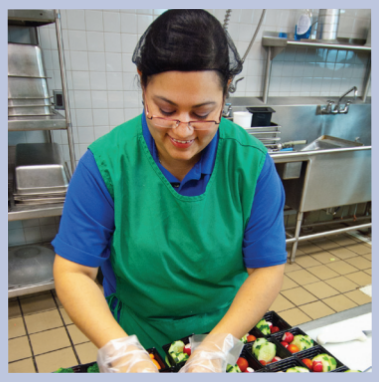

United States Department of Agriculture
Food and Nutrition Service

# School Nutrition Dietary Assessment Study-IV

## Volume II: Sampling and Data Collection Methods

Office of Research and Analysis • November 2012

United States Department of Agriculture
Food and Nutrition Service

November 2012
Child Nutrition Programs
Report No. CN-12-SNDA

# School Nutrition Dietary Assessment Study IV
# Volume II
# Sampling and Data Collection Methods

**Authors:**
From Mathematica Policy Research, Inc.:
John Hall
Eric Zeidman
Mary Kay Crepinsek
Elizabeth Condon

**Submitted by:**
Mathematica Policy Research, Inc.
955 Massachusetts Avenue
Cambridge, MA 02139

**Submitted to:**
Office of Research and Analysis
USDA, Food and Nutrition Service
3101 Park Center Drive, Room 1014
Alexandria, VA 22302-1500

**Project Director:**
Mary Kay Fox

**Project Officer:**
Fred Lesnett

This study was conducted under Contract number GS-10F-0050L with the Food and Nutrition Service.

This report is available on the Food and Nutrition Service website: http://www.fns.usda.gov/ora.

**Suggested Citation:**
U.S. Department of Agriculture, Food and Nutrition Service, Office of Research and Analysis, *School Nutrition Dietary Assessment Study IV*, Vol. II: *Sampling and Data Collection Methods,* by John Hall, Eric Zeidman, Mary Kay Crepinsek, and Elizabeth Condon. Project Officer, Fred Lesnett Alexandria, VA: November 2012.

# ACKNOWLEDGEMENTS

Many people and organizations contributed to the fourth School Nutrition Dietary Assessment (SNDA-IV) study. First and foremost, we would like to thank the school foodservice professionals, principals, and other staff in the school districts and schools across the country who participated in the study. Without their cooperation and hard work, the study could not have been completed.

Staff in the Office of Research and Analysis (ORA), Food and Nutrition Service (FNS), U.S. Department of Agriculture (USDA) provided oversight for the study. Fred Lesnett, as Project Officer, and others at FNS, including Jay Hirschman, Margaret Applebaum, Laura Walter, Eileen Ferruggiaro, Gabrielle Serra, and other Child Nutrition staff, made valuable contributions to the study's design and implementation, provided valuable insights and useful guidance throughout the project, and carried out a diligent review of all study reports.

FNS also sent the second draft of the report to 12 distinguished experts for review, all of whom provided helpful comments. The external reviewers included:

- Jamie Chriqui (Institute for Health Research and Policy), Goulda A. Downer (Howard University School of Medicine), Patricia Guenther (USDA, Center for Nutrition Policy and Promotion), Joanne Guthrie (USDA, Economic Research Service), Ellen Harris (USDA, Agricultural Research Service), Caitlin Merlo (U.S. Department of Health and Human Services, Centers for Disease Control and Prevention), Claire H. Miller (formerly with USDA's CN Division), Mary Nettles (National Food Service Management Institute), Constance Newman (USDA, Economic Research Service), Colette Rihane (USDA, Center for Nutrition Policy and Promotion), Marlene Schwartz (Yale University, Rudd Center for Food Policy and Obesity), and Katie Wilson (National Food Service Management Institute).

Staff in the Food Surveys Research Group at USDA's Agricultural Research Service, including Meghan Adler, Jaspreet Ahuja, Shanthy Bowman, James Friday, Staci McGonigal, Grace Omolewa-Tomobi, Donna Rhodes, and Lois Steinfeldt, developed full nutrient and food group profiles for 200 commonly offered commercially prepared school foods. This assistance allowed SNDA-IV researchers to accurately characterize the nutrient and food group content of specially formulated foods used by schools to promote healthier meals.

Many staff at Mathematica Policy Research contributed to SNDA-IV. Mary Kay Fox served as project director. John Hall, assisted by Cathy Lu and Thomas Decker, designed and implemented the sampling methodology and developed sampling weights. Eric Zeidman directed all data collection activities. He was ably assisted by Rebecca Weiner, Amanda Reiter, Mindy Hu, Stacie Feldman, Jessica Droz, and Brittany Vas, as well as Loring Funaki, Marcia Tobias, and many staff at Mathematica's Survey Operations Center. Staff in Mathematica's Information Services division, including April Villone, Roland Scurato, Carlo Caci, Mark Beardsley, Neil DeLeon, Shilpa Khambhati, Mark Brinkley and Jason Weber developed the programming that made the complex data collection possible. School Food Authorities were recruited into the study by Larry Vittoriano, Ralitsa Slavenkova, Yumiko Jolly, Bryce Onaran, Mindy Hu, Marcia Comly, Premini Sabaratnam, and Charlotte Cabili.

Elizabeth Condon and Mary Kay Crepinsek trained technical assistants who maintained frequent telephone contact with school foodservice managers as they completed the week-long menu survey. They also trained nutrition coders and managed the coding operation. Nutritionists Denise Mercury,

Katherine Niland, Dolores Wolongevicz, and Megan Miraglia supervised a team of nutrition coders in Mathematica's Cambridge office and conducted quality control checks of their work.

Elizabeth Condon and Mary Kay Crepinsek led analysis of the menu survey data, and were assisted by Mary Kay Fox, Denise Mercury, and Katherine Niland. Sarah Forrestal, Alexandra Killewald, and Charlotte Cabili led the analysis of survey data. A cadre of systems analysts and programmers, including Nora Paxton, Will Crumbley, Jessica Galin, Ryan Callahan, Rachel Thompson, Andrew McGuirk, Sabrina Rahman, Adrienne Martz, and Mike Gagnon, worked with these researchers to create analysis variables and generate study findings. Anne Gordon and Nathan Wozny coordinated analyses that compared SNDA-IV findings to previous SNDA studies. Anne Gordon and Patricia McKinney, a former FNS project officer who oversaw previous rounds of the SNDA studies, provided quality assurance review of preliminary tables, figures, and report chapters. Ms. McKinney also served as a technical advisor during the design phase of the project, along with Tracy Fox of Food, Nutrition & Policy Consultants, LLC.

Harshini Shah, of Mathematica, and Anne Garceau, an independent consultant at the time, contributed to early drafts of the final report. The final report was edited by John Kennedy and produced by Autumn Parker. Kari Beckmann, Alexandra Clifford, and Donna Dorsey provided production assistance on early drafts of the report. Deirdre Sheean designed the report cover as well as the separate Summary of Findings report.

# CONTENTS

**This page left blank for double–sided printing.**

# TABLES

**This page left blank for double–sided printing.**

# CHAPTER 1
# INTRODUCTION

The National School Lunch Program (NSLP) and the School Breakfast Program (SBP) provide meals to children during the school year. The overarching goal of both programs, known collectively as the school meal programs, is to ensure that children do not go hungry—that they have access to nutritious meals that support normal growth and development. The programs provide a safety net for children from low-income families, who are eligible to receive school meals free or at a reduced price. In recent years, program administrators at the Federal, State, and local levels have worked to enhance the nutritional quality of school meals, to better align them with the dietary practices recommended in the *Dietary Guidelines for Americans*.

The U.S. Department of Agriculture (USDA), which administers the school meal programs, has assessed the programs on a periodic basis since the 1980s. The fourth School Nutrition Dietary Assessment study (SNDA-IV) was conducted by Mathematica Policy Research under contract with USDA's Food and Nutrition Service (FNS). This report is the second of two volumes of the SNDA-IV final report. This volume describes the study methodology, including sample design, data collection, coding procedures for school menu data, and construction of sample weights necessary to obtain nationally representative estimates from the study sample. Study findings are presented in Volume I and in a separate summary report.

SNDA-IV included a small supplementary sample of schools participating in USDA's HealthierUS Schools Challenge (HUSSC) program. All of the data collected in SNDA-IV were collected for these schools, and the data were processed and analyzed using comparable approaches. However, this sample of schools was completely separate from the main SNDA-IV sample. Thus, details provided in this report about sampling (Chapter 2), response rates (Chapter 3), and calculation of sampling weights (Chapter 4) do not apply to the supplementary sample of HUSSC schools. See Volume I, Chapter 12 for information about the supplementary sample of HUSSC schools.

