racial Group Estimates*

In order to support separate analysis of the data for white, black and Hispanic students, in total and by school level, adequate sample sizes were required for subgroups defined 1) by school level by racial grouping or 2) by sex by racial grouping. Sample sizes were not designed, however, to support detailed analyses by sex and school level within racial/ethnic subgroups (e.g., middle school Hispanic males).

The design used selection with Probability Proportional to Size (PPS) sampling methods with a weighted MOS that increased the chances of selecting PSUs and schools with relatively high minority enrollment, thereby increasing the likelihood of selecting classrooms with relatively high concentrations of minority students.

## **1.4 School Strata: Defining Small and Medium Schools**

Another design parameter was the number of students randomly selected per school. This determination required balancing schools' student enrollment size and the degree of burden that would be placed on schools if substantially more students were selected from those with higher enrollment counts. While there are a fair number of small schools in the U.S., they account for only a nominal proportion of the total eligible student population. Further, they are more expensive per student survey as they incur all of the same costs involved with recruiting a school and dispatching a field data collector, but produce smaller samples. Thus, for the 2011 NYTS, the stratification by school size was refined to include a medium school stratum as well as a small school stratum.

## **1.5 Actual Sample Yield and Participation Rates**

Participation rates for this survey were high, and generally comparable to participation rates achieved in the prior cycles of NYTS. Of the 214 selected schools, 178 (83.2%) participated in the 2011 NYTS. The school response rate was 83% and the student response rate was 88%, resulting in an overall response rate (the school response rate multiplied by the student response rate) of 73%.

## SECTION II — NYTS SAMPLING METHODS

The sampling frame included public, Catholic, and other private schools in the 50 states and the District of Columbia with at least one grade in the 6<sup>th</sup> through 12<sup>th</sup> grade range. The sampling process employed a three-stage stratified cluster design whereby PSUs, schools within PSUs, and classrooms (i.e., students) within schools were probabilistically selected without replacement. The sampling approach utilized Probability Proportional to Size (PPS) sampling methods, in which the probability of a particular PSU being selected was proportional to the total number of eligible students enrolled in all eligible schools located within the PSU, and the probability of a particular school being selected proportional to the total number of eligible students enrolled in the school. A weighted measure of size (MOS) that increased the chances of selecting PSUs and schools with relatively high minority enrollment was also used, thereby increasing the likelihood of selecting classrooms with relatively high concentrations of minority students.

### 2.1 Measure of Size (MOS)

The sampling approach utilized Probability Proportional to Size (PPS) sampling methods to achieve over-sampling of blacks and Hispanics. In PPS sampling, when the MOS is defined as the count of final-stage sampling units, and a fixed number of units are selected in the final stage, the result is an equal probability of selection for all members of the universe. For the NYTS, we approximate these conditions, and thus obtain a roughly self-weighting sample. This section describes the type of MOS employed for selecting PSUs and schools, including over-sampling of blacks and Hispanics. We note that in the 2009 cycle we moved to the use of a total enrollment, rather than a per-grade average for the MOS. This moved the actual performance of the sample close to the ideal of a self-weighting sample, resulting in a modest precision gain.

This oversampling method employs a function of the form  $r_h H + r_b B + r_o O$  where the  $r$ 's are the weighting factors for enrollment by three racial/ethnic groupings--Hispanic, black, and Other (H, B, and O, respectively). The function increases the chances of schools with relatively large minority enrollments entering the sample, and also increases the probability of selection for high-minority PSUs.

The effectiveness of a weighted MOS in achieving oversampling is dependent upon the distributions of blacks and Hispanics in schools. For example, if U.S. schools had identical percentages of minorities in every school, then the sample of students from any sample of schools would mirror the national percentages and use of a weighted MOS would fail to oversample blacks and Hispanics. We know this is not the case, however, as the distribution of high school students with respect to race and ethnicity follows that of the general population, where there is geographic clustering by race and ethnicity. As such, use of a weighted MOS, as has been employed in prior NYTS cycles, is an effective means of oversampling black and Hispanic students.

As in previous NYTS cycles, the coefficients for the MOS weighting function were developed using simulation studies that ensure that target sample sizes are met for the two racial/ethnic groups of analytic interest (black and Hispanic students). These simulations are performed in iterative steps that 1) adjust for racial/ethnic composition, and 2) equalize the expected sample size by grade. Prior to full implementation, sampling design parameters are fine-tuned through simulated sample draws. The simulations involve the repeated selection of independent samples of schools (e.g., 1,000 samples) in an iterative application of the entire sampling process up to the selection of schools.

The simulations verified that these parameters would also achieve the estimation goals of the NYTS, i.e., meet the target sample sizes by school level established for the two minority groups. As a result of these computations, the following weighting function was used for the MOS:

$$2.5 H + 3.0 B + O$$

Compared to NYTS cycles prior to 2009, this MOS used less extreme weighting of minorities to produce the same level of oversampling in terms of yields —leading to more precise estimates (smaller sampling errors.)

The MOS was used also to compute stratum sizes and PSU sizes. Assigning an aggregate MOS to PSU, the sample allocation oversamples PSU that have higher minority student concentrations.

## **2.2 Stage 1 – Selection of PSUs**

### *PSU Definition*

In defining PSUs, several issues are considered:

- a. Each PSU should be large enough to contain the requisite numbers of schools and students by grade, yet not so large as to be selected with near-certainty.
- b. Each PSU should be compact geographically so that field staff can go from school to school easily.
- c. There should be recent data available to characterize the PSUs.
- d. PSU definitions should be consistent with secondary sampling unit (school) definitions.

Generally, counties were equivalent to PSUs with two exceptions: (1) low population counties are combined to provide sufficient numbers of schools and students, and (2) counties that are very large may be split to avoid becoming certainty or near-certainty PSUs. County population figures were aggregated from school enrollment data for the grades of interest.

The 2011 NYTS frame of PSUs was constructed using new enrollment data and clustering methods that ensure that the compact PSUs have the correct number of schools and students. For the coordinated NYTS/YRBS surveys, we constructed new PSUs that would accommodate the selection of the two samples, where one of the two samples (NYTS) contains middle schools as well as high schools. The new PSUs were constructed to contain at least 4 middle schools and 5 high schools. The frame had 1,276 PSUs, 536 of which were comprised of one single county.

County population figures were aggregated from school enrollment data for the grades of interest. Enrollment data were obtained from the most recent Common Core of Data from the National Center for Education Statistics, which are merged on a rolling basis into the current school and school district data files of Quality Education Data, Inc. (formerly QED), acquired by MCH Strategic Data.<sup>§</sup>

### *Stratification and Selection of PSUs*

The PSUs were organized into 16 strata, based on urban/rural location (as defined above) and minority enrollment. The approach involves the computation of optimum stratum boundaries using the cumulative square root of “F” method developed by Dalenius and Hodges.<sup>11, 12</sup> The boundaries or cutoffs change as the frequency distribution (“F”) for the racial groupings change from one survey cycle to the next. These rules are summarized below.

- If the percentage of Hispanic students in the PSU exceeded the percentage of black students, then the PSU is classified as Hispanic. Otherwise it is classified as black.
- If the PSU is within one of the 54 largest MSA in the U.S. it is classified as 'Urban', otherwise it is classified as 'Rural'.
- Hispanic Urban and Hispanic Rural PSUs were classified into four density groupings depending upon the percentages of Hispanics in the PSU.
- Black Urban and black Rural PSUs were also classified into four groupings depending upon the percentages of blacks in the PSU.

### *Allocation of the PSU sample*

We designed and selected a sample of 82 PSUs. In order to stay as close as possible to maximum sample efficiency in terms of precision, the initial allocation was made proportional to student enrollment. Using simulations as in previous studies, we then made adjustments to the initial allocation to meet minority targets.

Exhibit 2-1 presents the allocation of the PSU sample to strata. Compared to previous cycles, this allocation is closer to proportional and therefore more efficient statistically; i.e., it leads to smaller variances and tighter confidence intervals.

---

<sup>§</sup> More information is available [at](http://www.mchdata.com/qed) Quality Education Data. <http://www.mchdata.com/qed>

**Exhibit 2-1 Stratum Definition and PSU Allocation to Strata**

<b>Predominant Minority</b>	<b>Urban/Rural</b>	<b>Density Group Number</b>	<b>Stratum Code</b>	<b>Student Population</b>	<b>Percent of Student Population</b>	<b>Number of Sample PSUs</b>
Black	Urban	1	BU1	1,793,003	11.0%	9
		2	BU2	1,143,917	7.0%	6
		3	BU3	485,659	3.0%	2
		4	BU4	174,403	1.1%	1
	Rural	1	BR1	2,482,218	15.2%	12
		2	BR2	1,142,065	7.0%	6
		3	BR3	719,469	4.4%	4
		4	BR4	239,154	1.5%	2
Hispanic	Urban	1	HU1	1,701,695	10.4%	8
		2	HU2	1,011,053	6.2%	5
		3	HU3	1,149,500	7.0%	6
		4	HU4	844,445	5.2%	4
	Rural	1	HR1	2,247,638	13.7%	11
		2	HR2	512,487	3.1%	2
		3	HR3	428,973	2.6%	2
		4	HR4	288,335	1.8%	2

*Selection of PSUs*

Within each first-stage stratum, the PSUs were sorted by five-digit zip code to attain a form of implicit geographic stratification. Implicit stratification, coupled with the probability proportional to size (PPS) sampling method described below, ensures broad geographic representation. With PPS sampling, the selection probability for each PSU is proportional to the PSU's MOS.

The following systematic sampling procedures were applied to the stratified frame to select a PPS sample of PSUs.

- Select 82 PSUs with a systematic random sampling method within each stratum. The method applies within each stratum a sampling interval computed as the sum of the measures of size for the PSUs in the stratum divided by the number of PSUs to be selected in the stratum.
- Subsample at random 9 of the sample PSUs for the medium school sample for each school level (middle school and high school)
- Subsample at random 6 of the sample PSUs for the small school sample for each school level (middle school and high school)

### 2.3 Stage 2 – Selection of Schools

Schools were stratified by school level—middle schools and high schools—and by size. Middle schools were those that contained any of grades 6, 7, or 8, and high schools were those that contained grades any of grades 9 through 12. Schools that contained a mix of high- and middle-school grades were split into two sampling units, one for each level. We constructed three school size strata-- small, medium and large—as defined in Section 1.1.

Schools were classified as "whole" for high schools if they have all high-school grades 9-12, and whole for middle schools if they had all grades 6 - 8. Otherwise, they were considered a "fragment" school. Fragment schools formed component schools that were linked with other schools (fragment or whole) to form a cluster school that has all four grades. This process is illustrated in Figure 1, where fragment school A is linked with whole school B, to form a cluster school, or Secondary Sampling Unit (SSU) XXX. We linked schools before sampling using an algorithm, developed for use in the national YRBS that links geographically proximate schools. Cluster schools were treated as a single school, or sampling unit, during sampling performed at the grade level as described below.



Figure 1 – Cluster School Construction & Grade Sampling for High Schools

In addition, for each school level, large schools were classified as high-minority or low-minority. High-minority schools were identified as those containing more than 2.0% of black or 6.0% Hispanic students. This threshold was set so that approximately 75% of the sample large schools would be selected to supply two sections (or classrooms) per grade, rather than one.

For large schools, one high school and one middle school were selected with PPS systematic sampling within a PSU. The schools were selected into the sample with probability proportional to the weighted MOS.

Small and medium schools were sampled independently from large schools; they were set in two separate strata that were sampled at lower rates. This approach was implemented by drawing a sub-sample of 12 PSUs for the sampling of small schools sampling and a subsample of 18 PSUs for medium school sampling at each level. Then one small school or medium school was selected in each sub-sampled PSU with probability proportional to weighted MOS.

