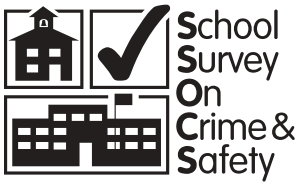
2000 School Survey on Crime and Safety

Detailed Data Documentation

U.S. Department of Education
Institute of Education Sciences
NCES 2004-307



2000 School Survey on Crime and Safety

Detailed Data Documentation

November 2003

U.S. Department of Education
Institute of Education Sciences
NCES 2004–307

Bradford Chaney
Sadeq Chowdhury
Adam Chu
Janice Lee
Peter Wobus
Westat

Kathryn Chandler
Project Director
National Center for Education Statistics

U.S. Department of Education
Rod Paige
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Val Plisko
A*ssociate Commissioner*

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006–5651

November 2003

The NCES World Wide Web Home Page address is *http://nces.ed.gov*
The NCES World Wide Web Electronic Catalog is: *http://nces.ed.gov/pubsearch*

Suggested Citation

U.S. Department of Education, National Center for Education Statistics. *2000 School Survey on Crime and Safety: Detailed Data Documentation,* NCES 2004–307, by Bradford Chaney, Sadeq Chowdhury, Adam Chu, Janice Lee, and Peter Wobus. Project Officer: Kathryn Chandler. Washington, DC: 2003.

For ordering information on this report, write:
U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794–1398

Call toll free 1–877–4ED–Pubs; or order online at www.edpubs.org.

Content Contact:
Kathryn Chandler
(202) 502–7486
Kathryn.Chandler@ed.gov

TABLE OF CONTENTS

List of Appendixes

List of Tables

List of Tables

List of Tables

List of Tables

List of Tables

List of Tables

List of Tables

List of Tables

List of Standard Error Tables

List of Figures

Figure

List of Standard Error Tables for Figures

1. INTRODUCTION

The School Survey on Crime and Safety (SSOCS) is a public school survey conducted by the National Center for Education Statistics (NCES). The survey builds on an earlier survey on school crime and safety conducted in 1997 using the Fast Response Survey System (FRSS),[1] and is one of several surveys on school crime and safety conducted by NCES. Funding for the SSOCS was provided by the Safe and Drug-Free Schools Program of the Office of Elementary and Secondary Education.

Conducted for the first time in the Winter/Spring of 2000, SSOCS:2000 is the only NCES survey to collect detailed information on crime and safety from the schools' perspective. As such, it fills an important gap in data collected by NCES. SSOCS:2000 collected information on:

- Characteristics of school policies,
- School violence prevention programs and practices,
- Violent deaths at school and elsewhere,
- Frequency of other incidents at school,
- Disciplinary problems and actions, and
- School characteristics that have been associated with school crime.

The SSOCS:2000 was developed in consultation with a Technical Review Panel consisting of some of the nation's top experts on school crime and school programs relating to crime and safety. As such, SSOCS:2000 provides a valuable tool to policymakers and researchers who need to know what policies and programs are in place, what the level of crime is and how it is changing, and what disciplinary actions schools are taking.

[1] Heaviside, S., Rowand, C., Williams, C., and Farris, E. Project Officers: S. Burns and E. McArthur. (1998). *Violence and Discipline Problems in U.S. Public Schools: 1996–1997* (NCES 98–030). Washington, DC: U.S. Department of Education, National Center for Education Statistics. .