**This page left blank for double–sided printing.**

# CHAPTER 2
# SAMPLE DESIGN AND SELECTION

The school meal programs are administered at the local level by School Food Authorities (SFAs), which usually are individual school districts or small groups of districts. The overall objective of the sampling plan was to provide nationally representative samples of public SFAs and schools that participate in the NSLP. The sample design included two samples—the SFA-only sample, which collected data only at the SFA level, and the SFA-plus-school sample, which collected data at both the SFA and school levels. A stratified two-stage sampling approach was used, with SFAs selected first and schools selected second, within a random subsample of sampled SFAs. In sampling terms, the primary selection unit (PSU) was the SFA and schools were the secondary selection units (SSUs). As in previous SNDA studies, the respondent universe included all public SFAs and schools participating in the NSLP and located in the contiguous 48 states and the District of Columbia.[1]

## A. Sampling Frames

Two sampling frames were required, one to select PSUs and the other to select schools (SSUs) within sampled PSUs. Developing the sample frame of PSUs required the use of multiple lists because no comprehensive frame of SFAs with all of the information needed for stratification exists. We relied primarily on the National Center for Education Statistics (NCES) 2006-2007 Common Core of Data (CCD) Local Education Agency (LEA) Universe Survey Data (http://nces.ed.gov/ccd/pubagency.asp).[2] Not all of the LEAs (school districts) identified in the CCD are SFAs, so we also employed a file provided by FNS containing data from the School Food Authority Verification Summary Report (FNS-742). Since the FNS-742 file contains records of SFAs, merging it to the CCD file of school districts enabled us to determine, in some cases, which school districts are SFAs. Districts that were not identified as SFAs via matching with FNS-742 were screened for SFA status. In addition, we used the U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE), (http://www.census.gov/hhes/www/saipe/district.html) to obtain district-level estimates of school age children in poverty.

The sampling frame for selecting the SFA sample was a list of PSUs. Before forming PSUs, districts on the CCD that were clearly ineligible were removed. These included districts that:

- were found only on the Census (SAIPE) file and not on the CCD
- were located outside the contiguous (48) United States plus the District of Columbia
- were State or federally operated agencies
- had ceased to operate (according to the CCD)

[1] SNDA-I, which included private schools, was an exception to this rule.

[2] This was the most recent version of the database available at the time the sampling frame was constructed.

- reported no schools or students and could not be connected to any other eligible district, to an operating school, or to students on the school-level CCD file.[3]

A PSU on the frame may be a single SFA (appears on FNS-742), a single district for which SFA status has not been determined (on CCD, but either not on or cannot be linked to FNS-742), or a group of districts or SFAs (those that are part of the same supervisory union).[4] The reason for keeping groups of districts or SFAs in a common supervisory union together was that within a supervisory union there may be a single SFA that serves multiple districts. If there were multiple SFAs in any PSU, we sampled a single SFA for data collection. Separate sampling frames of SSUs (schools) were constructed within each SFA selected for the SFA-plus-school sample. The school- level frames employed the CCD 2006-2007 Public Elementary/Secondary School Universe Survey (http://nces.ed.gov/ccd/pubschuniv.asp) as the main source of information.

## B. Stratification

Two samples of PSUs were selected using somewhat different methods: (1) a sample large enough to yield approximately 300 SFA Director Surveys, but no school surveys (the SFA-only sample); (2) a large enough sample of SFAs so that, in addition to approximately 300 SFA additional SFA director surveys, school-level data could be obtained from approximately 900 schools in those SFAs (the SFA-plus-school sample). To select these samples we first stratified the entire frame of PSUs, and then randomly divided the frame in half. Stratified samples were then selected from each frame, using the same strata used in dividing the frame. The stratifying variables used were region, urbanicity, poverty level, enrollment, and number of schools in the SFA. Each PSU sample was selected using probability proportional to size (PPS) sampling methods with different measures of size (MOS) used for the two samples. The MOS for the SFA-only sample was the square root of the number of schools; for the SFA-plus-school sample, the MOS was the number of schools.[5]

After the MOS had been assigned, the next step was to define certainty selections—those with a MOS so large that their probability of selection in a PPS sample would be 1.0 or close to 1.0. There were two levels of certainty selection. Some SFAs had a large enough number of schools to be designated as a certainty selection for the SFA-only sample; however, not all of these had enough schools to be selected with certainty for the SFA-plus-school sample. Thus, the first two strata were:

- SFAs with enough schools to be designated as certainty selections for both the SFA-only and SFA-plus-school samples—these were assigned to the SFA-plus-school sample.

[3] Under this criterion, districts that are not part of a supervisory union were considered ineligible if the district level report (on the CCD) did not indicate any schools or any students in grades K-12, and (a) the district did not have the same NCES identifier, or Local Education Agency ID (LEAID), as any school in the school-level file or (b) any school having the district's LEAID was closed or had no students. Districts that are part of a supervisory union were considered ineligible if the district met the ineligibility criteria for the non-supervisory-union districts and, in addition, did not link to any other eligible district (through its UnionID).

[4] Districts where, on the CCD, TYPE06 = 2 or 3 belong to supervisory unions.

[5] Use of the square root measure for the SFA-only sample assured representation of large SFAs and more precise SFA-level estimates. The SFA-plus-school sample was intended to provide estimates for both SFAs and schools. Using the number of schools as the MOS for this sample increased the precision of school-level estimates.

- SFAs large enough to be designated as certainty selections for the SFA-only sample but not the SFA-plus-school sample—these could be randomly assigned to either sample and were treated as certainty selections if they were assigned to the SFA-only sample.

The SFAs large enough to be certainty selections for the SFA-only sample but not the SFA-plus-school sample were further stratified when the subsample of SFAs was selected for the SFA-plus-school sample, using the same stratifying variables as those used for PSUs not large enough to be selected with certainty.

PSUs not large enough to be designated as certainty selections (referred to below as non-certainty PSUs) were assigned to non-certainty strata before selection of the SFA-only sample. In addition to including FNS region (of which there are seven), the following stratifying variables were constructed:

- **Degree of Urbanicity.** The CCD defined 12 levels. We defined three levels: in a city, in a suburb or town, or in a rural area.
- **District Child Poverty Level.** We defined two levels of poverty: high poverty, which included PSUs where prevalence of school-age children in poverty was estimated to be 30 percent or more, and lower poverty, which included the remainder of PSUs. We derived poverty estimates first from the U.S. Census SAIPE files. In cases where there SAIPE data were not available, we imputed the prevalence of children in poverty using data on the CCD, including district type, number of students certified for free meals, and degree of urbanicity.
- **Enrollment.** Because we sampled with PPS and had certainty strata, the value of stratifying the non-certainty PSUs by size is diminished. However, to ensure that smaller SFAs were represented, we formed two size categories in each FNS region: above or below the median enrollment among non-certainty PSUs for that region.
- **Number of Schools.** We formed four categories: 1 to 4 schools, 3 to 5 schools, 6 or 7 schools and more than 7 schools.

## C. Sample Allocation and Selection

Before selecting the two samples, the overall frame was randomly divided into 2 frames. PSUs with enough schools to be designated as certainty selections (see preceding discussion) for both the SFA-only and SFA-plus-school samples were assigned to the frame for the SFA-plus-school sample. Half of all remaining SFAs were randomly assigned to the SFA-only sample frame and the remainder to the SFA-plus-school sample frame. From each frame, we selected a sample of PSUs using PPS methods. An initial sample of PSUs was expanded to allow for ineligibility (not all PSUs defined in the frame contained a study-eligible SFA) and nonresponse.

Certainty selections were made first. Then, a sample of pairs of non-certainty PSUs was made. Selections were made so that the PSUs in a pair were similar with respect to characteristics used for stratification. Within each pair, one PSU was randomly designated as the main selection and the other

as reserve. The reserve PSU was typically used only if the main selection in its pair was ineligible or declined to participate in the study.[6] This method helps assure that the final sample resembles the initial sample on characteristics used for stratification. Because there were instances where both members of a pair did not participate, the initial sample included 21 extra (back-up) pairs within each stratum, defined by region and degree of urbanicity. These extra pairs were used only in cases where complete pairs did not participate (due to ineligibility or nonresponse).

Of the certainty PSUs with enough schools to be retained with certainty for both the SFA-only and SFA-plus-school samples, three were considered large enough to receive a double allocation of schools (and to represent two SFAs each).[7] Because of the double allocation, the number of unique SFAs in the SFA-plus-school sample was reduced by three.

Non-certainty pairs of PSUs (with the exception of back-up pairs) were randomly assigned to be part of the SFA-only sample or the SFA-plus-school sample. Schools were sampled in the designated subsample of SFAs (298 SFA equivalents) that were sampled for the SFA-plus-school sample. Strata were defined within SFA by school level (elementary, middle, or high), and schools were selected with equal probability within strata, within SFAs. The target was one school of each type within an SFA. However, because some SFAs had fewer than three schools and some SFAs did not have schools in all strata, some SFAs were allocated extra schools. Thus, while most SFAs in the SFA-plus-school sample had three sampled schools, some had one or two and others had four. Those with a double allocation had a target of six.

For PSUs where the target was three schools and each stratum contained at least two schools, the initial sample included two from each stratum, for a total of six. Allocations for PSUs that had a target of three schools but had other school configurations were as follows:

- If the PSU contained at least six schools but one stratum contained no schools, then three were selected from each of the other two strata, for a total of six.
- If a PSU contained at least six schools, all in one stratum, then six were selected from that stratum.
- If a PSU contained at least six schools, but one stratum contained only one school, then the only school in that stratum was selected and the other school that would have been allocated to that stratum was assigned to another stratum.[8]
- If a PSU included at least six schools, but two of the strata had only one school, then four schools were selected from the other stratum.

[6] In a few instances, SFAS participated in the study after earlier indications of their intent not to do so. If their corresponding reserve selection had already been released, both selections then remained in the sample.

[7] We calculated the number of "PSU equivalents" for each PSU, where one PSU equivalent is equal to the sum of all the PSU sizes (in the SFA-plus-school frame) divided by 300 (the desired number of SFAs participating in the study from this frame). For PSUs with more than 1.8 PSU equivalents, we allocated a double sample of schools.