#### *Replacement of Schools/School Systems*

We did not replace refusing school districts, schools, classes, or students. We allowed for school and student non-response by inflating the sample sizes to account for non-response. With this approach, all schools can be contacted in a coordinated recruitment effort, which is not possible for methods that allow for replacing schools.

## 2.4 Stage 3 – Selection of Grades, Classes and Students

### *Selection of Grades*

Except for cluster schools, all eligible grades were included in the class selection for each school. In cluster schools, grades were selected independently. One component school was selected to provide classes at each grade level, and grades within component schools were drawn with probability proportional to grade enrollment. This is illustrated in Figure 1, where grades 9 and 10 were drawn into the sample from component school QED ID (QID) B, and grades 11 and 12 were drawn into the sample from component school QID A. Thus, the SSU XXX is represented by two actual schools QID A and QID B, each supplying two of the four required grades.

### *Selection of Classes and Students*

Students were selected for participation by default via the selection of whole classes (e.g., all students enrolled in a selected class were eligible to take the survey). The frames from which classes were chosen were constructed such that eligible students had only one chance of being selected. However, at times the specific method of selecting classes varied from school to school according to how a school's class schedule was structured and implemented. Typically, classes were selected from a list of required core courses such as English, social studies, math, or science. Among middle school students, and among high school students in a few states, physical education and/or health also were considered core courses. In a small number of schools, however, it was extremely difficult to develop an appropriate frame using this particular approach. Therefore, in such schools, classes were selected by using a time of day (i.e., second period) when all eligible students were scheduled to attend a class of one kind or another as the frame, and randomly selecting from all classes held at this time. Lastly, in some schools, school homerooms were used as the frame for class selection. Doing so is not ideal, though, as relatively few schools hold homeroom of duration sufficient enough for conducting the survey.

In large schools, we selected an average of 1.75 classes per grade by selecting 2 classes per grade in 75% of the selected large schools, and one class per grade in the remaining schools. While all classes were selected in eligible grades in small schools, one class per grade was selected in medium schools.

All students in a selected class who could complete the survey without special assistance were considered eligible and offered the opportunity to participate in the survey. Refusing students were not replaced. Non-response at the student level was accounted for in the sample size using an average per class yield that assumed student response rates derived from historical experience with the NYTS.

## 2.5 Probabilities of Selection

This section describes the probabilities of selection associated with the various sampling stages including:

- Probability of selecting PSUs
- Probability of selecting schools
- Probability of selection of grades
- Probability of selecting classes and students

These probabilities provide the basis for the sampling weights discussed in Section IV, Weighting of NYTS Response Data.

The probability of selection for a student is the product of the probability of selection of the PSU, which is a group of schools, multiplied by the conditional probability of selecting the student's school, multiplied by the conditional probability of selecting the student's class. All students in a selected class are included in the sample.

### *Probability of Selecting PSUs*

If  $MOS_{klm}$  is the measure of size for school  $k$  in PSU  $l$  in stratum  $m$  and if  $K_m$  is the number of PSUs to be selected in stratum  $m$ , then  $P^P_{lm}$  is the probability of selection of PSU  $l$  in stratum  $m$ :

$$P^P_{lm} = K_m \left( \frac{MOS_{.lm}}{MOS_{..m}} \right)$$

As noted below, 12 of the 82 sample PSUs were sub-sampled for the sampling of small schools, with 6 PSUs sub-sampled for each school level. Similarly, 18 PSUs were sub-sampled for the sampling of medium schools with 9 for each school level. Thus, the sub-sample PSUs are assigned an additional sampling factor in their probability of selection for small schools (6/82) and for medium schools (9/82).

### *Probability of Selecting Schools*

The probability of selecting large school  $k$  in PSU  $l$  and stratum  $m$ ,  $P^{LS}_{klm}$ , was computed as follows:

$$P^{LS}_{klm} = \left( \frac{MOS_{klm}}{MOS_{.lm}} \right)$$

For small schools, one school was drawn from sub-sampled PSU, so the probability of selection of a small school,  $P_{klm}^{SS}$ , then becomes:

$$P_{klm}^{SS} = (6/82) \left( \frac{MOS_{klm}}{MOS_{.lm}} \right)$$

For medium schools, one school was drawn from sub-sampled PSU, so the probability of selection of a medium school,  $P_{klm}^{MS}$ , then becomes:

$$P_{klm}^{MS} = (9/82) \left( \frac{MOS_{klm}}{MOS_{.lm}} \right)$$

### *Probability of Selection of Grades*

For most schools in the sample all eligible grades were selected so that the probability of selecting a grade was 1.0. In some cases, however, a school unit was really a set of several schools combined in the sampling process because they lacked one or more eligible grades. These units may be called cluster schools. Some of the cluster schools contain more than one school that can be selected for a given grade. The selection of the school for that grade is made with PPS sampling. For example, suppose that there were two schools with 10<sup>th</sup> grades in a cluster, and their 10<sup>th</sup> grade enrollments were in a 2-to-1 ratio. In this way, the larger school in the cluster had twice the probability of yielding a 10<sup>th</sup> grade to the sample than the smaller school. The school selections (from the school cluster) at each grade level were made independently.

We denote by  $P_{jklm}^G$  the probability of selecting grade-j in school-k, in PSU-l, stratum-m.

### *Probability of Selecting Classes and Students*

A set of classes was identified for each school at each grade level such that every student in a given grade level was enrolled in exactly one of the classes in the set. For example, a required English course might be used. If the school's estimated black enrollment exceeded 40 percent or its estimated Hispanic enrollment exceeded 50 percent, then two classes were randomly selected, without replacement, from the list. Otherwise, one class was randomly selected. Selections were made at all eligible grade levels in the school.

The probability of selection of a class when there are  $C_{jklm}$  classes at grade j in school k, PSU l, stratum m is just  $1/C_{jklm}$  or  $2/C_{jklm}$  depending on whether 1 or 2 classes are taken in the school. All students in a selected class were chosen so the probability of selection of a student is the same as the class (i.e.  $1/C_{jklm}$  or  $2/C_{jklm}$ ). Note that the probability of student selection within a class does not vary by race, ethnicity or sex. We denote this probability as  $P_{ijklm}^C$  as the probability of selecting class-i in grade-j, school-k, PSU-l, stratum-m. Since every student in a selected class is also selected, the probability of selecting any student in class-i, grade-j, school-k, PSU-l, stratum-m, is also equal to  $P_{ijklm}^C$ .

If  $n_i$  is the number of PSUs to be selected from stratum  $i$ ,  $S_i$  is the size of stratum  $i$ , and  $S_{ij}$  is the size of  $PSU_i$  in stratum  $i$  (in all cases "size" refers to MOS), then the probability of selection of  $PSU_i$  is  $n_i S_{ij}/S_i$ . For the two schools to be selected in  $PSU_i$ ,  $S_{ijk}$  is the size of school  $k$  in  $PSU_j$  in stratum  $i$ , and then the conditional probability of selection of the school given the selection of the PSU is  $2S_{ijk}/S_{ij}$ . If  $C_{ijk}$  is the number of classes in school  $ijk$  and  $m$  is the number of classes to be selected, then the conditional probability of selection of a class is  $m/C_{ijk}$ . Since all students were selected, the conditional probability of selection of a student given the selection of the class is unity.

$$\left( \frac{n_i S_{ij}}{S_i} \right) \left( \frac{2S_{ijk}}{S_{ij}} \right) \left( \frac{m}{C_{ijk}} \right)$$

## **SECTION III — NYTS INSTRUMENT AND DATA COLLECTION**

### **3.1 Survey Instrument**

The NYTS collects data on key short-term, intermediate, and long-term tobacco prevention and control outcome indicators. The 2011 survey instrument included a total of 81 questions, with the first seven collecting student demographic information and the remaining measuring a comprehensive set of tobacco-related topics (Appendix B). Specifically, areas covered by the survey included prevalence of tobacco use, knowledge of and attitudes towards tobacco use, pro- and anti-tobacco media and advertising, minors' access to tobacco products, nicotine dependence, cessation attempts, exposure to secondhand smoke, and tobacco use prevention school curricula.

### **3.2 Recruitment and Scheduling of Schools**

The schools selected to participate in the 2011 NYTS fell across 34 different states. Recruitment began in May 2010 with calls to State Departments of Education and Health. Letters of support were obtained from various State agencies and used in mailings to districts and schools. A date for survey implementation was selected to optimize the efficiency of data collection while accommodating school schedules. In selecting a date, convenience to the school and its calendar were considered while also trying to schedule groups of schools from the same school district or PSU around the same time to facilitate efficient travel to and survey implementation within selected schools. Recruiters used an electronic calendar on a secure, shared drive to facilitate communication and to avoid scheduling two schools for the same data collector on the same day.

### **3.3 Mailings to Schools**

After schools had been recruited, classes selected, and a date scheduled, each school received a packet of pre-survey materials. These materials included all the information necessary to prepare the school for data collection. Teacher packets contained the parental permission forms that had to be given out to all students in the selected classes prior to data collection. The timing of these pre-survey packet mailings was determined in part by the type of permission form being used by the school. Passive parental permission forms, or forms returned only if the parents do not want their child to participate, were provided to students to give to their parents approximately one week prior to the scheduled date of data collection. Active parental permission forms, forms that must be returned with the parent's signature in order for the child to participate, were provided to students to give to their parents at least two weeks prior to the scheduled date of data collection. Passive consent is the primary way to receive consent from student's parents for participation in the NYTS and active consent was used in school districts that require active parental consent. Follow-up calls were made to the selected schools to answer any questions and to make sure materials were received and distributed to selected classes and students.

### **3.4 Hiring and Training Data Collectors**

Data collectors were recruited from a pool of previously trained data collectors as well as retired teachers' associations, school health networks, and a variety of health education listservs. Data collector training was conducted on February 2-5, 2011. Initially, data collectors observed everything they would have to say or do as "experts" performed it. Then they acquired these data collection skills through practice, demonstrated them to each other, and finally refined each other's performance through constructive feedback. Appendix C presents the training agenda from the 2011 NYTS Data Collector Training.

### **3.5 Management and Support of Data Collectors in the Field**

On a weekly basis, data collectors received mailings containing their assignments for the coming week, travel and logistics to get them where they needed to be, and their must-read weekly bulletin. Weekly bulletins underlined key performance issues, corrected misconceptions, provided consistent direction on any procedural changes, and kept everyone abreast of the latest must-have information. In addition to these mailings, boxes of survey supplies were sent to data collectors either to the data collector's home or hotel. These boxes contained all supplies necessary for completing the data collection, including questionnaires, data envelopes, field forms, and pencils. Data collectors were supplied with extra materials for emergency packs as well, which they carried with them at all times.

In addition to receiving multiple mailings, supervisors remained in close contact with the data collectors by phone and email, including daily contact during the first week and no less than twice-weekly contact thereafter. Phone calls were specifically used to review performance, provide reminders, and give emotional support. Further, the calls gave data collectors the opportunity to ask questions, provided feedback from schools, and discussed difficult or rewarding data collection experiences.

### **3.6 Survey Administration**

Survey administration in the schools began on February 8, 2011 immediately after data collector training and continued until June 17, 2011. Each data collector visited an average of 3 schools per week. While the details of each data collection varied, there were six core steps followed for every school: (1) Precontact call with the principal or lead contact prior to arrival at the school; (2) Entry meeting with the principal or lead contact; (3) Entry meeting with teacher or group of teachers prior to survey administration; (4) Survey administration; (5) Post-survey meeting with the teacher or teachers; and (6) Post-survey meeting with the principal or lead contact prior to leaving the school. Most survey administrations could be completed in one day, while at other times, due to the number of classes selected or alternating block schedules, the data collector needed to return for a second day. Procedures were designed to protect students' privacy by assuring that student participation was anonymous and voluntary. Students completed a pencil and paper, self-administered, scannable, questionnaire booklet.