By many measures, the issues of crime and safety are some of the more critical issues faced by U.S. school systems:

- The National Crime Victimization Survey indicated that 2.7 million violent crimes take place annually either at school or near schools.[2]

- A Phi Delta Kappa poll in 1996 found that teachers said that discipline is the main reason that teachers leave the profession.[3]

- A National Institute of Justice study found that one-third of male inner-city students were shot at, stabbed, or injured with a weapon at school or on the way to and from school over the past few years.[4]

- The Teacher Survey on Safe, Disciplined, and Drug-Free Schools indicated that 8 percent of teachers said in 1990-91 that they were threatened with injury by a student in the last 12 months.[5]

- The National Household Education Survey revealed that 56 percent of students said they had witnessed bullying, physical attack, or robbery at school or on the way to or from school.[6]

- In 1999, about 5 percent of students said that they had been bullied at school in the past 6 months, according to the School Crime Supplement to the National Crime Victimization survey.[7]

- The Longitudinal Study of Selected School Districts found that 37 percent of eighth and ninth graders were afraid of attacks at school.[8]

- A survey by the National Association of Secondary School Principals found that 52 percent of secondary school principals said their schools are facing serious gang problems.[9]

---

[2] Linquanti, R. and Borliner, B. *Rebuilding Schools as Safe Havens: A Typology for Selecting and Integrating Violence Prevention Strategies (Contract NO. S188A00001)* (NCES 2001-017), (ERIC Documentation Reproduction Service No. 376 600), Portland, OR: Western Regional Center for Drug-Free Schools and Communities, sponsored by the U.S. Department of Education, 1994; Kaufman, P., et al. (2000). *Indicators of School Crime and Safety, 2000*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

[3] The Third Phi Delta Kappa Poll of Teachers' Attitudes Toward Public Schools, *Phi Delta Kappan*, 1996.

[4] *Weapon-Related Victimization in Selected Inner-City School Samples.* National Institute of Justice, 1995.

[5] Mansfield, W., Alexander, D., and Farris, E. (1991). *Teacher Survey on Safe, Disciplined, and Drug-Free Schools* (NCES 91-091). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

[6] Nolin, M., Davies, E., and Chandler, K. (1995). *Student Victimization at School* (NCES 95-204). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

[7] Kaufman, P., et al. (2000). *Indicators of School Crime and Safety, 2000* (NCES 2001-017, 13). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

[8] Silvia, S., and Thorne, J. *School-Based Drug Prevention Programs: A Longitudinal Study in Selected School Districts*, Executive summary, Final report. Research Triangle Park, NC: Research Triangle Institute, 1997.

[9] *Safety Issues Serious in Nation's Schools: Principals Taking Action*, National Association of Secondary School Principals, 1997.

Providing a safe and disciplined environment is one of the core responsibilities of our school system. By acting in the role of parents, schools take on responsibility for the welfare of the children they serve. Clearly students' physical safety is basic to their welfare. However, students' welfare can be harmed even if they are not the ones being assaulted: "even youth who are not direct victims of violence may be victimized by the chronic presence of violence in their communities."[10] Providing safety and discipline is also critical in maintaining schools' educational function: students' attention to learning is considerably hindered when they are fearful for their own safety or welfare. Indeed, one of the determinants of classroom effectiveness is teachers' ability to maintain discipline within their classes.

For these reasons, the federal government has made safety and discipline one of its main priorities. SSOCS:2000 provides statistics on the frequency of violence, the nature of the school environment, and the characteristics of school violence prevention programs. Such national data are critical given the tendency to focus on anecdotal evidence of crimes without knowing the true frequency of problems in the schools. Without accurate information, policymakers may make misinformed decisions about school policy, and the public might lose confidence in public schools.

NCES has conducted one-time surveys that have provided much useful information. For example, at the time of the 1997 Fast Response Survey System survey on school crime, safety, and discipline, anecdotal news reports suggested that serious crime had become widespread in American schools, but the survey found that serious crime is actually quite rare. However, until the 2000 and subsequent SSOCS collections, NCES has not had a systematic approach to collecting such data. The amount of data collected has been relatively small, and the lack of a periodic survey has made it difficult to measure change over time.

This survey fills major gaps in our current information about the frequency and types of crime at schools, the nature of schools' policies and programs to reduce crime, and the disciplinary actions that schools take in response to crime. The data were used by NCES to prepare a summary descriptive report of the findings, and were made available as a public use database (following the removal of identifying information) for use by researchers on school crime and safety. NCES worked with the Safe and Drug-Free Schools Program within the Department of Education to design this survey, and the data will be used by that program. The survey also was developed in consultation with ED's Office of Special Education and Rehabilitative Services (OSERS) and provides that office with valuable data concerning special education students.