[8] If the elementary or high school stratum had only one, then the extra school was assigned to the middle school stratum; if the middle school stratum had only one, it was assigned to the elementary stratum.

If the PSU contained fewer than six schools, all schools were selected. In cases where PSUs received an allocation of four school interviews, the initial sample was eight schools. (Only PSUs with 8 or more schools received an allocation of four schools.) The distribution of the schools selected in these PSUs depended on the distribution of the expected shortfall among small SFAs with few schools or with no schools in some strata. The samples of schools were selected in two steps, each with equal probability within stratum, within SFA. First we selected a sample from the 2006–2007 CCD. After that selection, the preliminary file for 2007–2008 became available.[9] If we found schools in sampled SFAs on the more recent CCD that did not appear on the earlier version, these were selected and the initial sample for the SFA was selected from among schools selected on the two versions of the CCD. If the initial sample was four, five, or six schools, three schools were randomly selected as the main sample, and the others were designated as a reserve to be used in case of ineligibility or nonresponse. Similarly, if the initial sample was eight or more, half were randomly selected as the main sample. To the extent possible, a non-participating or ineligible school in the main sample was replaced by a reserve from the same stratum.

As discussed further in Chapter 3, there was some nonresponse at both the SFA and school levels, as well as variation in nonresponse across the different data collection instruments administered at the school level. Table 2.1 shows the respondent universe, initial samples, and completed samples for each level and instrument. Data collection instruments are described in Chapter 3.

**Table 2.1. Respondent Universe, Initial Samples, and Completed Samples**

| | Respondent Universe | Initial Sample | Completed Sample |
|---|---|---|---|
| SFAs — recruited | 14,500 | 747 | 595 |
| SFA Director Survey | 14,500 | 595 | 578 |
| | | | |
| Schools — recruited | 102,000 | 1,059 | 902 |
| Menu Survey | 102,000 | 902 | 884 |
| Foodservice Manager Survey | 102,000 | 902 | 876 |
| A la Carte Checklist | 102,000 | 902 | 895 |
| Principal Survey | 102,000 | 902 | 721 |
| Vending Machine Form | 102,000 | 902 | 680 |
| Other Food Sources Form | 102,000 | 902 | 732 |

Note: Recruited SFAs includes SFAs in both the SFA–only and SFA–plus–school samples. SFAs in the SFA–only sample were not formally recruited into the study.

[9] We did not believe it advisable to wait for the later, preliminary file for the main sampling because the preliminary CCD files may not be as complete as the final versions and are more likely to contain incorrect information.

This page left blank for double–sided printing.

# CHAPTER 3
# DATA COLLECTION

As described in Chapter 2, the study included two samples—the SFA-only sample, for which data were collected only at the SFA level, and the SFA-plus-school sample, for which data were collected at both the SFA and school levels. For the SFA-plus-school sample, the first step in the data collection process was recruiting SFAs to participate in the study. SFAs included in the SFA-only sample were not formally recruited into the study. Rather, SFA directors (the only respondents in the SFA-only sample) were invited by e-mail to complete the web-based SFA director survey.

## A. Recruiting SFAs in the SFA-Plus-School Sample

Recruitment began by securing support for the study at the national, regional, and State levels. Endorsements were obtained from the School Nutrition Association (SNA) and the American Dietetic Association's School Nutrition Service dietetic practice group. The SNA provided a letter for inclusion with study recruitment materials. The recruiting team contacted Child Nutrition (CN) liaisons in each of FNS's regional offices and State CN directors by e-mail and telephone. State CN directors were requested to provide contact information for each of the SFAs sampled in their States. They were also asked to directly encourage sampled SFAs to participate in the study.

Recruitment materials were mailed to directors of sampled SFAs. The mailing included an introductory study letter listing the sampled schools within the SFA, the SNA letter of support, and a study fact sheet. Followup telephone calls were made by recruiters to confirm receipt of the mailing, describe the study objectives and participation requirements, and address any questions or concerns the SFA director might have. Recruiters then reviewed the list of sampled schools and sought the SFA director's approval for each school's participation. In cases where individual sampled schools in an SFA were closed, ineligible under the study design, or refused to participate, replacement sampled schools were presented to the SFA as an alternative for study participation. A target week was agreed upon for menu survey reporting, and the SFA recruitment interview was completed. This interview gathered basic information about the SFA and sampled schools within the SFA that was needed for planning data collection.

A followup mailing was sent to SFA directors who agreed to participate in the study. The mailing included a letter to the SFA director that confirmed the schools participating in the study and the agreed upon target week. It also included letters and copies of the study fact sheet for the foodservice managers and principals in each of the sampled schools, which SFA directors were asked to distribute.[1]

A total of 382 SFAs in the SFA-plus-school sample were released for recruitment. Twenty SFAs were found to be ineligible and 298 agreed to participate in the study, resulting in an 82 percent recruitment rate among SFAs in the SFA-plus-school sample (Table 3.1). This rate is based on all SFAs ever part of the recruitment effort, including replacements for SFAs in the main sample that refused to participate. SFA directors generally agreed to have all of the sampled schools in their

[1] Direct contact was made with school foodservice managers, principals, and other school-level respondents as part of the various data collection tasks. A separate data collection contact was also made with SFA Directors to complete their own survey.

district participate in the study. In SFAs that agreed to participate, 98 percent of the sampled schools were successfully recruited.

**Table 3.1. Final Recruitment Samples**

| | Number of SFAs/Schools | | | | | |
|---|---|---|---|---|---|---|
| | Recruited | Closed | Ineligible | Refused | Total | Percent of Eligible SFAs/Schools Recruited |
| SFAs | 298 | N/A | 20 | 64 | 382 | 82.3 |
| Schools | 902 | 39 | 102 | 16 | 1,059 | 98.3 |
| Elementary | 316 | 14 | 18 | 3 | 351 | 99.1 |
| Middle | 297 | 13 | 20 | 6 | 336 | 98.0 |
| High | 289 | 12 | 64 | 7 | 372 | 97.6 |

Note: Table includes only SFAs in the SFA-plus-school sample.

## B. Data Collection Procedures

Data were collected from January through June 2010. Respondents included SFA directors, school foodservice managers (FSMs), and principals. In addition, an individual designated by the principal provided information about foods available in vending machines, school stores, and other venues outside of the school meal programs. Table 3.2 shows the data collection instruments used, along with information about respondents and mode of data collection.

**Table 3.2. Data Collection Instruments**

| Instrument | Respondent | Mode |
|---|---|---|
| **SFA Level** | | |
| Recruitment Interview | SFA director (only SFAs in the SFA-plus-school sample) | Telephone |
| SFA Director Survey | SFA director | Web, with telephone followup |
| **School Level** | | |
| Menu Survey | School foodservice manager | Mail with intensive telephone-based training, technical assistance, and followup |
| A la Carte Checklist | School foodservice manager | Mail |
| Foodservice Manager Survey | School foodservice manager | Mail |
| Principal Survey | Principal | Web, with telephone followup |
| Competitive Foods Checklists<br>Vending Machines<br>Other Sources of Foods/Beverages | Principal's designee | Fax-back, with training module[a] and telephone followup |

[a]A PowerPoint (converted to pdf format when necessary) training module discussed the data collection forms in detail, described the protocol for completing and returning the forms, raised ambiguous situations and provided instructions on how to address them, and answered frequently asked questions.

### 1. SFA-Level Data

The recruitment interview was completed only for SFAs in the SFA-plus-school sample. This interview was completed with SFA directors as soon as they agreed to participate in the study. The interview focused on selected schools within the SFA and requested basic information required to assess study eligibility and the accuracy of sample frame data and to plan for and support data collection at the school level. Information collected for sampled schools included whether the school participated in the NSLP (only schools that participated in the NSLP were eligible for inclusion in the study) and the SBP; whether they offered afterschool snacks through the NSLP; the grades included in the school; the type of menu-planning and meal-preparation systems used; and contact information for the school's FSM. Few variables were used for analytic purposes and those that were used were added to other school-level files.

The web-based SFA director survey collected data on SFA policies and practices regarding menu planning, a la carte foods, food purchasing, food safety and sanitation, nutrition promotion, and school wellness policies.

### 2. School-Level Data

At the school level, data were collected from the FSM, the principal, and a school staff member designated by the principal. The central component of the data collection—the menu survey—is described in detail below. In addition to the menu survey, the following instruments were used:

- **A la Carte Checklist.** The a la carte checklist documented whether a la carte foods were available to students at breakfast or lunch and, if so, the specific foods and beverages that were available. The checklist was completed by the FSM on one randomly assigned day during the target week.
- **School Foodservice Manager Survey.** The FSM survey collected information about the characteristics of school kitchens, availability and revenue from school foodservice-operated vending machines, meal pricing, scheduling of meal periods, nutrition promotion activities, and practices used to count reimbursable meals and to distribute and count afterschool snacks.
- **Principal Survey.** The web-based principal survey collected information on mealtime policies (including whether students were allowed off campus and what the rules were about buying a la carte foods), other activities scheduled during mealtimes, vending machines, school stores and snack bars, requirements for nutrition education and physical education, opportunities for physical activity during the school day, and school wellness policies.
- **Competitive Foods Checklists.** The competitive foods checklists were completed by a member of the school staff designated by the principal. The checklists documented the presence of vending machines (vending machines checklist), school stores, snack bars, fundraisers, and other sources of foods and beverages (other sources of foods and beverages checklist), and the specific foods available in each venue. Respondents received a training module, which could be accessed using a web link or received by e-mail. The training module discussed the data collection forms in detail, described the protocol for completing and returning the forms, raised ambiguous situations and provided instructions on how to address them, and answered frequently asked questions.