### **3.7 Make-up Sessions**

If students in selected classes were eligible to participate but were absent on the day of the survey, a Make-Up List was completed. Copies of the Make-Up List were given to the applicable teacher and to the principal/lead contact; one also was kept by the data collector. Every effort was made to have the data collector return at least once to each visited school to conduct a make-up session if necessary. However, sometimes there was no opportunity for the data collector to return to a school prior to leaving for the next site; or, coming back while the data collector was still in the area was too soon for a make-up session. In these and other similar situations, teachers were asked to administer make-up surveys and send them back in the provided pre-paid business reply envelopes. These supplementary make-ups closely followed the National YRBS model for teacher-administered make-up sessions.

Student anonymity was protected throughout the make-up process, with no means of connecting a make-up survey with a particular student. There were no student-level identifiers on the questionnaire, and school-level identifiers remain confidential and have not been released. Once data collection was completed, all Make-Up Lists were destroyed.

The use of make-up sessions ensured comparability of data and protected respondent anonymity while at the same time maximizing response rates. After the completion of all data collection in a school, data collectors sent thank you letters to each of the participating teachers and the principal/lead contact. This allowed the data collector an opportunity to connect once more with the staff at the school, inquire as to the status of make-ups expected, and encourage further participation.

## SECTION IV — WEIGHTING OF NYTS RESPONSE DATA

This section describes the procedures used to weight the data collected in NYTS 2011. The process involved the steps outlined below:

- Sampling weights
- Non-response adjustments
- Weight trimming
- Post-stratification to national estimates of racial totals by grade, sex and school type

This section focuses on the development of the weights for the student response data. The final student level response data was weighted to reflect the initial probabilities of selection and non-response patterns, to mitigate large variations in sampling weights, and to post-stratify the data to known sampling frame characteristics.

### 4.1 Sampling Weights

The sampling weight attached to each student response is the inverse of the probability of selection for that student. This basic weight can be adjusted to compensate for non-response, to alleviate excess weight variation, and to match the weighted data to known control totals. A convenient way of computing the basic weight is by inverting the probabilities of selection at each stage, to derive a partial weight or stage weight. The stage weights are then multiplied together to form the overall weight.

#### 4.1.1 Conditional Student Weights

The conditional student weight is the student weight given the selection of the PSU, school and grade. This weight is the product of the inverse of the probability of selection, a non-response adjustment and a ratio adjustment to control to known school enrollment totals. These adjustments are explained further.

**Basic Weight.** If  $K_{klm}^C$  classes per grade were selected the conditional probability of selection of the student was  $K_{klm}^C / C_{jklm}$  and the corresponding weight is the inverse of this probability,  $W_1 = C_{jklm} / K_{klm}^C$

**Nonresponse Adjustment.** The conditional student weight is adjusted for non-response by multiplying the basic weight by the inverse of the proportion of students selected that responded to the survey,  $W_2 = W_1 \times f_{NR}$ , where:

$$f_{NR} = \frac{\sum_{\text{Select}} W_1}{\sum_{\text{Resp}} W_1} = W_1 \times \frac{n_{\text{select}}}{n_{\text{resp}}}$$

**Enrollment Ratio Adjustment.** Next, the nonresponse adjusted student weights are ratio adjusted to conform to known school enrollment totals for each grade and sex,  $W_3 = W_2 \times f_{\text{enroll}}$ , where:

$$f_{\text{enroll}} = \frac{N_{\text{enroll}}}{\sum_{\text{Resp}} W_2} = \frac{N_{\text{enroll}}}{W_2 \times n_{\text{resp}}}$$

We have shown this as a three-step process for illustration purposes. However, to simplify calculations, we combine these steps to calculate ratio-adjusted student weights directly. Using the adjustment factors above, it is clear that ratio-adjusted student weight is equal to the ratio of the enrollment total to the number of students responding,  $W_3 = N_{\text{enroll}}/n_{\text{resp}}$ .

We form adjustment cells by grade and sex within each school. To avoid any excessively large weights, we apply a collapsing algorithm for nonresponse cells. We collapse if the nonresponse adjustment for the cell exceeds a cap, which we've set to be equal to:

$$2 \times \text{enrollment}/\text{min}(\text{enrollment}, 10)$$

If exceeded, we initially collapse males and females within grade within the school. If the adjustment still exceeds the cap, we collapse across grades. Finally, if necessary, we collapse all grades within the school.

#### 4.1.2 School Sampling Weights

For large schools the partial school weight is the inverse of the probability of selection of the school given that the PSU was selected:

$$W_{klm}^{LS} = \left( \frac{MOS_{.lm}}{MOS_{klm}} \right) = \frac{1}{P_{klm}^{LS}}$$

For small schools the partial school weight is:

$$W_{klm}^{SS} = (82/6) \left( \frac{MOS_{.lm}}{MOS_{klm}} \right) = \frac{1}{P_{klm}^{SS}}$$

For medium schools the partial school weight is:

$$W_{klm}^{MS} = (82/9) \left( \frac{MOS_{.lm}}{MOS_{klm}} \right) = \frac{1}{P_{klm}^{MS}}$$

The overall weights for a given PSU, school and grade combination were the product of the adjusted PSU, school and grade level weights.

#### 4.1.3 Grade Sampling Weights

The partial weight for a grade, given the selection of the cluster school containing it, is simply the inverse of the probability of selection described in Chapter 2. In a non-cluster school the weight is 1.0. We denote the grade weight as  $W_{jklm}^G$ .

#### 4.1.4 PSU Sampling Weights

The weight of the PSU is the inverse of the probability of selection of that PSU:

$$W^P_{lm} = \frac{1}{K_m} \left( \frac{MOS_{.m}}{MOS_{lm}} \right) = \frac{1}{P^P_{lm}}$$

For small schools and medium schools, the PSU component of the weight was multiplied by the ratio of the number of PSUs to the number of PSUs selected to have a sample small school, or medium school; specifically, this ratio is 82/6 for small schools, and 82/9 for medium schools.

## 4.2 Non-response Adjustments

Let  $W^{T1}_{hijklm}$  be the product of the weights computed to this point, i.e.:

$$W^{T1}_{hijklm} = W^P_{lm} W^{LS}_{klm} W^G_{jklm} W^R_{hijklm}$$

This is the weight for students in large schools with a non-response/post-stratification adjustment made for non-responding students in participating schools. The expression for small school students is similar. This section describes how weights are adjusted for nonparticipation by entire schools, using strata as weighting classes. The adjustment factor for large schools in stratum m is:

$$A_m^{LS} = \frac{\sum_{k,l \in \text{schoolssampled}} W_{lm}^P W_{klm}^{LS} MOS_{klm}^{LS}}{\sum_{k,l \in \text{schoolswithrespondents}} W_{lm}^P W_{klm}^{LS} MOS_{klm}^{LS}}$$

The student weight, adjusted for non-response, is  $A_{lm} W^{T1}_{hijklm}$ .

To avoid very large weight adjustment factors, which may lead to variance increases, weighting classes combined the top two sampling strata in terms of minority concentrations. These weighting cells were created for computing non-response adjustments only – the collapsed strata not kept on the analytic file. Specifically, weighting cells combined the following pairs of strata: B3U and B4U; B3R and B4R; H3U and H4U; and H3R and H4R. School response rates by weighting class, and the resulting non-response adjustment factors, are detailed in Exhibit 4-1. Note that the weighting classes are defined using collapsed sampling strata.

### Exhibit 4-1 School Non-response Adjustment

School Level	Stratum (for Non-Response)	Sampled Schools	Responding Schools	Percent Responding	Non-response Adjustment
HS	BR1	15	15	100.0	1.0
HS	BR2	6	4	66.7	1.5
HS	BR3	6	5	83.3	1.1
HS	BU1	6	5	83.3	1.2
HS	BU2	9	6	66.7	1.4
HS	BU3	3	3	100.0	1.0
HS	HR1	12	9	75.0	1.3
HS	HR2	2	2	100.0	1.0
HS	HR3	4	4	100.0	1.0
HS	HU1	10	6	60.0	1.5
HS	HU2	5	2	40.0	2.3
HS	HU3	10	6	60.0	1.7
HS	Small/Med	12	8	70.0	
Total HS		100	75	75.0	
MS	BR1	14	12	85.7	1.1
MS	BR2	6	5	83.3	1.2
MS	BR3	7	7	100.0	1.0
MS	BU1	7	7	100.0	1.0
MS	BU2	12	11	91.7	1.1
MS	BU3	3	2	66.7	1.4
MS	HR1	13	12	92.3	1.1
MS	HR2	2	2	100.0	1.0
MS	HR3	5	5	100.0	1.0
MS	HU1	10	10	100.0	1.0
MS	HU2	5	3	60.0	1.4
MS	HU3	12	11	91.7	1.1
MS	Small/Med	18	16	88.9	
Total MS		114	103	90.4	
Gross Total		214	178	83.2	

### 4.3 Weight Trimming

Extreme variation in sampling weights can cause inflated sampling variances, and offset the precision gained from a well-designed sampling plan. One strategy to compensate for this is to trim extreme weights and distribute the trimmed weight among the untrimmed weights. The method we used is based on a similar procedure done for the National Assessment of Educational Progress (NAEP).<sup>9, 10</sup>

The trimming is an iterative procedure. During each iteration, an optimal weight,  $W_o$  is calculated from the sum of the squared weights in the sample. Then, each weight  $W_i$  is marked and trimmed if it exceeds that optimal weight. The trimmed weight is summed within grade and spread out proportionally over the unmarked cases in the grade. This process is repeated until little or no weight is being trimmed.

$W_{ok}$  is determined by the following:

$$W_{ok} = \left( c \sum_{k=1}^n \frac{w_k^2}{n} \right)^{\frac{1}{2}}$$

The constant  $c$  is arbitrary. Setting it to a low level will generate high levels of trimming; increasing it will reduce the level of trimming. For the current study,  $c$  has been set so that approximately 5% of the weight will be trimmed in the first iteration of the trimming algorithm.

The results of the first iteration of the trimming operation are summarized in Exhibit 4-2.