[10] American Psychological Association Commission on Violence and Youth, quoted in M. Nolin, et al., 1995.

2. SAMPLE DESIGN

A stratified sample design was used to select schools for the 2000 School Survey on Crime and Safety (SSOCS:2000). For sample allocation purposes, strata were defined by instructional level, type of locale, and enrollment size. Minority status and region were also used as sorting variables in the sample selection process to induce additional implicit stratification. The sample was designed to provide reasonably precise cross-sectional estimates for selected subgroups of interest. Various aspects of the sample design such as construction of the sampling frame, stratification (including the choice of stratification variables), and sample allocation are described in detail in the following sections.

Sampling Frame

The sampling frame for the SSOCS:2000 was constructed from the public school universe file created for the 2000 Schools and Staffing Survey (SASS). However, only the approximately 81,000 "regular" schools (excluding schools in the outlying U.S. territories, ungraded schools, and those with a high grade of kindergarten or lower) in the 1997–98 NCES Common Core of Data (CCD) Public School Universe File within the SASS frame were eligible for the study. The SASS frame was derived primarily from the 1997–98 CCD, which includes charter schools. These CCD schools, including charter schools, were included in the SSOCS:2000 study. (The SASS frame also includes a supplement made up of additional charter schools as well as a small number of Bureau of Indian Affairs and Department of Defense schools not represented in the 1997–98 CCD file. Schools from this supplement were not included in the SSOCS:2000 study.) Tables 2-1 and 2-2A through 2-2C summarize the distribution of the eligible regular schools in the SASS/CCD frame by grade span, instructional level, type of locale, enrollment size, and minority status. Note that the percentage minority categories used in table 2-2C serve to illustrate how widely the schools in the 1997–98 CCD frame vary by minority status. However, it is not necessary to use these same categories for analysis purposes.

Table 2-1. Number of regular schools in the SASS frame,* by instructional level: 1997–1998

Low grade	High grade 1	2	3	4	5	6	7	8	9	10	11	12	Total
PK	219	753	1,063	2,010	13,960	10,718	276	3,128	43	27	31	647	32,875
K	211	534	550	971	6,862	3,761	113	1,231	27	19	16	457	14,752
1	14	104	160	201	622	331	33	159	5	4	4	18	1,655
2		10	111	74	158	84	9	32	2		1	12	493
3			15	131	768	218	14	83	4	3	7	16	1,259
4				17	378	554	40	261	3	2	4	20	1,279
5					39	404	74	1,344	5	4	4	40	1,914
6						123	133	7,900	117	27	25	404	8,729
7							31	2,745	797	26	37	2,774	6,410
8								34	111	12	11	266	434
9									94	68	75	10,725	10,962
10										10	6	560	576
11											6	42	48
12												19	19
Total	444	1,401	1,899	3,404	22,787	16,193	723	16,917	1,208	202	227	16,000	81,405

Elementary (49,691)	Lowest grade <= 3 and highest grade <=8
Middle/junior high (15,204)	Lowest grade >=4 and highest grade <=9
Secondary/senior high (11,511)	Lowest grade >= 9 and highest grade <= 12 (secondary) or lowest grade >=9 and highest grade = 12 (senior high)
Combined (4,999)	Lowest grade <= 8; highest grade > 8 if lowest grade <= 3; highest grade > 9 if lowest grade > 3

*Counts exclude schools in the outlying U.S. territories, nonregular schools such as special education, vocational, alternative/other schools, ungraded schools and schools with a high grade of kindergarten or lower. Includes charter schools listed in the 1997–98 CCD, but not the extra charter schools added to the SASS frame.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-2A. Number of regular schools and enrollment in the SASS/CCD public school universe file,[1] by instructional level and type of locale: 1997–1998