For some schools, the competitive foods checklists were completed by telephone. To obtain cooperation in these cases, data collection was limited to documenting the types of competitive food venues available. Detailed information about the specific foods and beverages available in the various venues was not collected.

**a. The Menu Survey**

The goal of the menu survey was to collect detailed data on all foods offered and served in NSLP lunches, SBP breakfasts, and afterschool snacks (if offered). Data needed to be sufficiently detailed to support a comprehensive assessment of nutrient content. Data were collected for one school week, referred to as the "target week." The target week typically included five school days. However, due to holidays and other school closings, some schools provided data for only four days and a very small number of schools provided data for only three days.

The menu survey was completed by FSMs who received training and intensive support from specially trained Mathematica technical assistants (TAs). The survey included the following five forms:

- **Daily Meal Counts Form.** The daily meal counts form was used to report the number of reimbursable breakfasts and lunches served, by reimbursement category, each day of the target week. It also captured information about total a la carte revenue during the target week.
- **Reimbursable Foods Form.** This form was used to identify foods and beverages offered to students in reimbursable meals each day of the target week. Separate forms were completed for breakfast and lunch. The form was designed to obtain, for each food and beverage offered, descriptive details needed for accurate nutrient analysis, portion sizes, and the number of portions served or sold in reimbursable meals.
- **Self-Serve/Made-to-Order Bar Form.** This form was used to list and describe foods offered in condiment/finishing bars, salad bars, sandwich bars, and other self-serve and made-to-order bars. For bars offered more than once during the target week, respondents were asked to list all ingredients only on the first day the bar was offered. Information provided on the ingredients offered on the bar was used to create a "recipe" to estimate the nutrient content of an average serving from the bar.
- **Recipe Form.** FSMs were asked to complete a recipe form for all foods prepared from scratch or by combining two or more foods or ingredients. The form collected information about ingredients, yield, and preparation methods. To minimize the level of effort needed to report recipes and reduce the potential for missing information, respondents were free to provide copies of their own printed recipes rather than copying them onto the form. However, instructions provided with the form emphasized the need to edit printed recipes if ingredients had been modified, for example, if ground beef had been substituted in a recipe that calls for ground turkey, or vice versa.
- **Afterschool Snack Form.** This form captured data on foods offered and served in reimbursable afterschool snacks during the target week, as well as information about the total number of snacks served each day.

To aid respondents in organizing this elaborate instrument, forms were assembled into a carefully designed packet. Key features of the Menu Survey packet include:

- **Color Coding.** Each form was a different color so forms could be easily identified by both title and color. In the instruction booklet, instructions for each form incorporated the corresponding color ink.
- **Simple, Clear Instructions, with Samples of Completed Forms.** Respondents received an instruction manual that provided simple, yet complete instructions for completing each form. The manual included clearly marked samples of completed forms which provided respondents with examples of how information should be entered on each form.
- **Tip Sheet.** A tip sheet, printed on cardstock, provided a one-page summary of key instructions for each form. The tip sheet provided a quick reference for respondents so they did not have to reference the full set of instructions each time they had a question about a form.
- **Portfolio.** Menu survey materials were presented in an attractive plastic, multi-pocket portfolio. The portfolio kept survey forms separate and neatly organized for each day of the target week.
- **Bar Codes.** Pre-printed bar code labels were used for each form, so that respondents did not have to label each form with the school name and study identification number.

The Menu Survey packet also included the FSM survey and a $50 incentive check to thank FSMs for their time and efforts.

### b. Menu Survey Data Collection Procedures

Prior to the target week, TAs initiated contact with the FSM in each sampled school. The TA introduced the study, established rapport with the FSM, and confirmed the target week. The TA then made arrangements for the menu survey packet to be shipped to the FSM. After the FSM had received the menu survey packet, but prior to the target week, the TA conducted a telephone training session. The training covered the contents of the menu survey, procedures and schedule for completing the survey, and frequently asked questions. Depending on SFA directors' preferences, the training call was conducted jointly for all schools in the SFA or separately for each school.

Following the training, TAs had direct responsibility for working with FSMs to ensure that the menu survey was completed in an accurate and timely manner. The protocol called for TAs to:

- Place a reminder call the day before the target week began to confirm that everything was on track to begin the survey and to highlight helpful hints about survey completion.
- Be reachable at a toll-free telephone number during normal foodservice operation hours to address any questions from respondents about survey instructions, forms, and procedures.
- Contact respondents periodically during the target week to review instructions, assist with completing forms, and answer questions as needed. (The final contact included a reminder to complete the FSM survey.)
- Issue reminders (as needed) following the target week to encourage prompt return of the completed survey.

- Perform a quality control review of the returned and completed forms, with prompt followup to obtain any missing or incomplete information, or to discuss corrections while the information was still recent.

## 3. Survey Results

Final completed sample sizes and response rates are shown in Table 3.3. SFA directors and school foodservice managers that agreed to participate in the study were very cooperative with the data collection. The response rate for the main component of the study—the menu survey—was very high, at 98 percent. Gaining cooperation from school principals was more challenging. The SFA directors who agreed to participate in the study did not have the authority to compel principals to participate, as they generally did with FSMs. The finite end date for the data collection period (the end of the school year) limited the amount of followup that could be done with nonresponding principals. The responsiveness of principals also affected response rates for the competitive foods checklists, since the data collection protocol called for the principal to designate a respondent for those instruments. For these reasons, response rates for the principal survey and the competitive foods checklists were lower than for the other components of the study.

**Table 3.3. Final Sample Sizes and Response Rates**

| | Completed Sample Size | Weighted Response Rate (%) |
|---|---|---|
| SFA Director Survey | 578 | 94.0 |
| Menu Survey | 884 | 97.7 |
| Foodservice Manager Survey | 876 | 96.7 |
| A la Carte Checklist | 895 | 99.5 |
| Principal Survey | 721 | 87.2 |
| Vending Machine Checklist | 680 | 79.0 |
| Other Sources of Foods and Beverages Checklist | 732 | 88.1 |

Notes: All response rates are weighted using raw sampling weights—that is, weights that correct for unequal probability of selection, before any nonresponse adjustments.

Sample size and response rate for the SFA director survey includes SFAs in both the SFA-only and SFA-plus-school samples.

Data collection response rates reflect the percentage of eligible SFAs/schools that completed each instrument, given that the SFA/school had been recruited and agreed to participate.

# CHAPTER 4
# PROCESSING OF SCHOOL MENU DATA

To assess the food and nutrient content of reimbursable school lunches, breakfasts, and afterschool snacks, the data collected in the menu survey had to be entered into a nutrient analysis system that ultimately provided nutrient amounts for every item included on the menus. We used USDA's Survey Net system for this purpose. Survey Net includes nutrient values from the USDA National Nutrient Database for Standard Reference, Release 20 (Agricultural Research Service, Nutrient Data Laboratory, 2008). Because Survey Net was developed for the analysis of individual dietary intake data, we faced some challenges in using the system for processing school menu data. Most of these challenges were met through the creative use of existing data fields and training.[1] In addition, we developed a separate food grouping system to describe the foods offered in school menus. The food grouping scheme is described in Volume I of the report (Appendix B) and is not repeated here.

## A. Staffing and Training

Menu survey data were processed in Mathematica's Cambridge, Massachusetts office under the direction of a senior nutritionist. A team of 20 nutrition coders and 3 coding supervisors was recruited and hired locally. Supervisors had advanced nutrition degrees, previous research experience, and had worked with computerized nutrient analysis systems in the past. Coders had at least an undergraduate degree in nutrition or previous experience in foodservice, as well as a range of computer skills.[2]

All nutrition coders and coding supervisors were trained by the senior nutrition staff to use the Survey Net food coding system and on the specific procedures developed for processing the menu survey data. Four 8-to-10-hour training sessions were conducted to cover each of the main components of the data processing task. Two initial training sessions held on consecutive days covered the process of reviewing and editing the menu surveys. Two months later an additional two training sessions were conducted to instruct coders on entering menu surveys into Survey Net. Training procedures included group instruction and demonstration, supervised hands-on practice, and exercises to be completed and checked by the supervisors before beginning work with "live" data. Detailed training and reference manuals were provided.

Training sessions covered the review, editing, and data entry of the menu survey forms. Prior to familiarizing the coders with the various menu survey forms, some background information was provided, such as the concepts of reimbursable versus a la carte menu items, meal patterns/components, menu-planning systems, and quantity recipes. Coders were then trained to review and prepare the menu surveys for data entry and, subsequently, to enter the menu items,

[1] Some data fields in Survey Net that applied to dietary intake data, such as time of day, eating occasion, and where the food was obtained, were not needed for analysis of school menu data. Therefore, these fields were used for the entry of other information essential to the menu analysis, including daily meal counts and the number of reimbursable portions of each menu item served.

[2] Six of our most experienced TAs edited and coded one of the menu survey forms (the Self-Serve Bar Forms). TAs were trained and their work was supervised and reviewed by of one of the study's co-investigators. Self-Serve Bar Forms were then entered into Survey Net by nutrition coders.

portion sizes, recipe modifications, and meal and food count data into Survey Net.