### Exhibit 4-2 Results of First Trimming Iteration

					Coefficient of Variation -CVs (in %)-		-Design Effects-	
Trimming Class	Number Cases	Trimming Factor	Total Weight	Percent Trimmed	After Trimming	Before Trimming	After Trimming	Before Trimming
BR106	252	10.2	440509.65	4.97108	96.067	116.711	1.91922	2.35673
BR107	290	7.9	419906.12	4.98287	95.753	112.793	1.9137	2.26784
BR108	293	6.8	415498.93	4.92763	86.278	101.848	1.74184	2.03376
BR109	306	4.9	558604.41	4.99921	66.343	82.828	1.4387	1.68381
BR110	304	3.8	554634.61	4.89222	64.935	77.901	1.42027	1.60486
BR111	282	4	515776.86	4.8624	62.883	75.648	1.39402	1.57023
BR112	275	3.8	532442.89	4.99185	63.239	77.459	1.39846	1.59781
BR206	208	3.1	273913.41	4.56887	56.757	66.811	1.32058	1.44422
BR207	242	2.7	455519.82	4.77921	62.062	69.955	1.38358	1.48735
BR208	265	4.1	434146.67	4.73023	71.954	83.664	1.51578	1.69733
BR209	87	1.5	262512.97	4.96012	17.869	29.972	1.03156	1.0888
BR210	151	4.6	238012.85	4.85898	63.757	77.732	1.4038	1.60023
BR211	135	7.1	195999.37	4.81389	90.249	106.376	1.80845	2.12321
BR212	194	7.5	192101.54	4.98473	72.978	92.557	1.52984	1.85226
BR306	159	1.4	136128.94	4.24554	21.697	29.86	1.04678	1.0886
BR307	153	1.1	140981.29	4.92728	6.622	23.73	1.00436	1.05594
BR308	165	1.3	130981.23	3.80828	22.811	28.154	1.05172	1.07878
BR309	169	1.6	210697.52	3.87391	25.558	32.576	1.06493	1.10549
BR310	173	2.3	198794.8	4.40588	50.986	59.315	1.25845	1.34979
BR311	170	5.3	181343.52	4.92886	68.174	83.354	1.46204	1.69071
BR312	156	1.7	181353.72	4.91281	46.224	53.879	1.21229	1.28844
BR406	72	3	30978.76	4.99362	53.449	64.414	1.28172	1.40916
BR407	95	1.4	42169.97	4.9022	17.349	26.614	1.02978	1.07009
BR408	101	1.1	42171.11	3.22881	10.588	17.863	1.0111	1.03159
BR409	33	1	26538.21	3.04625	0.175	6.434	1	1.00401
BR410	33	1	26390.78	4.15593	0.396	8.841	1.00002	1.00758
BR411	41	1.1	19977.38	4.69625	6.038	15.55	1.00356	1.02359
BR412	48	2.1	22999.78	4.88312	44.21	54.856	1.19138	1.29465
BU106	226	3.2	393272.93	4.86954	52.322	66.43	1.27255	1.43935
BU107	274	4.9	420844.39	4.8567	54.833	72.666	1.29956	1.52611
BU108	314	2	401120.45	4.98235	34.425	46.964	1.11813	1.21986
BU109	128	1.9	478276.29	4.8341	38.785	47.193	1.14926	1.22097
BU110	142	2	449618.31	4.68799	41.678	52.153	1.17248	1.27008
BU111	161	1.7	418996.34	4.4073	35.458	42.589	1.12495	1.18025
BU112	152	1.6	404999.59	4.65367	30.973	38.259	1.0953	1.14541
BU206	398	14.3	781268.77	4.94974	253.354	264.615	7.4027	7.98451
BU207	402	13.7	516997.8	4.96931	191.862	206.939	4.67196	5.27174
BU208	353	13.1	480314.97	4.88957	213.125	226.283	5.52937	6.10588
BU209	282	5	443072.02	4.76753	99.117	111.045	1.97893	2.22872
BU210	292	5.1	479376.9	4.85613	73.688	89.349	1.54113	1.7956
BU211	278	5.2	444358.52	4.7686	92.331	105.114	1.84944	2.10091
BU212	277	6.5	385918.17	4.97636	82.695	98.711	1.68139	1.97088
BU306	78	1.3	74640.41	4.55377	27.512	32.418	1.07472	1.10375
BU307	61	3.9	64026.27	4.90416	65.098	77.805	1.41683	1.59545
BU308	81	1.4	74460.53	4.44779	33.557	39.474	1.11121	1.1539
BU309	87	1.5	99062.84	4.78457	25.185	32.859	1.0627	1.10673
BU310	103	2.3	95613.98	4.5436	33.191	44.209	1.1091	1.19355
BU311	124	1.3	81266.11	4.8873	24.582	31.638	1.05994	1.09929
BU312	74	1.3	49351.36	4.45635	19.273	26.456	1.03664	1.06904

					Coefficient of Variation -CVs (in %)-		-Design Effects-	
Trimming Class	Number Cases	Trimming Factor	Total Weight	Percent Trimmed	After Trimming	Before Trimming	After Trimming	Before Trimming
BU406	31	1	18139.92	2.12905	0.088	4.437	1	1.0019
BU407	36	1	17777.12	2.3243	0.115	4.829	1	1.00227
BU408	46	1.2	50903.87	3.61037	9.366	16.779	1.00858	1.02754
BU409	35	1.5	52184.47	4.3449	23.782	32.781	1.05494	1.10439
BU410	46	1	33402.33	2.22372	0.096	4.61	1	1.00208
BU411	27	1.5	30520.98	4.58204	22.113	31.782	1.04709	1.09727
BU412	311	3.5	331738.07	4.92981	49.895	65.407	1.24815	1.42644
HR106	349	4.1	360659.68	4.90252	64.886	81.192	1.41981	1.65732
HR107	326	2.2	354950.15	4.73517	48.394	58.008	1.23348	1.33546
HR108	281	2.6	482993.54	4.66045	46.832	58.335	1.21854	1.33909
HR109	287	8.3	414696.07	4.87293	84.291	105.47	1.70802	2.10851
HR110	282	2.5	413108.48	4.83774	51.504	62.232	1.26433	1.38591
HR111	292	4.1	383052.15	4.98164	69.293	82.662	1.47851	1.68096
HR112	98	2.8	519296.91	3.5009	113.082	120.466	2.2657	2.43639
HR206	157	4.9	391304.65	4.89719	142.227	148.977	3.00997	3.20527
HR207	141	3	372481.53	4.18391	102.391	111.549	2.04095	2.2355
HR208	129	2.5	138393.38	4.8601	43.173	53.435	1.18494	1.28332
HR209	110	1.4	123321.95	4.97267	26.534	34.042	1.06977	1.11483
HR210	86	1.3	113759.19	4.80249	14.964	25.225	1.02213	1.06289
HR211	171	3.4	119611.08	4.74142	68.199	77.361	1.4624	1.59497
HR212	47	1.4	68522.14	4.69969	23.293	30.316	1.0531	1.08995
HR306	57	1.2	69979.67	2.91486	15.809	19.732	1.02455	1.03825
HR307	61	1.5	68212.09	3.96733	29.362	34.698	1.0848	1.11842
HR308	93	2.3	113738.99	4.63744	56.179	65.925	1.31221	1.42994
HR309	84	1.9	106546.48	4.45163	50.645	55.762	1.25344	1.30724
HR310	87	3.4	115204.2	4.59301	65.886	75.436	1.42911	1.56252
HR311	88	2.1	116022.42	4.38178	45.684	52.435	1.20633	1.27182
HR312	106	2.1	25528.28	4.55265	65.021	69.863	1.41878	1.48347
HR406	62	1	74404.39	4.83014	5.319	13.717	1.00278	1.01851
HR407	78	2.5	73128.65	4.37142	54.547	62.152	1.29372	1.38133
HR408	102	1	71666.78	3.22446	1.8	11.78	1.00032	1.01374
HR409	74	2	64115.01	4.50302	37.185	45.586	1.1364	1.205
HR410	100	1.2	67218.43	3.64936	16.29	24.133	1.02627	1.05766
HR411	97	1.1	63458.23	2.28899	12.41	15.731	1.01524	1.02449
HR412	281	6.1	527698.87	4.99139	76.461	96.211	1.58255	1.92236
HU106	281	2.6	556096.71	4.83174	39.81	52.693	1.15792	1.27667
HU107	384	2.1	566679.43	4.35368	54.779	60.859	1.29929	1.36941
HU108	330	3.2	591717.4	4.87095	51.394	64.209	1.26333	1.41103
HU109	220	2.2	437405.4	4.8674	39.273	49.15	1.15354	1.24048
HU110	169	1.8	344918.02	4.77462	31.372	40.396	1.09784	1.16222
HU111	166	2.4	347448.52	4.7886	52.445	60.812	1.27339	1.36758
HU112	100	4.4	286920.55	4.92754	74.064	86.159	1.54307	1.73491
HU206	190	8.3	333738.69	4.87696	94.5	112.791	1.88833	2.26549
HU207	204	3	333822.56	4.61837	41.314	54.818	1.16985	1.29903
HU208	126	5.1	351013.06	4.88734	76.18	89.378	1.57574	1.79251
HU209	78	1.4	216107.05	3.33243	21.709	27.634	1.04652	1.07538
HU210	153	4.4	352374.85	4.6375	77.965	91.508	1.60388	1.8319
HU211	152	2.7	268780.03	4.58305	37.825	50.014	1.14213	1.2485
HU212	233	3.7	214710.09	4.91531	61.663	73.276	1.37861	1.53464
HU306	255	3	222358.84	4.59507	76.565	84.036	1.58391	1.70343
HU307	197	1.5	230868.47	4.64561	51.064	54.496	1.25943	1.29547
HU308	212	1.6	299148.11	3.73911	26.233	33.732	1.06849	1.11325
HU309	201	1.5	286036.47	4.67607	21.477	30.789	1.0459	1.09433
HU310	211	1.4	250248.36	4.46236	15.506	27.057	1.02393	1.07286

					Coefficient of Variation -CVs (in %)-		-Design Effects-	
Trimming Class	Number Cases	Trimming Factor	Total Weight	Percent Trimmed	After Trimming	Before Trimming	After Trimming	Before Trimming
HU311	249	1.3	250419.76	4.912	12.417	24.309	1.01536	1.05886
HU312	181	3	222532.44	4.93549	62.066	71.415	1.38309	1.50719
HU406	185	2.5	204348.75	4.97739	26	48.461	1.06723	1.23358
HU407	202	1.4	192037.63	4.33487	28.772	34.829	1.08237	1.1207
HU408	100	1.7	180308.16	4.89923	16.972	30.039	1.02852	1.08933
HU409	73	2.2	155530.67	4.65612	28.982	40.646	1.08284	1.16294
HU410	106	1.2	137292.76	4.46459	12.772	20.088	1.01616	1.03997
HU411	111	1.3	151527.05	3.85442	25.924	30.341	1.0666	1.09123
HU412	252	10.2	440509.65	4.97108	96.067	116.711	1.91922	2.35673

Let  $W_{ik}$  and  $W_{ok}$  be the weight for the  $i$ th case and the optimum weight for the  $k$ th iteration, respectively, and define  $T_{ik}$  as 1 if  $W_{ik}$  is greater than or equal to  $W_{ok}$ , and zero otherwise. Then the trimmed weight for the  $k + 1$  iteration is defined as follows:

$$W_{i,k+1} = \begin{cases} W_{ok} & \text{if } W_{ik} \geq W_{ok} \\ \frac{\sum_{i=1}^n W_{ik} \left( 1 - \frac{t_{ik} \times W_{ok}}{W_{ik}} \right)}{\sum_{i=1}^n W_{ik} (1 - t_{ik})} & \text{if } W_{ik} < W_{ok} \end{cases}$$

#### 4.4 Post-stratification to national estimates of racial/ethnic totals by grade, sex and school type

To obtain accurate counts of high school students in schools considered eligible for the NYTS by grade, sex, and race/ethnicity for use in post-stratification, we turned to two school universe surveys conducted by the National Center for Education Statistics (NCES)\*\*. Raw school level data files were downloaded and processed to mirror eligibility requirements imposed on the sampling frame.

National estimates of racial/ethnic percentages were obtained from the two sources. Private schools enrollments by grade and five racial/ethnic groups were obtained from the Private School Universe Survey (PSS), and public school enrollments by grade, sex, and five racial/ethnic categories were obtained from the Common Core of Data (CCD), both produced by the National Center for Education Statistics (NCES). These databases were combined to produce the enrollments for all schools, and to develop population percentages to use as controls in the post-stratification step.

Specifically, population control totals for public school enrollments were taken from the most recent NCES Common Core of Data (CCD) Public Elementary/Secondary School Universe Survey (2009-10). Records for special education, vocational, and other/alternative schools were deleted prior to computing control totals. Control totals for private school enrollments were taken from the NCES Private School Universe Survey (PSS), School Year 2007-08 (most recent PSS data); this file was also restricted to “regular” schools.