Instructional level	Type of locale[2]	Number of regular schools	Total enrollment in schools
Elementary	City	14,958	8,114,496
	Urban Fringe	17,051	8,854,689
	Town	6,397	2,523,617
	Rural	11,285	3,313,656
	Total Elementary	*49,691*	*22,806,458*
Middle	City	3,812	2,960,615
	Urban Fringe	5,504	3,879,747
	Town	2,685	1,322,262
	Rural	3,203	970,137
	Total Middle	*15,204*	*9,132,761*
Secondary	City	2,441	3,575,163
	Urban Fringe	3,702	4,554,666
	Town	2,075	1,543,139
	Rural	3,293	1,226,898
	Total Secondary	*11,511*	*10,899,866*
Combined	City	522	351,738
	Urban Fringe	733	549,083
	Town	504	272,335
	Rural	3,240	1,051,183
	Total Combined	*4,999*	*2,224,339*
All Levels	City	21,733	15,002,012
	Urban Fringe	26,990	17,838,185
	Town	11,661	5,661,353
	Rural	21,021	6,561,874
	Total All Levels	*81,405*	*45,063,424*

[1]Counts exclude schools in the outlying U.S. territories, nonregular schools such as special education, vocational, alternative/other schools, ungraded schools, and schools with a high grade of kindergarten or lower. See table 2-1 for definition of instructional levels used in this table.

[2]The following definitions in the 1997–98 CCD file apply to the type of locale. City: a central city of a consolidated metropolitan statistical area. Urban fringe: any incorporated place, Census-designated place, or non-place territory within a CMSA or MSA of a city, and defined as urban by the Census Bureau. Town: any incorporated place or Census-designated place with 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 designated as rural by the Census Bureau. For SSOCS: 2000, the CCD types of locale codes were collapsed as follows: large city (1) and mid-size city (2) = city; urban fringe of a large city (3) and urban fringe of a mid-size city (4) = urban fringe; large town (5) and small town (6) = town; and rural, outside MSA (7) and rural, inside MSA (8) = rural.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-2B. Number of regular schools and enrollment in the SASS/CCD public school universe file,[1] by instructional level and enrollment size: 1997–1998

Instructional level	Enrollment size[2]	Number of regular schools	Total enrollment in schools
Elementary	1. Under 300	13,300	2,396,408
	2. 300 to 499	16,811	6,731,291
	3. 500 to 999	18,204	12,041,452
	4. 1000+	1,376	1,637,307
	Total Elementary	*49,691*	*22,806,458*
Middle	1. Under 300	3,243	542,577
	2. 300 to 499	3,191	1,280,625
	3. 500 to 999	6,884	4,939,959
	4. 1000+	1,886	2,369,600
	Total Middle	*15,204*	*9,132,761*
Secondary	1. Under 300	2,387	366,029
	2. 300 to 499	1,563	622,628
	3. 500 to 999	2,953	2,169,958
	4. 1000+	4,608	7,741,251
	Total Secondary	*11,511*	*10,899,866*
Combined	1. Under 300	2,370	353,714
	2. 300 to 499	1,057	414,025
	3. 500 to 999	1,167	800,156
	4. 1000+	405	656,444
	Total Combined	*4,999*	*2,224,339*
All Levels	1. Under 300	21,300	3,658,728
	2. 300 to 499	22,622	9,048,569
	3. 500 to 999	29,208	19,951,525
	4. 1000+	8,275	12,404,602
	Total All Level	*81,405*	*45,063,424*

[1]Counts exclude schools in the outlying U.S. territories, nonregular schools such as special education, vocational, alternative/other schools, ungraded schools, and schools with a high grade of kindergarten or lower. See table 2-1 for definition of instructional levels used in this table.