## B. Coding Procedures

Completed menu surveys were forwarded to Mathematica's Cambridge office by TAs (see Chapter 3), after they had completed data retrieval and final editing. The surveys were logged into an Excel database as they were received, and tracked through each step of data processing. Coding supervisors assigned all surveys from a given SFA to the same coder because of the potential for similarities in the menus, recipes, and purchased products across schools.

### 1. Review and Editing

Each menu survey was reviewed in a systematic manner to identify occurrences of missing information, inconsistencies within and across the various forms, and instances where the number of reimbursable portions was not directly reported but could be calculated from the data provided. During the initial review, coders also identified unambiguous linkages between food items (for example, syrup served with pancakes) and commonly offered pre-prepared foods (for example, pizza, chicken nuggets, or burritos). Coders also assigned numerical codes, needed for data processing, to identify entrees and accompaniments. Questions regarding missing, unusual or ambiguous data provided on the menus survey (such as missing meal counts, unusually large portion sizes, and ambiguous linkages) were flagged by the coders for supervisor review. Six TAs were responsible for the specialized coding of self-serve salad bars and other food bars. A checklist was used to promote consistency across coders and to ensure all review and editing steps were completed.

#### a. Missing Data

Attempts were made to reconcile missing data problems by cross-referencing with other menu forms in the survey and with surveys completed by other schools within the same SFA.[3] For example, if a food description or the portion size of a food was vague or incomplete, coders checked if the same or a similar food was served on other days of the week and filled in the information accordingly. When it appeared that condiments had been omitted, coders checked the forms completed for other menu days to determine if the school usually offered condiments when they served certain items and added them, if appropriate. The same procedure was used for salad dressings served with salads. Incomplete or missing manufacturer or brand information was obtained from forms for other days on which the food was served or from menu surveys completed for other schools in the district that offered the same items.

When portion size information could not be obtained from other survey forms and in cases where the students served themselves, coders assigned a standard default portion size. The default portion sizes used for lunch and breakfast menus were based on those used in the SNDA-III study.[4]

[3] Nutrition coders did not directly contact school foodservice staff to inquire about menu information that was missing or needed clarification. However, supervisors did contact TAs, who were often able to answer the coders' questions.

[4] With the exception of salad dressing, default portion sizes for SNDA-III and SNDA-IV were the same as those used in SNDA-II (see Fox et al. 2001, Appendix E). In SNDA-III, the default portion size for salad dressing was increased from ¾ tablespoon (originally defined in SNDA-I) to 2 tablespoons. The revised default portion, which was also used in SNDA-IV, reflects the average portion of salad dressing consumed by school-age children in the Continuing Survey of Food Intakes by Individuals 1994–1996, 1998.

### b. Linked Menu Items

When a menu item, such as a topping or condiment, was clearly offered with another food item, the items were "linked" for analysis purposes.[5] Coders assigned special link codes to identify and categorize linked items. Salad dressings were always linked to salads. Other menu items were linked when the school foodservice manager reported offering the items together, as opposed to each item being available to all students (for example, spaghetti served with garlic bread, crackers served with salad, and rice served with stir-fried beef and vegetables).

Link codes were also assigned to the individual components of pre-plated meals, bag lunches, and multi-component foods to facilitate aggregation for nutrient analysis. A multi-component food was defined as a menu item for which ingredient and portion size information was provided, but which could not be entered into Survey Net as a single item. For example, chili cheese fries were not in the Survey Net database and could not be coded by modifying an existing recipe. Instead, this entree was entered as three separate items—french fries, cheese, and chili—and a link code was assigned to each item. Different link codes were assigned based on the types of foods being linked (for example, bread with additions and entrees with accompaniments).

### c. Pre-prepared School Foods

Schools use many commercially prepared (pre-prepared) foods that are formulated specifically for school foodservice, sometimes with more whole grains, less fat, more vitamins or minerals, or added protein. As a result, the nutrient content of the pre-prepared school foods reported on the menu surveys may not be accurately represented by a similar product in the Survey Net nutrient database. During the review of the menu surveys, coders entered pre-prepared foods into a centralized database for tracking the most commonly served pre-prepared school foods. Each pre-prepared food was then assigned to one of 70 pre-prepared food groups used to categorize foods based on similar nutrient content. When coding was completed, this list was used to obtain accurate information about nutrient content, as well as USDA food group equivalents, from USDA's Agricultural Research Service (ARS). The process of working with ARS to obtain these data is described later in this chapter.

[5] All condiments that could have been taken with more than one food (that is, there was no indication on the menu survey that a condiment was linked to a specific food) were considered "unlinked" and were not assigned special link codes.

### d. Self-Serve Food Bars

Coding the self-serve salad bars, theme bars (for example, Mexican, Italian, and potato bars), and condiment or fixins' bars was particularly challenging and was overseen by one of the study's co-investigators. By definition, students served themselves from these bars, there were few pre-portioned items, and the combinations of foods taken were not known. For example, entree salad bars offered the option to take different types of meats, cheeses, eggs, vegetables, and other items. It was unknown what types, combinations, and amounts of different food items each student truly selected from the food bar. Therefore, in order to define an average serving, detailed coding rules were developed for each type of food bar and for each meal component offered on the food bar, using a methodology employed by the previous SNDA studies. This approach assumes that students are offered everything on the bar and assigns default portion sizes to individual items on the bar based on minimum portions required for each specified meal component in food-based menu planning or on default portion sizes for items such as condiments and toppings.

### e. Production Records

Some schools were unable or unwilling to complete the menu survey forms. To facilitate participation in these schools, we agreed to accept production records in the place of the menu survey forms. This accommodation was only made when the production records were detailed enough to provide essentially the same data as the menu survey forms and/or when SFA directors or FSMs were willing to provide missing information during followup contacts. A total of 55 schools in the final sample provided production records rather than completed menu surveys. For one of these schools, data on the number of portions served in reimbursable meals were not provided. This school had to be excluded from the analysis of meals served, leaving a total of 54 schools with production records included in the analysis.[6]

Production records provided by some schools were very similar in structure to the Reimbursable Foods Form and provided information about the number of individual portions of each menu item served in reimbursable meals. However, other schools provided information about foods served to students as information about the total quantities of food prepared and left over. In these instances, nutrition coders had to convert the data on bulk quantities to estimates of the number of individual portions. For example, if the form indicated that 30 pounds of raw carrots were prepared, 2 pounds were left over, and the portion size was ¼ cup cooked carrots, the coder calculated the number of ¼ cup servings of cooked carrots that 28 pounds of raw carrots would yield. Coders used the USDA Food Buying Guide for Child Nutrition Programs (U.S. Department of Agriculture, 2008) and measurement equivalents and conversion charts to minimize errors. After these calculations were completed, coders compared the total numbers of reimbursable servings of entrees and milk, and the number of servings of individual menu items to the total number of meals reported for that day. Large discrepancies were flagged for supervisor review to ensure they were not due to miscalculation of the number of portions served.

[6] One school provided production records that were too incomplete to substitute for the menu survey. This school was ultimately considered a nonresponder for the menu survey component of the study.

## 2. Entering Data into Survey Net

After a menu survey was reviewed, edited, and cross-checked by a supervisor or lead coder, it was ready for entry into Survey Net. Coders entered the information using procedures developed specifically for this study (building on the procedures used in SNDA-III). A separate file was created for each school, with separate records for each daily lunch and breakfast menu. Food items from the Reimbursable Foods Form were matched to the closest food in the database, considering characteristics such as the form of the food (fresh, canned, frozen), the preparation method (baked or fried), and characteristics that might affect nutrient content—particularly fat (regular versus low-fat or nonfat versions). To expedite the process of selecting the appropriate item in the database, coders were provided with search terms and food codes for commonly served foods. Information on portion size (reported or the assigned default) and the total number of reimbursable portions served was also entered for each menu item. In addition, for selected menu items, the link codes and entree and accompaniment flags that were added during editing, along with any special instructions pertaining to how a food should be treated in the analysis, were entered into Survey Net.

A set of coding guidelines was developed to assist coders and standardize entry of foods that were not thoroughly described. These guidelines were designed to reflect common school foodservice practices, which did not always correspond to the Survey Net "not further specified" option that is typically used in coding such foods. For example, if a school reported serving cooked carrots but did not specify whether fat was added in cooking, the options for entering the carrots into Survey Net included fat added, no fat added, and not further specified (NFS), which assumes fat was added. The menu coding guideline for this scenario was to assume that fat was not added (that is, select the "cooked carrots, fat not added" code).

Special procedures were developed for entering school recipes, self-serve food bars, and pre-prepared school foods (discussed in the next three subsections). For self-serve food bars and pre-prepared school foods, "placeholder" food codes were entered in the Survey Net menu files to flag the items for subsequent replacement of nutrient data.

### a. Dealing with Recipes

Survey Net was not designed to allow users to add recipes to the database. However, existing recipes can often be modified to more closely match the foods reported. Coders followed specific guidelines to decide if recipe modification was appropriate. These guidelines (summarized in Table 4.1) were developed for and used in SNDA-III and were based on guidelines provided by USDA's Food Survey Research Group.