\*\* Common Core of Data, National Center for Education Statistics <http://nces.ed.gov/ccd/>. School Year 2009-10.

Exhibit 4-3 gives counts of schools and students by grade for public and non-public schools based on both the raw file and the resulting set of eligible schools. The latter set of student totals was used as control totals for the post-stratification adjustments described next.

**Exhibit 4-3 Counts by Schools and Students by Type and Eligibility Status**

Type	Grade	Raw (all)		Eligible	
		Schools	Students	Schools	Students
Public	6	38,633	3,684,782	35,323	3,661,585
	7	31,482	3,684,685	27,404	3,654,283
	8	31,979	3,690,588	27,579	3,650,623
	9	27,401	4,115,243	20,520	3,995,571
	10	26,602	3,842,645	19,421	3,708,200
	11	26,416	3,568,333	19,178	3,411,507
	12	26,240	3,447,097	19,016	3,237,649
	Total		26,033,373		25,319,418
Private	6	17,574	342,675	15,541	324,372
	7	16,560	339,249	14,662	320,914
	8	16,415	337,350	14,469	317,943
	9	8,513	312,502	6,863	290,626
	10	8,124	303,841	6,525	281,885
	11	7,771	294,436	6,233	273,818
	12	7,551	286,063	6,096	266,841
	Total		2,216,116		2,076,399
Total	6	56,207	4,027,457	50,864	3,985,957
	7	48,042	4,023,934	42,066	3,975,197
	8	48,394	4,027,938	42,048	3,968,566
	9	35,914	4,427,745	27,383	4,286,197
	10	34,726	4,146,486	25,946	3,990,085
	11	34,187	3,862,769	25,411	3,685,325
	12	33,791	3,733,160	25,112	3,504,490
	Total		28,249,489		27,395,817

The weighted data were then post-stratified to the joint national distribution of school type, grade, sex and race/ethnicity. Exhibit 4-4 presents the post-stratification cells as well as the calculations involved in the post-stratification process. Within the Private school adjustment cells, sex was omitted, as enrollments by sex were not available for these schools. This is indicated by a “Combined” sex in Exhibit 4-4. Also within private schools, the racial/ethnic groups were collapsed to preclude small numbers of students in the adjustment classes. For the public schools, five racial/ethnic categories were used: white, black, Hispanic, Asian/ Pacific Islander, and Native American.

Given a national estimate of student counts  $R_a$  and a weighted response total of  $P_a$  for post-stratification adjustment class “a”, the post-stratification factor was the ratio of  $R_a$  to  $P_a$ . Exhibit 4-3 gives the population control totals used in post-stratification adjustments side by side with the sum of the weights in each post-stratum cell, as well as the adjustment factors calculated as the ratio of these two totals. In other words, the adjustments in column G in this exhibit are computed as  $E/F$ , control total for the cell divided by the weight estimate in the cell.

The control population total matches the sum of the adjusted weights. (The match is approximate due to rounding.) Note that in Exhibit 4-4, column F contains a total representing the sum of unadjusted weights, prior to post-stratification. Following post-stratification, adjusted weights sum to the control population totals.

Because estimates are typically reported separately for middle schools and high schools, the weights were adjusted separately for both subpopulations.

### Exhibit 4-4 Post-Stratification Adjustments

School Type	Grade	Race/Hispanic Origin	Sex	(E) Control Total	(F) Weighted Estimate	No. of Cases	(G) Post- stratification Adjustment
Private	6	Combined	Combined	324,372	121,373	104	2.67252
Private	7	Combined	Combined	320,914	359,699	221	0.89217
Private	8	Combined	Combined	317,943	365,680	111	0.86946
Private	9	Combined	Combined	290,626	301,740	247	0.96317
Private	10	Combined	Combined	281,885	298,812	156	0.94335
Private	11	Combined	Combined	273,818	320,417	189	0.85457
Private	12	Combined	Combined	266,841	215,625	128	1.23753
Public	6	Asian and Pacific Islander	Female	88,031	76,177	72	1.15561
Public	6	Black	Female	299,004	335,250	302	0.89188
Public	6	Hispanic	Female	410,404	574,759	440	0.71404
Public	6	Native American	Female	23,359	70,744	49	0.3302
Public	6	White	Female	965,646	933,933	641	1.03396
Public	6	Asian and Pacific Islander	Male	90,804	67,288	80	1.34949
Public	6	Black	Male	309,461	277,351	286	1.11578
Public	6	Hispanic	Male	427,504	718,582	416	0.59493
Public	6	Native American	Male	24,130	112,369	49	0.21474
Public	6	White	Male	1,022,069	1,039,834	714	0.98292
Public	7	Asian and Pacific Islander	Female	88,352	85,744	85	1.03042
Public	7	Black	Female	296,234	279,990	304	1.05802
Public	7	Hispanic	Female	405,764	618,844	478	0.65568
Public	7	Native American	Female	23,126	86,921	47	0.26606
Public	7	White	Female	970,030	955,719	686	1.01497
Public	7	Asian and Pacific Islander	Male	92,484	96,353	97	0.95985
Public	7	Black	Male	305,186	321,414	265	0.94951
Public	7	Hispanic	Male	423,195	546,283	499	0.77468
Public	7	Native American	Male	23,480	126,773	30	0.18522
Public	7	White	Male	1,026,280	813,739	717	1.26119
Public	8	Asian and Pacific Islander	Female	88,917	76,058	57	1.16907
Public	8	Black	Female	295,211	298,672	71	0.98841
Public	8	Hispanic	Female	397,892	599,360	88	0.66386
Public	8	Native American	Female	22,714	128,502	79	0.17676
Public	8	White	Female	981,228	884,781	62	1.10901
Public	8	Asian and Pacific Islander	Male	92,355	75,501	124	1.22322
Public	8	Black	Male	300,572	343,401	95	0.87528
Public	8	Hispanic	Male	411,469	492,816	88	0.83493
Public	8	Native American	Male	23,200	53,042	52	0.43738
Public	8	White	Male	1,036,937	870,839	295	1.19073
Public	9	Asian and Pacific Islander	Female	92,403	114,106	439	0.8098
Public	9	Black	Female	343,631	270,610	41	1.26984
Public	9	Hispanic	Female	428,546	634,183	450	0.67574
Public	9	Native American	Female	24,761	57,980	47	0.42706
Public	9	White	Female	1,045,606	915,846	294	1.14168
Public	9	Asian and Pacific Islander	Male	98,280	89,401	511	1.09931
Public	9	Black	Male	368,351	296,933	58	1.24052
Public	9	Hispanic	Male	459,163	660,890	492	0.69476
Public	9	Native American	Male	26,418	66,248	57	0.39877

School Type	Grade	Race/Hispanic Origin	Sex	(E) Control Total	(F) Weighted Estimate	No. of Cases	(G) Post-stratification Adjustment
Public	9	White	Male	1,108,382	950,710	323	1.16585
Public	10	Asian and Pacific Islander	Female	89,948	85,689	538	1.0497
Public	10	Black	Female	312,344	232,175	53	1.3453
Public	10	Hispanic	Female	378,513	521,951	601	0.72519
Public	10	Native American	Female	23,208	54,336	74	0.42713
Public	10	White	Female	1,019,269	897,439	334	1.13575
Public	10	Asian and Pacific Islander	Male	96,265	97,821	468	0.98409
Public	10	Black	Male	313,968	258,110	58	1.21641
Public	10	Hispanic	Male	394,255	518,334	490	0.76062
Public	10	Native American	Male	23,942	55,146	64	0.43415
Public	10	White	Male	1,056,460	878,573	336	1.20247
Public	11	Asian and Pacific Islander	Female	86,591	75,440	534	1.14781
Public	11	Black	Female	279,248	206,772	61	1.35051
Public	11	Hispanic	Female	332,290	481,375	587	0.69029
Public	11	Native American	Female	20,837	40,942	66	0.50893
Public	11	White	Female	976,283	832,360	357	1.17291
Public	11	Asian and Pacific Islander	Male	91,560	91,909	472	0.9962
Public	11	Black	Male	263,654	233,702	45	1.12816
Public	11	Hispanic	Male	334,879	513,862	600	0.65169
Public	11	Native American	Male	20,943	39,544	64	0.5296
Public	11	White	Male	1,005,201	848,921	207	1.18409
Public	12	Asian and Pacific Islander	Female	84,175	56,636	455	1.48625
Public	12	Black	Female	259,755	213,693	33	1.21555
Public	12	Hispanic	Female	298,664	470,628	499	0.63461
Public	12	Native American	Female	20,335	44,352	55	0.45848
Public	12	White	Female	958,319	825,484	207	1.16092
Public	12	Asian and Pacific Islander	Male	87,428	78,155	435	1.11865
Public	12	Black	Male	235,920	194,349	33	1.2139
Public	12	Hispanic	Male	290,431	505,220	461	0.57486
Public	12	Native American	Male	19,955	34,331	51	0.58126
Public	12	White	Male	982,622	861,534	195	1.14055

#### 4.5 Analysis Strata and Variance Estimation

Sampling variances for complex sampling designs can be estimated using one of several methods, including linearized estimators and balanced repeated replication. These methods may be implemented with a variety of software packages, including SUDAAN and WesVar, and with Stata and SAS using special sample survey procedures (such as Proc SurveyMeans in SAS Version 8). The 2011 NYTS data were prepared for estimating variances using the method of linearized estimators. Because estimates are typically reported separately for middle schools and high schools, analysis strata need to ensure that each stratum has two or more PSUs for variance estimation within each subpopulation (middle schools and high schools separately).

Exhibit 4-5 displays the correspondence between the sampling strata and the analysis strata. Stratum codes used in sampling and weighting were converted to a numeric “analysis stratum” code for use in SUDAAN. All strata but one had at least two PSUs. Stratum BU3 was combined with BU4 to result in an analysis stratum with at least two PSUs (analysis stratum coded 113). In addition, and for the same reason, strata BR4 and BR3 were combined for middle schools under the

same analysis stratum, 103. Thus, the analytic file contains 15 analysis strata for high schools, and 14 for middle schools.

#### Exhibit 4-5 Sampling and Analysis Stratum Coding Schemes

High Black		High Hispanic	
Sampling Stratum Code	Analysis Stratum Code	Sampling Stratum Code	Analysis Stratum Code
BU1	111	HU1	211
BR1	101	HR1	201
BU2	112	HU2	212
BR2	102	HR2	202
BU3	113	HU3	213
BR3	103	HR3	203
BU4	113	HU4	214
BR4	104(MS)/103(MS)	HR4	204

Exhibit 4-6 presents key survey estimates and their sampling errors estimated using Taylor series linearization methods which are usually employed by NYTS data analysts, and implemented with SUDAAN or similar software (e.g., SAS Proc SurveyMeans). Specifically, the exhibit presents standard errors for prevalence estimates of current use and ever use of tobacco products separately for high schools (4-6a) and middle schools (4-6b).