[2]Enrollment size categories are not necessarily optimized for analytic purposes. Different size categories for the various levels can be used in analysis if desired.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-2C. Number of regular schools and enrollment in the SASS/CCD public school universe file,[1] by instructional level and percentage minority enrollment: 1997–1998

Instructional level	Percentage minority enrollment[2]	Number of regular schools	Total enrollment in schools
Elementary	1. <5 percent or unknown	12,474	4,174,547
	2. 5 to 19 percent	12,632	5,511,455
	3. 20 to 49 percent	10,628	5,244,125
	4. 50 percent +	13,957	7,876,331
	Total Elementary	*49,691*	*22,806,458*
Middle	1. <5 percent or unknown	3,769	1,599,973
	2. 5 to 19 percent	4,191	2,546,443
	3. 20 to 49 percent	3,455	2,314,901
	4. 50 percent +	3,789	2,671,444
	Total Middle	*15,204*	*9,132,761*
Secondary	1. <5 percent or unknown	3,489	2,014,679
	2. 5 to 19 percent	3,104	2,955,622
	3. 20 to 49 percent	2,533	2,849,326
	4. 50 percent +	2,385	3,080,239
	Total Secondary	*11,511*	*10,899,866*
Combined	1. <5 percent or unknown	2,322	908,457
	2. 5 to 19 percent	1,018	470,363
	3. 20 to 49 percent	779	429,857
	4. 50 percent +	880	415,662
	Total Combined	*4,999*	*2,224,339*
All Levels	1. <5 percent or unknown	22,054	8,697,656
	2. 5 to 19 percent	20,945	11,483,883
	3. 20 to 49 percent	17,395	10,838,209
	4. 50 percent +	21,011	14,043,676
	Total All Levels	*81,405*	*45,063,424*

[1]Counts exclude schools in the outlying U.S. territories, nonregular schools such as special education, vocational, alternative/other schools, ungraded schools. and schools with a high grade of kindergarten or lower. See table 2-1 for definition of instructional levels used in this table.

[2]Minority enrollment derived from racial/ethnic counts on the 1997–98 CCD. Included in the minority counts are the following racial/ethnic groups: American Indian or Alaskan Native, Asian or Pacific Islander, Hispanic, and Black non-Hispanic. These categories were chosen as those that have commonly been used. Other definitions of "minority" can be used in analysis.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Sample Size

The target sample size for the SSOCS:2000 was approximately 3,000 responding schools. The target sample size was determined to be sufficiently large to allow detection of a 10 percent relative change in a 25 percent population characteristic with 95 percent confidence, after allowance for possible design effects.

In general, the standard error of a difference in proportions, $\hat{p}_1 - \hat{p}_2$, is given approximately by

$$SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{2DP(1-P)}{n}}$$

where D is the design effect, P is the underlying proportion being estimated, and n is the sample size (assumed to be roughly equal for the two proportions $\hat{p}_1$ and $\hat{p}_2$). This implies that the sample size needed for an estimate of change (or difference) to be subject to a relative "margin of error" of no more than 10 percent at the 95 percent confidence level should be at least

$$n = \frac{2(1.96)^2 DP(1-P)}{(.10P)^2} .$$

The required sample size, n, as computed from the above formula is summarized in table 2-3 for selected values of D and P. For example, for $P = 0.25$, a design effect of 1.10, and a margin of error of 10 percent, the required total sample size would be about 2,535. On the other hand, if the design effect is as high as 1.30, a sample size of 2,996 would be needed to achieve a 10 percent margin of error. Thus, as long as the design effect is no greater than 1.30, a sample size of around 3,000 schools will satisfy the 10 percent relative difference criterion.

Similarly, to detect a relative difference between subgroups (i.e., within a given survey) of 15 percent on a 30 percent characteristic ($P = 0.30$), a sample size of 876 would be required per subgroup if the design effect is no more than $D = 1.10$. An implication of the latter calculation is that with a total sample size of 3,000, generally only one-way comparisons involving no more than 3 or 4 subgroups (e.g., comparisons by instructional level or by enrollment size class, but not necessarily by the cross classification of level and size class) will satisfy the 15 percent precision requirement.