The decision to modify a recipe was based primarily on the importance of the modification to the overall fat content of the food and presence of whole grains. For example, if the school provided a recipe for a ham and cheese sandwich that was comprised of turkey ham and reduced-fat cheese, an existing recipe for a ham and cheese sandwich was modified to account for the lower-fat foods included in the school's recipe. Another consideration was the amount of the meat/meat alternate in school-prepared sandwiches, entree salads, and some Mexican foods, compared with the standard recipes for these foods in Survey Net. Single serving recipes for sandwiches, Mexican entrees and entree salads were modified when the amount of meat, cheese or bread provided in the school recipe differed from the Survey Net recipe by more than one-half ounce. When modified recipes were created, the ingredients and/or amounts that were changed were noted in the name assigned to the new recipe.

**Table 4.1. Recipe Modification Guidelines**

| Menu Item | Allowed Modifications to *Type* of Ingredient | | | | | |
|---|---|---|---|---|---|---|
| | Type of Fat | Type of Meat | Type of Cheese | Type of Bread/grain | Type of Milk | Type of Mayonnaise or Salad Dressing |
| Vegetables, Dry Beans or Peas | √ | | | | | |
| Rice and Pasta | √ | | | | | |
| Eggs and Omelets | √ | | | | | |
| Whipped/Mashed Potatoes | √ | | | | √ | |
| Garlic Bread | √ | | | | | |
| Cooked Cereal | √ | | | | √ | |
| Macaroni and Cheese, Other Mixed Dishes with Cheese and Grain | √ | | √ | √ | √ | |
| Mixed Dishes with Meat and Grain | | √ | √ | √ | | |
| Pudding or Cream Soups | | | | | √ | |
| Salads – NOT Lettuce-based | | | | | | √ |
| Sandwiches | √ | √ | √ | √ | | √ |
| Mexican Entrees | √ | √ | √ | √ | | |
| Entree Salads | | √ | √ | | | |

| Menu Item | Allowed Modifications to Ingredient *Amounts*[a] | | | |
|---|---|---|---|---|
| | Amount of Meat/ Meat Alternate[b] | Amount of Cheese[c] | Amount of Bread/Grain[d] | Higher-fat Ingredients[e] |
| Sandwiches | √ | √ | √ | √ |
| Mexican Entrees | √ | √ | √ | √ |
| Entree Salads | √ | √ | | √ |

[a] Modifications to ingredient amounts were made only when the school recipe and the Survey Net recipe were single-serving recipes.

[b] Amounts of meat/meat alternates were modified only if the difference between the school recipe and the Survey Net recipe was more than ½ oz.

[c] Amounts of cheese were modified only if the difference between the school recipe and the Survey Net recipe was more than ½ oz.

[d] Amounts of bread/grain were modified only if the difference between the school recipe and the Survey Net recipe was more than ½ oz.

[e] Higher-fat ingredients (butter, margarine, mayonnaise, salad dressing, cheese) were deleted from Survey Net recipes if they were not included in school recipes.

There were limits to the feasibility of modifying recipes depending on how the recipe existed in Survey Net. For single-serving recipes (for example, recipes for sandwiches), both the amounts and types of ingredients could be modified easily. However, for recipes that yielded more than one serving, modifications were limited to ingredient substitutions. Changes to ingredient amounts could not be made because there was no way to account for the effect on the recipe's yield. Complications also arose when changing the type of meat in a quantity recipe. The form of the food (raw versus cooked) to be substituted was not always comparable to what was in the recipe. For example, cooked ground turkey (the only form of ground turkey in Survey Net) could not be substituted for raw ground beef in a recipe due to the effect on fat and moisture losses. In order to calculate the yield of a recipe, Survey Net takes into account the moisture and fat retention of each ingredient after cooking. Substituting a different form of an ingredient and/or altering the ingredient amount in quantity recipes would have required entering retention factor codes for each altered ingredient, which is not a simple or straightforward process.

### b. Self-Serve Food Bars

Each unique self-serve food bar was entered separately from the rest of the menu survey, as if it were a "menu" of all of the food items offered on the bar. Default portion sizes were assigned to individual items on the bar based on the minimum portions required for specific meal components in food-based meal patterns: fruits/vegetables, bread/grain products, and meat/meat alternates. (Milk was not usually included on food bars.) For non-meal-pattern food items, such as condiments, toppings, salad dressings, and desserts, the same default portion sizes were used as for self-serve menu items not on bars.

If more than one option within a meal component group was offered, a recipe was created for the meal component group. The recipe "ingredients" consisted of a full portion of each item from the meal component group available on the food bar, and the recipe yield (number of servings) equaled the total number of items or ingredients. For example, a sandwich bar offered a choice of turkey, ham, or tuna, and a choice of white bread, a hoagie roll, or wheat bread. The recipe created to represent one average serving of meat from the bar would have a yield of three servings (since there are three meat ingredients). The coding rules for a sandwich bar also called for *two* average servings of breads/grains. In cases where the coding rules called for more than one serving from a meal component group, the yield of the recipe was equal to the total number of ingredients, divided by the desired number of servings. Thus, in this example, an average serving of breads/grains would have a yield of 3 bread/grain choices divided by 2 servings, or 1.5 servings. An average serving from the entire self-serve bar was the simple sum of the average nutrients per serving for each of the meal components included in the bar.

### c. Imputing Missing Data on the Number of Portions Served

Many reported accompaniments (condiments, salad dressings, and toppings) were missing data on the number of portions served. This was mainly due to the nature of the data being reported as "self-serve." For linked accompaniments (for example, salad dressings and accompaniments such as cheese on broccoli or toppings on a taco) data on the number of portions served was imputed based on the number of portions reported for the menu item to which the item was linked. For unlinked accompaniments, data on the number of portions served was imputed based on the mean/median number of servings of accompaniments per meal, in schools that provided servings data.

### 3. Pre-Prepared School Foods

Since manufacturer food labels were not collected from individual schools, nutrient and ingredient information for pre-prepared school foods was researched on the Internet and obtained from selected manufacturers. The most frequently logged items were selected for additional research on nutrient and ingredient information by contacting manufacturers. Seventy pre-prepared food-type groups were created to identify which products needed further research. Food-type groups were defined as foods that seemed essentially "the same" based on their food description and any nutrients available. For example, four pre-prepared food-type groups were created to capture each type of cheese pizza served in schools, "cheese pizza," "cheese pizza reduced fat," "cheese pizza whole grain" and "cheese pizza reduced fat, whole grain." Two hundred of the most commonly reported pre-prepared foods, at least one for each of the 70 food-type groups, were sent to ARS for further analysis. ARS returned complete nutrient and food group profiles for each food. These data were used to replace the profiles for the placeholder foods that had been used in coding the menus.

### 4. Quality Control Procedures

During the initial phases of menu data processing, supervisors reviewed each coder's editing and entry for one SFA (three to four schools). Coders received detailed feedback and the process was repeated until a level of accuracy greater than 90 percent was achieved. In addition, during the editing phase, each menu survey was cross-checked by a second coder and any discrepancies were resolved by supervisors.

Menu data entry was also carefully reviewed by supervisors to ensure that the appropriate food selections were made from the database, portion sizes were entered correctly, coding rules were applied when necessary, and recipe modification guidelines were followed. Overall, full quality review checks were conducted for 15 percent of all menu surveys. A similar procedure was followed for the quality review of coding and entry of self-serve food bars. Every recipe modification created by coders was individually reviewed by a coding supervisor. Recipes were checked for compliance to guidelines and approved when acceptable. Incorrect or unnecessary recipe modifications were adjusted or deleted.

In order to maintain standardized procedures, coders attended periodic meetings and received "coder updates" clarifying issues that were identified or changes to procedures. Throughout the editing and entry phases, coders documented issues that arose in a central location, which facilitated supervisor followup. Coding guidelines were updated regularly, and coders were required to review updates each day. Supervisors were available at every shift to answer questions and resolve emerging issues. The senior nutritionist met weekly with coding supervisors to discuss coding progress and resolve coding issues they needed help with.

After all of the menu information was entered in Survey Net, a set of detailed data checks were performed to identify potential coding errors. Problem cases were identified, and each was reviewed and corrected by coding supervisors. The cleaning runs included the following types of checks:

- **Basic Data Integrity.** Daily menus were checked for missing data, duplicate entries, and valid values for the following items: school ID numbers, consecutive menu days and dates, minimum number of meal components, and number of meals served. Individual menu items were checked for valid and non-missing portion sizes, number of portions served, appropriate linking codes, and entree and accompaniment identifiers. Problem

cases were identified and checked against hard copy menu surveys, and corrections were made as necessary. Afterschool snack forms were checked for the minimum of at least two snack items offered each day. Self-serve food bars were also checked for valid values and the minimum expected items based on the type of bar.

- **Out-of-Range Menu Items.** Estimated per-serving nutrient values for individual foods were reviewed for calories, total fat, and sodium to identify possible outliers. Foods with nutrient values that were below the 5th or above the 95th percentile were identified and checked against hard copy menu surveys to verify entry. Corrections were made where appropriate. The same procedure was followed for self-serve food bars, with the assumption that the total nutrients for any particular food bar "menu" would be reasonably close to the expected range for one serving from the particular meal component group in which the bar falls. For example, the range of nutrients for entree salad bars, Mexican bars, and sandwich bars should have approximated the nutrients for other "entrees" on the main menus.

