

Technical Notes

Technical Notes — Survey on Foods and Physical Activity in Public Elementary Schools: 2005

Fast Response Survey System

The Fast Response Survey System (FRSS) was established in 1975 by the National Center for Education Statistics (NCES), U.S. Department of Education. FRSS is designed to collect issue-oriented data within a relatively short time frame. FRSS collects data from state education agencies, local education agencies, public and private elementary and secondary schools, public school teachers, and public libraries. To ensure minimal burden on respondents, the surveys are generally limited to three pages of questions, with a response burden of about 30 minutes per respondent. Sample sizes are relatively small (usually about 1,000 to 1,500 respondents per survey) so that data collection can be completed quickly. Data are weighted to produce national estimates of the sampled education sector. The sample size permits limited breakouts by classification variables. However, as the number of categories within the classification variables increases, the sample size within categories decreases, which results in larger sampling errors for the breakouts by classification variables.

Sample and Response Rates

The sample for the survey consisted of 1,198 regular public elementary schools in the 50 states and the District of Columbia. It was selected from the 2002–03 NCES Common Core of Data (CCD) Public School Universe file, which was the most current file available at the time of selection. The sampling frame included 50,980 regular elementary schools. For the purposes of the study, an elementary school was defined as a school with a high grade of 1 to 8 and a low grade of prekindergarten, kindergarten, or grades 1 to 3. Excluded from the sampling frame were schools with a high grade of prekindergarten or kindergarten and ungraded schools, along with special education, vocational, and alternative/other schools, schools outside the 50 states and the District of Columbia, and schools with zero or missing enrollment.

The public school sampling frame was stratified by enrollment size (less than 300, 300 to 499, 500 to 599, 600 to 749, and 750 or more) and percent eligible for free or reduced-price lunch (less than 35 percent, 36 to 49 percent, 50 to 74 percent, and 75 percent or more). Schools in the frame were then sorted by type of locale (city, urban fringe, town, and rural) and region (Northeast, Southeast, Central, and West) to induce additional implicit stratification.

All survey data were reported for spring 2005 when the survey was conducted. Questionnaires and cover letters for the study were mailed to the principal of each sampled school in early March 2005. Principals were told that the survey was designed to be completed by the person most knowledgeable about the availability of foods and opportunities for physical activity at the school. Respondents were encouraged to consult with the school's food service personnel and physical education staff to complete relevant sections of the questionnaire, as necessary. Telephone follow-up for survey nonresponse and data clarification was initiated in late March 2005 and completed in late June 2005. During data collection, 37 schools were found to be ineligible for the survey because they were closed or did not meet the grade requirements for inclusion as an elementary school. For the eligible institutions, the response rate was 91 percent (1,055 responding institutions divided by the 1,161 eligible schools in the sample). The weighted response rate was also 91 percent. Of the schools that completed the survey, 19 percent

completed it by Web, 53 percent completed it by mail, 27 percent completed it by fax, and 1 percent completed it by telephone.

Although item nonresponse for key items was very low (less than 1 percent), missing data were imputed for the 32 items with a response rate of less than 100 percent. The missing items included both numerical data such as total minutes per day of scheduled recess, as well as categorical data such as whether soft drinks were available at vending machines. The missing data were imputed using a “hot-deck” approach to obtain a “donor” school from which the imputed values were derived. Under the hot-deck approach, a donor school that matched selected characteristics of the school with missing data (the recipient school) was identified. The matching characteristics included enrollment size, percent of students in the school eligible for free or reduced-price lunch, and type of locale. In addition, relevant questionnaire items were used to form appropriate imputation groupings. Once a donor was found, it was used to obtain the imputed values for the school with missing data. For both categorical and numerical items, the imputed value was simply the corresponding value from the donor school. All missing items for a given school were imputed from the same donor. Imputation flags are included in the data.

Weighting Procedures and Sampling Errors

The response data were weighted to produce national estimates (see table 1). The weights were designed to adjust for the variable probabilities of selection and differential nonresponse. FRSS survey data are based on complex sample designs that require the use of weights to compensate for variable probabilities of selection, differential response rates, and possible deficiencies in the sampling frame. The reciprocal of the probability of selection, referred to as the “base weight,” will produce unbiased (or consistent) estimates of population totals and ratios if there is no nonresponse in the survey. Since a stratified sample design was employed for the survey, the base weight for the i -th school in stratum h was computed as $w_{hi} = 1/f_h$ where f_h is the overall sampling rate used to select schools in stratum h .