**Exhibit 4-6a Ever Use Estimates for High School Students**

Ever Use – High School			Sex		Race/ethnicity		
Product	Estimate	Overall	Female	Male	Whites	Blacks	Hispanics
Any tobacco	Prevalence	47.46	42.32	52.45	47.28	44.72	52.38
	Standard Error	1.64	1.80	1.75	2.16	2.72	1.71
Cigarettes	Prevalence	39.60	37.00	42.21	38.54	36.95	46.52
	Standard Error	1.78	2.02	1.82	2.30	3.01	1.73
Cigar	Prevalence	28.27	21.49	35.00	29.80	24.14	29.22
	Standard Error	1.10	1.05	1.35	1.32	1.81	1.67
SLT	Prevalence	15.25	6.22	24.07	19.39	6.74	10.38
	Standard Error	1.40	0.73	2.20	1.74	1.27	1.68
Pipe	Prevalence	9.84	6.96	12.65	9.84	5.64	12.94
	Standard Error	0.57	0.46	0.88	0.79	0.74	0.95
Bidis	Prevalence	4.31	2.86	5.69	4.13	2.63	5.90
	Standard Error	0.31	0.41	0.35	0.44	0.48	0.61
Kreteks	Prevalence	4.25	2.61	5.83	4.93	1.87	3.84
	Standard Error	0.35	0.41	0.46	0.50	0.46	0.47

**Exhibit 4-6b Ever Use Estimates for Middle School Students**

Ever Use – Middle School			Sex		Race/ethnicity		
Product	Estimate	Overall	Female	Male	Whites	Blacks	Hispanics
Any tobacco	Prevalence	21.27	18.76	23.57	18.26	26.68	27.55
	Standard Error	0.93	1.32	0.97	1.17	2.01	0.97
Cigarettes	Prevalence	16.20	15.33	16.96	13.26	20.87	22.50
	Standard Error	0.96	1.30	1.01	1.28	1.76	1.04
Cigar	Prevalence	9.32	6.90	11.67	7.17	12.49	14.72
	Standard Error	0.57	0.67	0.68	0.51	1.32	0.82
SLT	Prevalence	5.51	3.52	7.34	5.78	3.01	6.51
	Standard Error	0.46	0.44	0.65	0.53	0.65	0.67
Pipe	Prevalence	4.39	3.45	5.31	3.32	2.49	9.05
	Standard Error	0.45	0.54	0.51	0.46	0.36	0.75
Bidis	Prevalence	2.41	2.11	2.72	1.80	2.61	4.47
	Standard Error	0.20	0.32	0.30	0.27	0.42	0.53
Kreteks	Prevalence	1.75	1.55	1.95	1.44	1.09	3.14
	Standard Error	0.21	0.31	0.29	0.27	0.29	0.50

Example specifications for applying the method with SUDAAN are provided below for computing prevalence. Two types of analyses are presented. The first analysis considers the two subpopulations defined by school level as domains defined by the variable for school type (SCHOOLTY). These were the specifications used for producing the estimates above (Exhibit 4-6). The first example below, restricted to ever use estimates, also matches the variables used in that analysis. The second analysis considers two data sets defined by the two school levels, middle school and high school. The second example includes current use as well as ever use of tobacco, cigarettes and cigars.

### Example 1 Using subpopulations to estimate ever use by school level

```
PROC DESCRIPT DATA="C:\NYTS2011.ssd" FILETYPE= SAS
DESIGN=WR;
NEST NSTRATUM NPSU/MISSUNIT;
WEIGHT FINWGT;
SUBPOPN SCHOOLTY=2;
VAR   ETOB ECIGT ECIGAR ESLT EPIPE EBIDIS EKRETEKS;
CATLEVEL 1 1 1 1 1 1 1;
TITLE "NYTS 2011 High Schools, Overall";
SETENV LABWIDTH = 30 LINESIZE=120 PAGESIZE=59 COLWIDTH=11;
PRINT NSUM="N" TOTAL="n" PERCENT=" %" SEPERCENT="SE"
      / NSUMFMT=F8.0 TOTALFMT=F10.1 PERCENTFMT=F6.2 SEPERCENTFMT=F6.2
STYLE=NCHS;
```

### Example 2 Using two data sets to estimate prevalence by school level

```
1 PROC DESCRIPT
DATA="C:\nyts09_sud_hs.ssd"
  FILETYPE = SAS DESIGN=WR;
2 NEST NSTRATUM NPSU/MISSUNIT;
3 WEIGHT FINWGT;
4 VAR   ETOB ECIGT ECIGAR CTOB CCIGT CCIGAR;
5 CATLEVEL 1 1 1 1 1 1;
```

## References

1. CDC. *Best Practices for comprehensive tobacco control programs*. Atlanta, GA: US Department of Health and Human Services, Public Health Service, CDC; 1999.
2. MacDonald G, Starr G, Schooley M, Yee SL, Klimowski K, Turner K. *Introduction to program evaluation for comprehensive tobacco control programs*. Atlanta, GA: US Department of Health and Human Services, CDC; 2001.
3. U.S. Department of Health and Human Services (2010). *Healthy People 2020*. Washington, DC: U.S. Department of Health and Human Services. Available at: <http://www.healthypeople.gov/2020/about/default.aspx>
4. CDC. *Tobacco use among middle and high school students—United States, 1999*. MMWR 2000;49(10):49–53.
5. CDC. *Youth tobacco surveillance—United States, 2000*. MMWR 2001;50(SS4).
6. CDC. *Youth tobacco surveillance—United States, 2002*. MMWR 2006; 55(SS3).
7. CDC. *Tobacco use, access, and exposure to tobacco in media among middle and high school students—United States, 2004*. MMWR 2005;54(12):297-301.
8. Ryan H, Arrazola R. *2006 National Youth Tobacco Survey and key prevalence indicators*. [http://www.cdc.gov/tobacco/data\\_statistics/surveys/NYTS/00\\_pdfs/indicators.pdf](http://www.cdc.gov/tobacco/data_statistics/surveys/NYTS/00_pdfs/indicators.pdf). Last accessed October, 2008.
9. Potter F. Survey of Procedures to Control Extreme Sampling Weights, in Proceedings of the Section on Survey Research Methods, American Statistical Association, pp 453-458. 1988.
10. Potter F. *A Study of Procedures to Identify and Trim Extreme Sample Weights*, Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 225-230. 1990.
11. Potter F. Survey of Procedures to Control Extreme Sampling Weights, in *Proceedings of the Section on Survey Research Methods*, American Statistical Association 1988;pp 453-458.
12. Potter F. *A Study of Procedures to Identify and Trim Extreme Sample Weights*, *Proceedings of the Section on Survey Research Methods*, American Statistical Association 1990;pp. 225-230.

# **Appendix A**

## **Common Core of Data (CCD) Race/Ethnicity Definitions**

## Appendix A: Common Core of Data Race/Ethnicity Definitions

**American Indian/Alaska Native**—A person having origins in any of the original peoples of North and South America (including Central America) and who maintains cultural identification through tribal affiliation or community recognition.

**Asian/Pacific Islander**—A person having origins in any of the original peoples of the Far East, Southeast Asia, the Indian subcontinent, or the Pacific Islands. This area includes, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, Thailand, Vietnam, Guam, the Philippine Islands, Samoa, and other Pacific Islands.

**Black**—A person having origins in any of the black racial groups of Africa; African American.

**Hispanic**—A person of Mexican, Puerto Rican, Cuban, Central or South American, or other Spanish culture or origin, regardless of race.

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

# **Appendix B**

## **2011 NYTS Survey Instrument**

# National Youth Tobacco Survey (NYTS) 2011 Questionnaire

**This survey is about tobacco. We would like to know about you and things you do that may affect your health. Your answers will be used for programs for young people like yourself.**

**DO NOT write your name on this survey. The answers you give will be kept private.**

**NO one will know what you write. Answer the questions based on what you really do and know.**

**Completing the survey is voluntary. Whether or not you answer the questions will not affect your grade in this class. Try to answer all the questions. If you do not want to answer a question, just leave it blank. There are no wrong answers.**

**The questions that ask about your background will only be used to describe the types of students completing this survey. The information will not be used to find out your name. No names will ever be reported.**

**Please read every question. Try to answer all the questions. Fill in the circles in the booklet completely. When you are finished, follow the instructions of the person giving you the survey.**

Public reporting burden for this collection of information is estimated to average 45 minutes per response, including time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. An agency may not conduct or sponsor, and a person is not required to respond to a collection of information unless it displays a currently valid OMB control number. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to: CDC Reports Clearance Officer, 1600 Clifton Road, NE, MS D-74, Atlanta, GA 30333, ATTN: PRA (0920-0621).

***Thank You Very Much For Your Help***

**DIRECTIONS**

Use a #2 pencil only.

Make dark marks.

Fill in a response like this: ○ ○ ● ○ ○

If you change your answer, erase your old answer completely.

**TODAY'S DATE**

Month	Day	
○ January	○ 0	○ 0
○ February	○ 1	○ 1
○ March	○ 2	○ 2
○ April	○ 3	○ 3
○ May		○ 4
○ June		○ 5
		○ 6
		○ 7
		○ 8
		○ 9

**The first six questions ask for some background information about you.**

- How old are you?
  - 9 years old
  - 10 years old
  - 11 years old
  - 12 years old
  - 13 years old
  - 14 years old
  - 15 years old
  - 16 years old
  - 17 years old
  - 18 years old
  - 19 years old or older
- What is your sex?
  - Female
  - Male
- What grade are you in?
  - 6th
  - 7th
  - 8th
  - 9th
  - 10th
  - 11th
  - 12th
  - Ungraded or other grade

- Are you Hispanic or Latino?
  - No
  - Yes, I am Mexican, Mexican American, or Chicano
  - Yes, I am Puerto Rican
  - Yes, I am Cuban or Cuban American
  - Yes, I am some other Hispanic or Latino not listed here
- What race or races do you consider yourself to be? (You can **CHOOSE ONE ANSWER** or **MORE THAN ONE ANSWER**)
  - American Indian or Alaska Native
  - Asian
  - Black or African American
  - Native Hawaiian or Other Pacific Islander
  - White
- During the past 30 days, about how much money did you have **each week** to spend any way you wanted to?
  - None
  - Less than \$1
  - \$1 to \$5
  - \$6 to \$10
  - \$11 to \$20
  - \$21 to \$50
  - More than \$50

**The next six sections ask about your use of different kinds of tobacco products.**

**Cigarettes**

- Have you **ever tried** cigarette smoking, even one or two puffs?
  - Yes
  - No
- Do you think you will smoke a cigarette at any time in the next year?
  - Definitely yes
  - Probably yes
  - Probably not
  - Definitely not
- Do you think that you will try a cigarette soon?
  - I have already tried smoking cigarettes
  - Yes
  - No
- If one of your best friends were to offer you a cigarette, would you smoke it?
  - Definitely yes
  - Probably yes
  - Probably not
  - Definitely not

11. How old were you when you **first tried** cigarette smoking, even one or two puffs?
- I have never smoked cigarettes, not even one or two puffs
  - 8 years old or younger
  - 9 years old
  - 10 years old
  - 11 years old
  - 12 years old
  - 13 years old
  - 14 years old
  - 15 years old
  - 16 years old
  - 17 years old
  - 18 years old
  - 19 years old or older
12. About how many cigarettes have you smoked in your **entire life**?
- I have never smoked cigarettes, not even one or two puffs
  - 1 or more puffs but never a whole cigarette
  - 1 cigarette
  - 2 to 5 cigarettes
  - 6 to 15 cigarettes (about 1/2 a pack total)
  - 16 to 25 cigarettes (about 1 pack total)
  - 26 to 99 cigarettes (more than 1 pack, but less than 5 packs)
  - 100 or more cigarettes (5 or more packs)
13. During the **past 30 days**, on how many days did you smoke cigarettes?
- 0 days
  - 1 or 2 days
  - 3 to 5 days
  - 6 to 9 days
  - 10 to 19 days
  - 20 to 29 days
  - All 30 days
14. During the past 30 days, **on the days you smoked**, how many cigarettes did you smoke per day?
- I did not smoke cigarettes during the past 30 days
  - Less than 1 cigarette per day
  - 1 cigarette per day
  - 2 to 5 cigarettes per day
  - 6 to 10 cigarettes per day
  - 11 to 20 cigarettes per day
  - More than 20 cigarettes per day
15. When was the last time you smoked a cigarette, even one or two puffs? (**PLEASE CHOOSE THE FIRST ANSWER THAT FITS**)
- I have never smoked cigarettes, not even one or two puffs
  - Earlier today
  - Not today but sometime during the past 7 days
  - Not during the past 7 days but sometime during the past 30 days
  - Not during the past 30 days but sometime during the past 6 months
  - Not during the past 6 months but sometime during the past year
  - 1 to 4 years ago
  - 5 or more years ago
16. During the past 30 days, what brand of cigarettes did you usually smoke? (**CHOOSE ONLY ONE ANSWER**)
- I did not smoke cigarettes during the past 30 days
  - I did not smoke a usual brand
  - American Spirit
  - Camel
  - GPC, Basic, or Doral
  - Kool
  - Lucky Strike
  - Marlboro
  - Newport
  - Parliament
  - Virginia Slims
  - Some other brand not listed here
17. Menthol cigarettes are cigarettes that taste like mint. During the past 30 days, were the cigarettes that you usually smoked menthol?
- I did not smoke cigarettes during the past 30 days
  - Yes
  - No
  - Not sure
18. During the past 30 days, how did you get your own cigarettes? (**You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER**)
- I did not smoke cigarettes during the past 30 days
  - I bought them myself
  - I had someone else buy them for me
  - I borrowed or bummed them
  - Someone gave them to me without my asking
  - I took them from a store or another person
  - I got them some other way