- **Over-Reporting of Portions Served.** Checks were run to identify cases where the number of servings reported for a menu item was greater than the total number of meals served. The number of servings of milk, side salad bars, french fries and tater tots, desserts, entrees, and salad dressings were adjusted to ensure that the weighted analysis would not overestimate the nutrient content of meals served. This was based on the assumption that students generally select no more than one milk, one entree and one serving of any particular side item per reimbursable meal. In addition, instances where the number of portions served for french fries, tater tots, and salad dressings was greater than the number of meals served were often cases where the manager provided the information as bulk amounts prepared and left over.

This page left blank for double–sided printing.

# CHAPTER 5
# CALCULATION OF SAMPLE WEIGHTS

All of the analyses conducted for SNDA-IV report were weighted to produce estimates that are representative of public SFAs or schools participating in the NSLP in the 48 contiguous States and the District of Columbia.[1] Given the complex, multi-stage sample design, as described in Chapter 2, sample weights adjust both for unequal probabilities of selection at each stage of sampling and for nonresponse at each stage of data collection. Weights were constructed at two levels: SFAs and schools. The weights at the two levels are not independent—the final SFA base weight adjusted for SFA participation served as the initial weight at the school level. At each level, two sets of weights were constructed, one to represent SFAs or schools, and one to represent the students enrolled in the SFA or schools.

One set of weights was adequate for the data collected at the SFA level—the SFA director survey. However, because several data collection instruments were used at the school level (see Chapter 3) and schools did not necessarily complete all instruments, the weights for use in analysis of school-level data had to be adjusted to reflect school nonresponse to different instruments.

## A. SFA-Level Weights

As discussed in Chapter 2, two samples of SFAs were selected, the SFA-only sample and the SFA-plus-school sample. Data from these samples were weighted separately and then a "composite" weighting factor was used to combine SFA directory survey cases from the two samples. For each sample, the initial weight at the SFA level is the sampling weight, which starts as the inverse of each PSU's probability of selection into the initial sample. At this point, each PSU included one or more school districts (not all districts are SFAs). Within each sample, the weights also incorporate: (1) adjustments for the selection of SFAs in multi-SFA PSUs; (2) adjustments for the release and participation of SFAs within PSU pairs; (3) nonresponse adjustments not accounted for by the PSU pair adjustments; and (4) poststratification. After making these adjustments, the weights from the two samples were combined using a compositing factor, then adjusted for nonresponse to the SFA director survey.

### 1. Initial Weights

The initial weight for the *kth* PSU in sample *j* (*j* identifying the SFA-only or SFA-plus-school sample) is:

$$SFAWGT1_{jk}=SWF1_{jk}*SWF2_{jk}$$

$SWF1_{jk}$ is the inverse of PSU *k*'s probability of being selected for frame *j*. Some large PSUs were selected with certainty for the SFA-plus-school frame; for these $SWF1_{jk}$=1.0. For all other PSUs, $SWF1_{jk}$=2.0 (since half of the PSUs not selected with certainty were assigned to each of the two frames).

[1] Analyses focused on the supplementary sample of HUSCC schools were an exception. Estimates for HUSSC schools were not weighted because the sample was not nationally representative.

$SWF2_{jk}$ adjusts for probability of selection into the SNDA-IV sample within each of the two frames and varies according to how the SFA was selected into the sample. Selection within the two SNDA-IV frames took place in three phases: (1a) selection from the SFA-plus-school frame of 640 PSUs, 86 with certainty and 544 with PPS, and (1b) the selection from the SFA-only frame of 642 PSUs with PPS; (2) within sampled PSUs with more than one SFA, random selection of one of those SFAs resulting in samples of SFAs (or potential SFAs) within each frame; and (3) the pairing of the selected SFAs and release of one or both for each contact.

For defining $SWF2_{jk}$, the SFAs in the SFA-plus-school frame sample were divided into two groups based on how they were selected into the sample. The groups were:

1. SFAs selected with certainty in the initial sample and into the main sample
2. SFAs that were paired and randomly selected to be released. Thus:

$$SWF2_{jk}=SWF2a_{jk}*SWF2b_{jk},$$

where $SWF2a_{jk}$ is the inverse of the probability of selection into the initial sample and $SWF2b_{jk}$ adjusts for release from a given pair. These terms are defined as follows for the two groups:

1. For those selected with certainty into the SFA-plus-school sample, $SWF2_{jk}=1$. For these SFAs, $SWF2a_{jk}$= 1 because of selection with certainty and $SWF2b_{jk}$ =1 because these SFAs were not placed into pairs (all were released).
2. For the non-certainty selections, $SWF2a_{jk}$ reflects the chance of being selected from the initial sample, and $SWF2b_{jk}$ is a pair adjustment. *SWF2b* takes on the value of 0, 1, or 2 and adjusts for selection into the sample as part of a pair, release within the pair, and nonresponse within the pair. The values of $SWF2b_{jk}$ for non-certainty SFAs are presented in Table 5.1. The sum of $SWF2b_{jk}$ for a pair will always equal 2. When only one district in a pair was released, $SWF2b_{jk}$ reflects subsampling within the pair; if both were released, the weight reflects no subsampling within the pair. If one of the pair was not completed, $SWF2b_{jk}$ adjusts for nonresponse within the pair.

**Table 5.1. Values of $SWF2b_{jk}$ for Non-certainty SFAs**

| Within a Pair | | |
|---|---|---|
| Number Released | Recruited | *SWF2b* |
| 1 | 0 | 2 for the released district (based on *1/p*; *p=1/2*); 0 for the other |
| 1 | 1 | 2 for the released district (based on *1/p; p=1/2);* 0 for the other |
| 2 | 0 | 1 for each of the districts |
| 2 | 1 | 2 for the completed district (*1/p x 1/rr where p=1/2 and rr=1/2*); 0 for the other |
| 2 | 2 | 1 for each of the districts |

## 2. Nonresponse Adjustment

For both samples, the next step was to form cells to adjust for nonresponse (not already accounted for by $SWF2b_{jk}$). For those selected with certainty into the main sample (group 1 above), only one weighting cell was used. But for other SFAs, the nonresponse weighting cell was the reserve zone within the sample (SFA-only or SFA-plus-school).[2] $SWF3\_c_j$ is the nonresponse adjustment factor with cell c:

$$SWF3\_c_j = \frac{\sum_{releasedSFAs \in c} SFAWGT1_{jk}}{\sum_{completedSFAs \in c} SFAWGT1_{jk}}$$

The values of $SWF3\_c_j$ are shown in Table 5.2. These weight factors are the inverse of the weighted response rate for each reserve zone. The SFA weight adjusted for nonresponse is: $SFAWGT\_NR_{jk}=SFAWGT1_{jk}*SWF3\_c_j$.

**Table 5.2. SFA Nonresponse Adjustment Factor**

| Reserve Zone | SFA Nonresponse Adjustment Factor ($SWF3\_c_j$) | |
|---|---|---|
| | SFA-Plus-School | SFA-Only |
| Certainty | 1.111111 | NA |
| 1 | 1.875000 | 1.214286 |
| 2 | 1.066667 | 1.117647 |
| 3 | 1.285714 | 1.133333 |
| 4 | 1.214286 | 1.148718 |
| 5 | 1.250000 | 1.214286 |
| 6 | 1.307692 | 1.250000 |
| 7 | 1.214286 | 1.000000 |
| 8 | 1.133330 | 1.156846 |
| 9 | 1.214286 | 1.545455 |
| 10 | 1.133333 | 1.000000 |
| 11 | 1.000000 | 1.000000 |
| 12 | 1.076923 | 1.230769 |
| 13 | 1.250000 | 1.066667 |
| 14 | 1.750000 | 1.000000 |
| 15 | 1.250000 | 1.307692 |
| 16 | 1.214286 | 1.235294 |
| 17 | 1.214286 | 1.062500 |
| 18 | 1.000000 | 1.000000 |
| 19 | 1.250000 | 1.134454 |
| 20 | 1.357143 | 1.071429 |
| 21 | 1.307692 | 1.125000 |

[2] After the initial samples of PSUs were selected and pairs formed, the file was sorted based on the sort variables used in the sampling and 21 zones were defined, each containing 15 or 16 pairs of PSUs. One pair was randomly selected within each zone to serve as a replacement in case of nonparticipation of both PSUs in a pair.

### 3. Poststratification

The SFA weights were ratio-adjusted (poststratified) so that the weighted total of the completed sample matched that of our estimated total of SFAs on the SNDA-IV sampling frame. The target total was 15,633. The poststratified SFA weight is:

$SFAWGT\_PS_{jk}=SFAWGT\_NR*RAF_{SFA}$.

Where:

$$RAF_{SFA} \text{ (ratio adjustment factor)} = \frac{15,633}{\sum_{k \in CompletedSFAs} SFAWGT\_NR_{jk}}$$

After this adjustment, the weights for sampled and recruited SFAs from each frame summed to the population total of SFAs.[2] The weights for the SFA-plus-school sample served as the base for the school-level weights. Weighting adjustments for SFA-level survey data then incorporated a factor to combine the samples from the two frames, which is discussed next.

### 4. SFA Director Survey Weight

The SFA director survey had its own level of nonresponse and required further weighting. The survey weight involved a nonresponse adjustment and a composite weight adjustment to bring the two SFA samples together. The poststratified SFA weight was the starting point. For the SFA-only sample, no nonresponse adjustment was necessary, because these SFAs were not recruited into the study so there was no additional nonresponse within these SFAs. For the SFA-plus-school sample, weighting cells were constructed using the FNS region and SFA size. For SFA size, two categories were defined: large SFAs were those with more than 10 schools and small SFAs were those with 10 or fewer schools. *SWFDir3_c* is the nonresponse adjustment for the SFA director survey. The values of the adjustment are shown in Table 5.3.