Although the survey had a high response rate, adjustment of the base weights was necessary to compensate for the survey nonrespondents (i.e., whole questionnaire or unit nonresponse). To compensate for unit nonresponse, an adjustment factor was computed within selected weighting classes, and this factor was then used to inflate the base weights of the schools in the weighting class. The weighting classes used for this purpose were subsets of the sample defined by enrollment size class, poverty status, and type of locale. Within each subset, a nonresponse-adjustment factor was computed as the inverse of the base-weighted response rate. The factor was then applied to the base weights of the responding schools in the subset to obtain the final nonresponse-adjusted weight.

The findings in the survey report titled *Calories In, Calories Out: Food and Exercise in Public Elementary Schools, 2005* are estimates based on the sample selected and, consequently, are subject to sampling variability. The standard error is a measure of the variability of an estimate due to sampling. It indicates the variability of a sample estimate that would be obtained from all possible samples of a given design and size. Standard errors are used as a measure of the precision expected from a particular sample. If all possible samples were surveyed under similar conditions, intervals of 1.96 standard errors below to 1.96 standard errors above a particular statistic would include the true population parameter being estimated in about 95 percent of the samples. This is a 95 percent confidence interval. For example, the estimated percentage of public elementary schools that sold foods to generate funds is 36.3 percent, and the standard error is 1.5 percent (tables 2 and 2a). The 95 percent confidence interval for the statistic extends from $[36.3 - (1.5 \times 1.96)]$ to $[36.3 + (1.5 \times 1.96)]$, or from 33.4 to 39.2 percent. The coefficient of variation (“c.v.,” also referred to as the “relative standard error”) of an estimate (y) is defined as $c.v. = (s.e. / y) \times 100$, where s.e. is the standard error of the estimate y .

Estimates of standard error were computed using a technique known as jackknife replication. As with any replication method, jackknife replication involves constructing a number of subsamples (replicates) from the full sample and computing the statistic of interest for each replicate. The mean square error of the replicate estimates around the full sample estimate provides an estimate of the variance of the statistic. To construct the replications, 50 stratified subsamples of the full sample were created and then dropped 1 at a time to define 50 jackknife replicates. Thus, the weight variables used to calculate standard errors were AWT1 through AWT50. A computer program (WesVar) was used to calculate the estimates of standard errors and the design applied to this calculation was JK1.

Table 1. Number and percent of responding public elementary schools in the study sample, and estimated number and percent of public schools the sample represents, by school characteristics: 2005

School characteristic	Respondent sample (unweighted)		National estimate (weighted)	
	Number	Percent	Number	Percent
All public elementary schools	1,055	100	49,390	100
Enrollment size				
Less than 300	310	29	13,170	27
300 to 499	312	30	17,670	36
500 or more	433	41	18,560	38
Locale				
City	286	27	13,810	28
Urban fringe	378	36	17,940	36
Town	92	9	4,220	9
Rural	299	28	13,420	27
Region				
Northeast	178	17	8,620	17
Southeast	224	21	10,260	21
Central	295	28	14,030	28
West	358	34	16,500	33
Percent minority enrollment				
Less than 6 percent	222	21	10,140	21
6 to 20 percent	234	23	10,970	23
21 to 49 percent	230	22	10,980	23
50 percent or more	348	34	16,260	34
Percent of students eligible for free or reduced-price lunch				
Less than 35 percent	372	35	17,540	36
35 to 49 percent	164	16	7,570	15
50 to 74 percent	253	24	11,690	24
75 percent or more	266	25	12,600	26

NOTE: Percent minority enrollment was not available for 21 schools. Those schools were included in the totals and in the breakouts by other school characteristics. Detail may not sum to totals because of rounding or missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Fast Response Survey System, "Foods and Physical Activity in Public Elementary Schools: 2005," FRSS 87, 2005.

Nonsampling Errors, Coding, and Editing

The survey estimates are also subject to nonsampling errors that can arise because of nonobservation (nonresponse or noncoverage) errors, errors of reporting, and errors made in data collection. These errors can sometimes bias the data. Nonsampling errors may include such problems as

misrecording of responses; incorrect editing, coding, and data entry; differences related to the particular time the survey was conducted; or errors in data preparation. While general sampling theory can be used to determine how to estimate the sampling variability of a statistic, nonsampling errors are not easy to measure and, for measurement purposes, usually require that an experiment be conducted as part of the data collection procedures or that data external to the study be used.

To minimize the potential for nonsampling errors, the questionnaire was pretested with respondents at institutions like those that completed the survey. During the design of the survey and the survey pretest, an effort was made to check for consistency of interpretation of questions and to eliminate ambiguous items. The questionnaire and instructions were extensively reviewed by NCES.