19. During the **past 30 days**, where did you **buy** your own cigarettes? (**You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER**)
- A. I did not buy cigarettes during the past 30 days
  - B. A gas station
  - C. A convenience store
  - D. A grocery store
  - E. A drugstore
  - F. A vending machine
  - G. Over the Internet
  - H. Through the mail
  - I. Some other place not listed here

20. During the **past 30 days**, did anyone **refuse** to sell you cigarettes because of your age?
- A. I did not try to buy cigarettes during the past 30 days
  - B. Yes
  - C. No

### Cigars

21. Have you **ever tried** smoking cigars, cigarillos, or little cigars, even one or two puffs?
- A. Yes
  - B. No
22. How old were you when you **first tried** smoking a cigar, cigarillo, or little cigar, even one or two puffs?
- A. I have never smoked cigars, cigarillos, or little cigars, not even one or two puffs
  - B. 8 years old or younger
  - C. 9 years old
  - D. 10 years old
  - E. 11 years old
  - F. 12 years old
  - G. 13 years old
  - H. 14 years old
  - I. 15 years old
  - J. 16 years old
  - K. 17 years old
  - L. 18 years old
  - M. 19 years old or older

23. During the **past 30 days**, on how many days did you smoke cigars, cigarillos, or little cigars?
- A. 0 days
  - B. 1 or 2 days
  - C. 3 to 5 days
  - D. 6 to 9 days
  - E. 10 to 19 days
  - F. 20 to 29 days
  - G. All 30 days

24. During the **past 30 days**, how did you get your own cigars, cigarillos, or little cigars? (**You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER**)
- A. I did not smoke cigars, cigarillos, or little cigars during the past 30 days
  - B. I bought them myself
  - C. I had someone else buy them for me
  - D. I borrowed or bummed them
  - E. Someone gave them to me without my asking
  - F. I took them from a store or another person
  - G. I got them some other way

25. During the **past 30 days**, where did you **buy** your own cigars, cigarillos, or little cigars? (**You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER**)
- A. I did not buy cigars, cigarillos, or little cigars during the past 30 days
  - B. A gas station
  - C. A convenience store
  - D. A grocery store
  - E. A drugstore
  - F. A vending machine
  - G. Over the Internet
  - H. Through the mail
  - I. Some other place not listed here

### Smokeless Tobacco

26. Have you **ever used** chewing tobacco, snuff, or dip, such as Redman, Levi Garrett, Beechnut, Skoal, Skoal Bandits, or Copenhagen, even just a small amount?
- A. Yes
  - B. No
27. How old were you when you **used** chewing tobacco, snuff, or dip for the first time?
- A. I have never used chewing tobacco, snuff, or dip
  - B. 8 years old or younger
  - C. 9 years old
  - D. 10 years old
  - E. 11 years old
  - F. 12 years old
  - G. 13 years old
  - H. 14 years old
  - I. 15 years old
  - J. 16 years old
  - K. 17 years old
  - L. 18 years old
  - M. 19 years old or older
28. During the **past 30 days**, on how many days did you use chewing tobacco, snuff, or dip?
- A. 0 days
  - B. 1 or 2 days
  - C. 3 to 5 days
  - D. 6 to 9 days
  - E. 10 to 19 days
  - F. 20 to 29 days
  - G. All 30 days

29. During the **past 30 days**, how did you get your own chewing tobacco, snuff, or dip? (**You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER**)
- I did not use chewing tobacco, snuff, or dip during the past 30 days
  - I bought it myself
  - I had someone else buy it for me
  - I borrowed or bummed it
  - Someone gave it to me without my asking
  - I took it from a store or another person
  - I got it some other way
30. During the **past 30 days**, where did you **buy** your own chewing tobacco, snuff, or dip? (**You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER**)
- I did not buy chewing tobacco, snuff, or dip during the past 30 days
  - A gas station
  - A convenience store
  - A grocery store
  - A drugstore
  - A vending machine
  - Over the Internet
  - Through the mail
  - Some other place not listed here

**A Pipe. Do NOT include smoking tobacco in a waterpipe or hookah.**

31. Have you **ever tried** smoking tobacco in a pipe, even one or two puffs?
- Yes
  - No
32. During the **past 30 days**, on how many days did you smoke tobacco in a pipe?
- 0 days
  - 1 or 2 days
  - 3 to 5 days
  - 6 to 9 days
  - 10 to 19 days
  - 20 to 29 days
  - All 30 days

**Bidis (small brown cigarettes wrapped in a leaf) and Kreteks (clove cigarettes)**

33. Have you **ever tried** smoking any of the following, even one or two puffs:
- I have never smoked bidis (small brown cigarettes wrapped in a leaf) or kreteks (clove cigarettes)
  - Bidis
  - Kreteks
  - I have tried both bidis and kreteks

34. During the **past 30 days**, on how many days did you smoke bidis?
- 0 days
  - 1 or 2 days
  - 3 to 9 days
  - 10 to 19 days
  - 20 to 29 days
  - All 30 days
35. During the **past 30 days**, on how many days did you smoke kreteks?
- 0 days
  - 1 or 2 days
  - 3 to 9 days
  - 10 to 19 days
  - 20 to 29 days
  - All 30 days

**Other Tobacco Products**

36. Which of the following tobacco products have you ever tried, even just one time? (**You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER**)
- Roll-your-own cigarettes
  - Flavored cigarettes, such as Camel Crush
  - Clove cigars
  - Flavored little cigars
  - Smoking tobacco from a hookah or a waterpipe
  - Snus, such as Camel or Marlboro Snus
  - Dissolvable tobacco products, such as Ariva, Stonewall, Camel orbs, Camel sticks, or Camel strips
  - Electronic Cigarettes or E-cigarettes, such as Ruyan or NJOY
  - Some other new tobacco products not listed here
  - I have never tried any of the products listed above or any new tobacco product
37. During the **past 30 days**, which of the following tobacco products did you use on **at least one day**? (**You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER**)
- Roll-your-own cigarettes
  - Flavored cigarettes, such as Camel Crush
  - Clove cigars
  - Flavored little cigars
  - Smoking tobacco from a hookah or a waterpipe
  - Snus, such as Camel or Marlboro Snus
  - Dissolvable tobacco products, such as Ariva, Stonewall, Camel orbs, Camel sticks, or Camel strips
  - Electronic Cigarettes or E-cigarettes, such as Ruyan or NJOY
  - Some other new tobacco products not listed here
  - I have not used any of the products listed above or any new tobacco product

**The next twelve questions ask about  
about different issues related to tobacco.**

38. How easy would it be for you to get tobacco products if you wanted some?
- A. Very easy
  - B. Somewhat easy
  - C. Not easy at all
39. Do you believe that tobacco companies try to get young people under 18 to use tobacco products?
- A. Yes
  - B. No
40. When you are using the Internet, how often do you see ads for tobacco products?
- A. I do not use the Internet
  - B. Never
  - C. Rarely
  - D. Sometimes
  - E. Most of the time
  - F. Always
41. When you read newspapers or magazines, how often do you see ads for cigarettes and other tobacco products?
- A. I do not read newspapers or magazines
  - B. Never
  - C. Rarely
  - D. Sometimes
  - E. Most of the time
  - F. Always
42. During the past 30 days, did you receive coupons from a tobacco company through... **(You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER)**
- A. The mail
  - B. E-mail
  - C. The Internet
  - D. Facebook
  - E. Myspace
  - F. A text message
43. During the past 30 days, did you receive ads from a tobacco company through... **(You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER)**
- A. The mail
  - B. E-mail
  - C. The Internet
  - D. Facebook
  - E. Myspace
  - F. A text message
44. When you go to a convenience store, supermarket, or gas station, how often do you see ads for cigarettes and other tobacco products?
- A. I never go to a convenience store, supermarket, or gas station
  - B. Never
  - C. Rarely
  - D. Sometimes
  - E. Most of the time
  - F. Always
45. During the past 30 days, how often did you see an ad for cigarettes or smokeless tobacco that was outdoors on a billboard or could be seen from outside a store?
- A. I did not see an ad for cigarettes or smokeless tobacco during the past 30 days
  - B. Never
  - C. Rarely
  - D. Sometimes
  - E. Most of the time
  - F. Always
46. When you watch TV or go to the movies, how often do you see actors using tobacco?
- A. I do not watch TV or go to the movies
  - B. Never
  - C. Rarely
  - D. Sometimes
  - E. Most of the time
  - F. Always
47. What is the name of the cigarette brand of your favorite cigarette ad? **(You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER)**
- A. American Spirit
  - B. Camel
  - C. GPC, Basic, or Doral
  - D. Kool
  - E. Lucky Strike
  - F. Marlboro
  - G. Newport
  - H. Parliament
  - I. Virginia Slims
  - J. Some other brand not listed here
  - K. I do not have a favorite cigarette ad
  - L. Not sure
48. A warning label tells you if a product is harmful to you and can be either a picture or words. During the past 30 days, how often did you see a warning label on a cigarette pack?
- A. I did not see a cigarette pack during the past 30 days
  - B. Never
  - C. Rarely
  - D. Sometimes
  - E. Most of the time
  - F. Always

49. During the past 30 days, how often did you see a warning label on a smokeless tobacco product?
- A. I did not see a smokeless tobacco product during the past 30 days
  - B. Never
  - C. Rarely
  - D. Sometimes
  - E. Most of the time
  - F. Always

**Some tobacco companies make items like sports gear, T-shirts, lighters, hats, jackets, sunglasses, or other items that people can buy or receive for free.**

50. During the **past 12 months**, did you buy or receive anything that has a tobacco company name or picture on it?
- A. Yes
  - B. No
51. How likely is it that you would ever use or wear something--such as a lighter, T-shirt, hat, or sunglasses --that has a tobacco company name or picture on it?
- A. Very likely
  - B. Somewhat likely
  - C. Somewhat unlikely
  - D. Very unlikely

**The next two questions are about questions and advice any doctor, dentist, or nurse might have asked.**

52. During the **past 12 months**, did any doctor, dentist, or nurse ask you if you use tobacco of any kind?
- A. I did not see a doctor, dentist, or nurse during the past 12 months
  - B. Yes
  - C. No
53. During the **past 12 months**, did any doctor, dentist, or nurse advise you not to use tobacco of any kind?
- A. I did not see a doctor, dentist, or nurse during the past 12 months
  - B. Yes
  - C. No