[2] The total for the SFA-plus-school sample was slightly higher because it contains the certainty selection.

**Table 5.3. SFA Director Survey Nonresponse Adjustment Factor**

| Region | Large | Nonresponse Adjustment Factor (*SWFDIR3_c*) |
|---|---|---|
| West | 1 | 1.028571 |
| West | 0 | 1.142857 |
| Southwest | 1 | 1.045455 |
| Southwest | 0 | 1.045455 |
| Southeast | 1 | 1.000000 |
| Southeast | 0 | 1.083333 |
| Northeast | 1 | 1.125000 |
| Northeast | 0 | 1.058824 |
| Mountain | 1 | 1.000000 |
| Mountain | 0 | 1.000000 |
| Midwest | 1 | 1.000000 |
| Midwest | 0 | 1.156250 |
| Mid-Atlantic | 1 | 1.100000 |
| Mid-Atlantic | 0 | 1.153846 |
| Certainty | | 2.625000 |

The SFA director survey weights for interviews from both SFA samples were combined using a composite weighting factor. The composite factor (*compadj*) was set to:

- 1.0 for those selected with certainty for the SFA-plus-school sample
- L for the SFA-only sample (O<L<1)
- (1-L) for those in the SFA-plus-school sample that were not selected with certainty

L was set to minimize the variance of the combined samples.

The SFA director survey weight is

$SFAdirWT_{jk}=SFAWT\_PS\_{jk}*SWFDir3\_c*compadj$ where:

$Deff_{sch}$=the estimated design effect for the SFA-plus-school sample

$Deff_{SFA}$=the estimated design effect for the SFA-only sample

$n(Dir)_{sch}$= the number of cases responding to the SFA director survey for SFA-plus-school sample

$n(Dir)_{SFA}$=the number of cases responding to the SFA director survey for the SFA-only sample

$neff_{sch}=n(Dir)_{sch}/Deff_{sch}$

$neff_{SFA}=n(dir)_{SFA}/Deff_{SFA}$

$L=neff_{sch}/(neff_{sch}+neff_{SFA})$.

## B. School-Level Weights

### 1. Initial Weights

The initial weight for school *i* in stratum *h* and $SFA_k$ is the variable $SFAWGT\_PS_{jk}$ for the SFA to which the school belongs. Since schools were only selected from SFAs in the SFA-plus-school sample, no composite adjustment was necessary. These initial weights were first adjusted for probability of selection of schools within the SFA, using two factors. The first adjustment factor, $W1_{ihk,}$ is the inverse of the probability of the first phase of selection of the school within its SFA:

$$W1_{ihk}=1/Psel_{ihk}$$

where:

$$Psel_{ihk} = n'_{hk}/ N'_{hk}$$

$n'_{hk}$ is the number of school selections made in stratum *h*, $SFA_k$

$N'_{hk}$ is the number of schools available for with PPS in stratum *h* and $SFA_k$

The next factor, $W2_{ihk,}$ accounts for subselection into the main and alternate samples. If there was no subselection within SFA (that is, if there was only one selection or all selections were treated as main), then $W2_{ihk}$=1.0. In other cases, the value of $W_{2ihk}$ would be 1 or 2, depending on the numbers released and cooperating within pairs, following the same pattern that was used for SFA pairs as shown in Table 5.1.

The initial school-level weight, before adjustment for nonparticipation (not already accounted for in the pair adjustment) is:

$$SCHWGT1_{ihk} = SFAWGT\_PS_j * W1_{ihk} * W2_{ihk}.$$

The nonparticipation adjustment factor is:

$$W3_c = \frac{\sum_{ihk\in(resp,c)} SCHWGT1_{ihk} + \sum_{ihk\in(nonrespjc)} SCHWGT1_{ihk}}{\sum_{ihk\in(resp,c)} SCHWGT1_{ihk}}$$

where the numerator is the sum of the initial school-level weights across participating and nonparticipating schools, and the denominator is the sum of these weights for the participating schools only.

The school-level weight, adjusted for nonparticipation, is $SCHWGT\_NR_{ihk}=SCHWGT1_{ihk}*W3_c$ for participating schools.

## 2. Poststratification

Finally, the school weights were ratio-adjusted so that the sum of weights for participating schools was 83,389, the best estimate of the number of schools in SFAs offering the NSLP or the SBP. Thus,

$$RAF_{school} = \frac{83{,}389}{\sum_{ihk \in complete} SCHWGT\_NR_{ihk}}$$

and

$$SCHWGT\_PS_{ihk} = SCHWGT\_NR_{ihk} * RAF_{school} .$$

## 3. Survey-Specific Weights

There were several school-level surveys. For each survey, separate school nonresponse adjustments were needed. Each survey started with the initial school weight and was then adjusted for nonresponse by weighting cells, and then poststratified to equal 83,389 as was done with the initial school weight.

The following weights were developed for use with the various school-level data files:

- **School-Level Data Collected in the SFA Director Survey.** For the school-level data collected in this survey, weighting cells were created using region, school level (elementary, middle, high), and SFA size (large or not).
- **Menu Survey, Foodservice Manager Survey, and Daily Meal Counts Form.** A single weight was created for these two surveys and the daily meal counts form (a component of the menu survey) because their nonresponse patterns were very similar. A school was considered a respondent if it completed either the menu survey or the foodservice manager survey. For the weighting cells, region, school level (elementary, middle, high) and size (large or not) were used.
- **Principal Survey.** For the weighting cells, region, school level (elementary, middle, high) and size (large or not) were used.
- **Competitive Foods Checklists.** For the three competitive foods checklists (a la carte, vending machine, and other sources of foods and beverages), the nonresponse adjustment required the use of the Chi Square Automated Interaction Detection (CHAID) branching logic procedure to determine the best combinations of variables to form weighting cells. CHAID allowed us to identify the variables that had the greatest influence upon nonresponse and use these to create the weighting cells.

  For the vending machine checklist, the weighting cells were created using the school level (elementary, middle, high). For the other sources of foods and beverages checklist, the weighting cells were created using the concentration of black students (high or low),

and size of the SFA to which the school belonged (large or not).[4] For the a la carte checklist, only two weighting cells were created and these were based on region (Mid-Atlantic region or not). Once again, each of these began with the school-level initial weight which was then adjusted by previously stated weighting cells.

Additional weights were required for the vending machine and a la carte checklists to adjust for nonresponse among schools that indicated that they had vending machines or sold a la carte foods and beverages but did not complete the portion of the checklist that identified the specific foods and beverages available.[5] For the vending machine checklist, we formed weighting adjustment cells based on the number of vending machines reported (1 machine, 2 machines, or more than 2 machines). For the a la carte checklist, we formed cells based on quartiles of reported a la carte revenue.

- **Afterschool Snack Menu Survey.** The afterschool snack menu survey was not provided to all schools because some schools did not provide afterschool snacks. As such, we did no poststratification adjustment because we do not know how many schools nationally provide afterschool snacks through the NSLP. For the nonresponse adjustment, CHAID was used to identify the most appropriate weighting cells. The final weighting cells created were based on the percentage of reduced price or free lunches that a school served (high or not).

Each of these weights (for survey *s*) is identified as $SCHWGT\_PS_{sihk,}$ and was derived in the same manner as $SCHWGITP\_S_{ihk}$, described above.

## C. Student-Enrollment-Adjusted Weights

For both SFA- and school-level instruments, we created weights adjusted to the student population (enrollment). We start with the final school-level weight for each survey (*s*) in school *i* in stratum *h* in SFA *k* is $SCHWGT\_PS_{sihk}$, the poststratified school-level weight. The school-level weight is then adjusted for the number of students that attended the school, which gave the student-level weights. Thus the enrollment adjusted weight was, for each survey:

$$ENRWGT_{sihk} = SCHWGT\_PS_{sihk} \left(enrollment_{ihk}\right)$$

where $enrollment_{ihk}$ is the number of students enrolled.

[4] Variables used in developing weighting classes do not have to be limited to those used in defining sampling strata. Use of the concentration of black students was indicated by the CHAID analysis High concentration was defined as greater than 25 percent of students; the percentage was estimated from the CCD.

[5] Comparable weights were not developed for the other sources of foods and beverages checklist because the sample of schools that reported these alternative sources of competitive foods (school stores and snack bars) was too small to produce reliable estimates.

## REFERENCES

Fox, M.K., E. Condon, M.K. Crepinsek, et al. "School Nutrition Dietary Assessment Study-IV: Volume I: School Foodservice Operations, School Environments, and Meals Offered and Served." Alexandria VA: U.S. Department of Agriculture, Food and Nutrition Service, 2012.

U.S. Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory. "USDA National Nutrient Database for Standard Reference, Release 20." Alexandria, VA: U.S. Department of Agriculture, 2008. Available at [http://www.ars.usda.gov/Services/docs.htm?docid=17476].

U.S. Department of Agriculture, Food and Nutrition Service. "USDA Food Buying Guide for Child Nutrition Programs." Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service, 2008.

This page has been left blank for double-sided copying.

**United States Department of Agriculture**
Food and Nutrition Service
3101 Park Center Drive
Alexandria, VA 22302

**Visit us at www.fns.usda.gov/fns**