Manual and machine editing of the questionnaire responses were conducted to check the data for accuracy and consistency. Cases with missing or inconsistent items were recontacted by telephone. A coding source file and editing manual were produced to identify cases requiring data retrieval or clarification and prepare cases for key entry. The source file served as a data dictionary and included the data file layout, a description of each data item, and a list of valid response codes, range formats, as well as codes for nonresponse, inapplicable responses, and defined skip patterns. The coding source file was used to develop the ACCESS database for data verification while the codebook served as the main tool for coding, editing, and processing questionnaires received by mail, fax, or telephone.

Logics, ranges, and validation checks were prepared prior to data collection and included online edit checks, manual logic checks, and automated checks using SAS. Online checks were incorporated into the online data entry system while manual edits were conducted to process cases received by mail, fax, or telephone. In both cases, where electronic and hardcopy survey data were submitted, steps were taken to ensure that the method of entering the data is the same, regardless of mode. For example, to enter survey data received by mail/fax or telephone, we accessed the survey website as “respondents” and “complete” the survey using the responses on the hardcopy survey. Subjecting all survey responses to the same set of built-in logics, ranges, and validation checks helps to ensure that data entry does not produce system differences in the survey data. In addition, all hardcopy data were subject to 100 percent verification and “double-keyed” in a simulated website interface for subsequent data checks.

Definitions of Selected Analysis Variables

Enrollment Size – This variable indicates the total number of students enrolled in the school based on data from the 2002–03 CCD. The variable was collapsed into the following categories:

- Less than 300 students (small)**
- 300 to 499 students (medium)**
- 500 or more students (large)**

School Locale – This variable indicates the type of community in which the school is located, as defined in the 2002–03 CCD (which uses definitions based on U.S. Census Bureau classifications). This variable was based on the eight-category locale variable from CCD, recoded into a four-category analysis variable for this report. Large and midsize cities were coded as city, the urban fringes of large and midsize cities were coded as urban fringe, large and small towns were coded as town, and rural areas outside and inside Metropolitan Statistical Areas (MSAs) were coded as rural. The categories are described in more detail below.

City – A large or midsize central city of a Consolidated Metropolitan Statistical Area (CMSA) or Metropolitan Statistical Area (MSA).

Urban fringe – Any incorporated place, Census-designated place, or non-place territory within a CMSA or MSA of a large or midsize city, and defined as urban by the Census Bureau.

Town – Any incorporated place or Census-designated place with a population greater than or equal to 2,500 and located outside a CMSA or MSA.

Rural – Any incorporated place, Census-designated place, or non-place territory defined as rural by the Census Bureau.

Region – This variable classifies schools into one of the four geographic regions used by the Bureau of Economic Analysis of the U.S. Department of Commerce, the National Assessment of Educational Progress, and the National Education Association. Data were obtained from the 2002–03 CCD School Universe file. The geographic regions are:

Northeast – Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont

Southeast – Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia

Central – Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin

West – Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming

Percent Minority Enrollment – This variable indicates the percentage of students enrolled in the school whose race or ethnicity is classified as one of the following: American Indian or Alaska Native, Asian or Pacific Islander, non-Hispanic Black, or Hispanic, based on data in the 2002–03 CCD School Universe file. Data on this variable were missing for 21 schools; schools with missing data were excluded from all analyses by percent minority enrollment. The percent minority enrollment variable was collapsed into the following four categories:

Less than 6 percent minority

6 to 20 percent minority

21 to 49 percent minority

50 percent or more minority

Percent of Students Eligible for Free or Reduced-Price Lunch—This variable was based on responses to question 16 on the survey questionnaire; if it was missing from the questionnaire (3.9 percent of all cases), it was obtained from the 2002–03 CCD School Universe File. This item served as a measurement of the concentration of poverty at the school. The categories are:

Less than 35 percent

35 to 49 percent

50 to 74 percent

75 percent or more

Data Disclosure Warning

Under law, public use data collected and distributed by the National Center for Education Statistics (NCES) may be used only for statistical purposes.

Any effort to determine the identity of any reported case by public-use data users is prohibited by law. Violations are subject to Class E felony charges of a fine up to \$250,000 and/or a prison term up to 5 years.

NCES does all it can to assure that the identity of data subjects cannot be disclosed. All direct identifiers, as well as any characteristics that might lead to identification, are omitted or modified in the dataset to protect the true characteristics of individual cases. Any intentional identification or disclosure of a person or institution violates the assurances of confidentiality given to the providers of the information. Therefore, users shall:

- Use the data in this dataset for statistical purposes only.
- Make no use of the identity of any person or institution discovered inadvertently, and advise NCES of any such discovery.
- Not link this dataset with individually identifiable data from other NCES or non-NCES datasets.
- To proceed you must signify your agreement to comply with the above-stated statutorily based requirements.

Data perturbations were conducted on some background data to preclude identification of individuals and institutions.