**The next six questions are about quitting tobacco products.**

54. Do you want to stop smoking cigarettes **for good**?
- A. I do not smoke now
  - B. Yes
  - C. No
55. I plan to stop smoking cigarettes **for good** within the next... **(PLEASE CHOOSE THE FIRST ANSWER THAT FITS)**
- A. I do not smoke now
  - B. 7 days
  - C. 30 days
  - D. 6 months
  - E. 1 year
  - F. I do not plan to stop smoking cigarettes within the next year
56. During the **past 12 months**, how many times did you stop smoking for **one day or longer** because you were trying to quit smoking cigarettes for good?
- A. I did not smoke during the past 12 months
  - B. I did not try to quit during the past 12 months
  - C. 1 time
  - D. 2 times
  - E. 3 to 5 times
  - F. 6 to 9 times
  - G. 10 or more times
57. When you **last tried to quit** for good, how long did you stay off cigarettes? **(PLEASE CHOOSE THE FIRST ANSWER THAT FITS)**
- A. I have never smoked cigarettes
  - B. I have never tried to quit
  - C. Less than a day
  - D. 1 to 7 days
  - E. More than 7 days but less than 30 days
  - F. More than 30 days but less than 6 months
  - G. More than 6 months but less than 1 year
  - H. 1 year or more
58. Are you seriously thinking about quitting the use of all tobacco?
- A. I have never used tobacco
  - B. Yes, within the next 30 days
  - C. Yes, within the next 6 months
  - D. Yes, within longer than 6 months
  - E. I am not thinking about quitting the use of all tobacco

59. During the **past 12 months**, did you do any of the following to help you quit using tobacco of any kind for good? (**You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER**)
- A. I did not use tobacco of any kind during the past 12 months
  - B. I did not try to quit during the past 12 months
  - C. Attended a program in my school
  - D. Attended a program in the community
  - E. Called a telephone help line or telephone quit line
  - F. Used nicotine gum
  - G. Used nicotine patch
  - H. Used any medicine to help quit
  - I. Visited an Internet quit site
  - J. Got help from family or friends
  - K. Used another method, such as hypnosis or acupuncture
  - L. Tried to quit on my own or quit "cold turkey"

**The next five questions ask about your exposure to other people's tobacco smoke.**

60. During the **past 7 days**, on how many days did someone smoke tobacco products in your home while you were there?
- A. 0 days
  - B. 1 day
  - C. 2 days
  - D. 3 days
  - E. 4 days
  - F. 5 days
  - G. 6 days
  - H. 7 days
61. During the **past 7 days**, on how many days did you ride in a vehicle where someone was smoking a tobacco product?
- A. 0 days
  - B. 1 day
  - C. 2 days
  - D. 3 days
  - E. 4 days
  - F. 5 days
  - G. 6 days
  - H. 7 days
62. During the **past 7 days**, on how many days did you breathe the smoke from someone who was smoking a tobacco product at your school, including school buildings, school grounds, and school parking lots?
- A. 0 days
  - B. 1 day
  - C. 2 days
  - D. 3 days
  - E. 4 days
  - F. 5 days
  - G. 6 days
  - H. 7 days

63. During the **past 7 days**, on how many days did you breathe the smoke from someone who was smoking tobacco products in the place where you work?
- A. 0 days
  - B. 1 day
  - C. 2 days
  - D. 3 days
  - E. 4 days
  - F. 5 days
  - G. 6 days
  - H. 7 days
64. During the **past 7 days**, on how many days did you breathe the smoke from someone who was smoking tobacco products in an indoor or outdoor public place? Examples of indoor public places are school buildings, stores, restaurants, and sports arenas. Examples of outdoor public places are school grounds, parking lots, stadiums, and parks.
- A. 0 days
  - B. 1 day
  - C. 2 days
  - D. 3 days
  - E. 4 days
  - F. 5 days
  - G. 6 days
  - H. 7 days

**The next five questions ask about smoking and other tobacco use in your home, your family's cars, and by your friends and family.**

65. Inside your home (not counting decks, garages, or porches) is smoking...
- A. Always allowed
  - B. Allowed only at some times or in some places
  - C. Never allowed
66. In the vehicles that you and family members who live with you own or lease, is smoking...
- A. Always allowed
  - B. Sometimes allowed
  - C. Never allowed
67. Does anyone who lives with you now...? (**You can CHOOSE ONE ANSWER or MORE THAN ONE ANSWER**)
- A. Smoke cigarettes
  - B. Use chewing tobacco, snuff, or dip
  - C. Use snus
  - D. Smoke cigars, cigarillos, or little cigars
  - E. Smoke tobacco using a hookah or waterpipe
  - F. Smoke tobacco out of a pipe other than a hookah or waterpipe
  - G. Smoke bidis (small brown cigarettes wrapped in a leaf)
  - H. Smoke kreteks (clove cigarettes)
  - I. Use any other form of tobacco
  - J. No one who lives with me now uses any form of tobacco

68. How many of your **four** closest friends smoke cigarettes?
- A. None
  - B. One
  - C. Two
  - D. Three
  - E. Four
  - F. Not sure

69. How many of your **four** closest friends use chewing tobacco, snuff, or dip?
- A. None
  - B. One
  - C. Two
  - D. Three
  - E. Four
  - F. Not sure

**The next six questions ask about your thoughts about tobacco.**

70. In your opinion, inside your home, smoking should....
- A. Always be allowed
  - B. Be allowed only at some times or in some places
  - C. Never be allowed

71. In your opinion, in their vehicles, people should...
- A. Always allow smoking
  - B. Sometimes allow smoking
  - C. Never allow smoking

72. Do you think that breathing smoke from other people's cigarettes or other tobacco products is...?
- A. Very harmful to one's health
  - B. Somewhat harmful to one's health
  - C. Not very harmful to one's health
  - D. Not harmful at all to one's health

73. Do you think smoking cigarettes makes young people look cool or fit in?
- A. Definitely yes
  - B. Probably yes
  - C. Probably not
  - D. Definitely not

74. Do you think young people who smoke cigarettes have more friends?
- A. Definitely yes
  - B. Probably yes
  - C. Probably not
  - D. Definitely not

75. How strongly do you agree with the statement 'All tobacco products are dangerous'?
- A. Strongly agree
  - B. Agree
  - C. Disagree
  - D. Strongly disagree

**The last six questions are about your experiences at home, in your community, and at school.**

76. During the **past 12 months**, did your parents or guardians talk with you, even once, about not using any type of tobacco product?
- A. Yes
  - B. No

77. During the **past 12 months**, were you involved in any organized activities to keep people your age from using any form of tobacco product?
- A. Yes
  - B. No

78. During **this school year**, were you taught in any of your classes about why you should not use tobacco products?
- A. Yes
  - B. No

79. During the **past 30 days**, to your knowledge, has anyone, including yourself, smoked a tobacco product on school property when he or she was not supposed to?
- A. Yes
  - B. No

80. During the **past 30 days**, to your knowledge, has anyone, including yourself, used some other type of tobacco product (**that is, one that is not smoked**) on school property when he or she was not supposed to?
- A. Yes
  - B. No

81. During the **past 30 days**, how many days did you miss **at least one class period** because you skipped or "cut" or just did not want to be there?
- A. 0 days
  - B. 1 day
  - C. 2 to 5 days
  - D. 6 to 10 days
  - E. 11 or more days

**Thank you for participating in this survey.**

**DO NOT WRITE IN THIS BOX**

**OFFICE USE ONLY**

---

School ID

0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9

# **Appendix C**

## **2011 NYTS Data Collector Training Agenda**

**Appendix C: Data Collector Training Agenda**  
National Youth Tobacco Survey 2011 Data Collector Training Agenda

**Day 1 - Wednesday, February 2, 2011**

Module 1 (3:00 - 4:00)	Introduction and Orientation
Module 2 (4:00 - 4:15)	Overview of Training
Module 3 (4:15 - 4:45)	Division of Adolescent and School Health (DASH) and the YRBSS
Module 4 (4:45 - 5:15)	Office on Smoking or Health (OSH) and the YTS
Module 5 (5:15 - 5:30)	Parental Permission
Module 6 (5:30 - 6:00)	Roles and Responsibilities
Module 7 (6:00 - 7:00)	Dinner
Module 8 (7:00 - 7:30)	Steps Leading to Data Collection
Module 9 (7:30 - 7:45)	Introduction and Brief Glimpse of the Case Management System
Module 10 (7:45 - 8:00)	Wrap Up and Quick Look at Day 2

**Day 2 - Thursday, February 3, 2011**

Module 11 (8:45 - 9:00)	Review of Day 1 – Q&A
Module 12 (9:00 - 9:20)	Overview of the Data Collection Process
Module 13 (9:20 - 10:00)	Receiving Assignments
Module 14 (10:00 - 10:30)	Starting the Data Collection Process – Advance Calls to Schools –
Module 15 (10:30 - 10:45)	Break
Module 16 (10:45 - 11:00)	Preparing Your Materials
Module 17 (11:00 - 11:30)	Entry Meeting with Principal/Contact Person
Module 18 (11:30 - 12:00)	Teacher Meeting Before Survey
Module 19 (12:00 - 1:00)	Lunch
Module 20 (1:00 - 2:00)	Survey Administration
Module 21 (2:00 - 3:00)	Classroom Forms/Essential Paperwork
Module 22 (3:00 - 3:15)	Break
Module 23 (3:15 - 3:45)	Exit Meeting with Teachers
Module 24 (3:45 - 4:15)	Exit Meeting with Principal/Contact Person
Module 25 (4:15 - 4:30)	Day 2 Wrap Up
OPTIONAL: (4:30 - 5:30)	Open Skill Building Session

### Day 3 - Friday, February 4, 2011

Module 26 (9:00 – 9:15)	Review from Day 2 – Q&A
Module 27 (9:15 – 9:30)	Distribution and Review of Materials Needed for Calls to Week 1 and 2 Schools
Module 28 (9:30 – 10:15)	Telephone Calls to Week 1 and 2 Schools
Module 29 (10:15 – 11:00)	Debriefing on Calls to Schools
Module 30 (11:00 - 11:15)	Break
Module 31 (11:15 – 11:45)	Unpacking and Checking Your Data
Module 32 (11:45 - 12:45)	Lunch
Module 33 (12:45 - 1:45)	Student Participation Rates and the CMS
Module 34 (1:45 – 2:00)	Break
Module 35 (2:00 – 2:30)	Employment Forms, Administrative Procedures, and Expense Reimbursement Procedures
Module 36 (2:30 - 3:15)	Expense Report Problems and Procedures
Module 37 (3:15 – 4:30)	Day 3 Wrap up, Role Play/Simulation Assignments, and Practice
OPTIONAL: (4:30 – 5:30)	Open Skill Building Session

### Day 4 - Saturday, February 5, 2011

Module 38 (9:00 – 9:15)	Review from Day 3 – Q&A
Module 39 (9:15 - 11:00)	Role Plays/Simulations – Advance Call, Entry Meeting with Principal/Contact Person, and Entry Meeting with Teacher - <b>Team</b>
Module 40 (11:00 – 11:15)	Break
Module 41 (11:15 - 12:30)	Continued Role Plays/Simulations – Survey Administration, Exit Meeting with Teachers, and Exit Meeting with Principal/Contact Person
Module 42 (12:30 - 1:30)	Lunch
Module 43 (1:30 - 2:00)	Wrap up Role Plays/Simulations
Module 44 (2:00 - 2:30)	Distribution and Return of Project Materials
Module 45 (2:30 – 2:45)	Break
Module 46 (2:45 - 3:15)	Areas for Priority Attention
Module 47 (3:15 - 3:30)	Day 4 Wrap Up, Q & A, and Closure
Dinner (5:00)	

# Appendix D