Finally, it should be noted that with a sample size of 3,000 responding schools, the probability of selecting at least one school for which a particular type of crime or incident has occurred is relatively high, except for the extremely rare events. As indicated in the last column of table 2-4, this probability exceeds 99 percent for events with prevalence rates as low as 0.5 percent. However, even though the probability of observing an occurrence of a particular crime is high, the expected number of schools reporting that crime may still be too small to support detailed analysis.

Table 2-3. Sample size required per group for an estimated difference to be subject to relative margin of error of specified size at the 95 percent confidence level

Population proportion (*P*)	Relative margin of error (%)	Design effect (*D*)			
		1.00	1.10	1.20	1.30
0.10	10	6,915	7,606	8,298	8,989
	15	3,073	3,381	3,688	3,995
0.20	10	3,073	3,381	3,688	3,995
	15	1,366	1,502	1,639	1,776
0.25	10	2,305	2,535	2,766	2,996
	15	1,024	1,127	1,229	1,332
0.30	10	1,793	1,972	2,151	2,331
	15	797	876	956	1,036
0.40	10	1,152	1,268	1,383	1,498
	15	512	563	615	666
0.50	10	768	845	922	999
	15	341	376	410	444

SOURCE: Statistical computations made for the U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-4. Probability of observing at least one school reporting the occurrence of a specific type of crime for alternative sample sizes and prevalence rates

Prevalence of crime (%)	Sample size (assumes simple random sampling)*				
	300 (%)	500 (%)	800 (%)	1000 (%)	3000 (%)
5.00	>99	>99	>99	>99	>99
1.00	95	>99	>99	>99	>99
0.50	78	92	98	>99	>99
0.10	26	39	55	63	95
0.05	14	22	33	39	78
0.01	3	5	8	10	26

*For the stratified sample design developed for SSOCS:2000, the actual probabilities may be smaller than those shown above.
SOURCE: Statistical computations made for the U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Stratification and Sample Allocation

Stratification refers to the process of subdividing the population frame into mutually exclusive subsets (called strata) from which samples of schools are selected at appropriate rates. There are two main goals of stratification. The first is to ensure that selected subdomains of interest are adequately represented in the sample for analysis purposes. For example, in the SASS/CCD universe file about 60 percent of public schools are elementary schools, with the remaining 40 percent roughly equally divided between middle and secondary schools. Thus, if a simple random sample of schools is selected without regard to level, the majority of the sampled schools will be elementary schools where the incidence of crimes and discipline problems is expected to be relatively low. Such a design would be inefficient for comparisons between the various levels of schools and for overall national estimates.

The second goal of stratification is to improve sampling precision by permitting a more nearly optimal allocation of the sample to the various sampling strata. For a fixed sample size, the optimum allocation (i.e., the allocation that produces the smallest sampling error) is a function of the number of schools in the stratum and the underlying within-stratum variance of the statistic of interest. Estimation of different types of statistics (e.g., the proportion of schools that report a particular type of incident *vs*. the total number of incidents reported by schools) can lead to vastly different sample

allocations. An important goal of the design process is to develop a sample allocation that is reasonably efficient for a range of different types of statistics.

Stratification Variables

An initial step in identifying potentially effective stratifiers was to examine the variation of selected crime and school violence statistics by school-level characteristics. For this purpose, estimates from the FRSS *Survey on Violence and Discipline Problems in U.S. Public Schools: 1996–97* were used.[11] Selected results from this study are summarized in tables 2-5 and 2-6. As can be seen in table 2-5, the percentage of schools reporting various types of incidents varied by instructional level, enrollment size, and type of locale (where, in general, the percentage of schools reporting crime incidents is higher for middle and secondary schools, schools with enrollment of 1,000 or more, and city schools). These variables were used to define the primary stratification variables. The reported numbers of incidents also varied by percentage minority enrollment and region. Therefore, these variables were used as sorting variables in the sampling process to induce additional implicit stratification (see Selection of the Sample).

Allocation of Sample to Strata

Tables 2-7A through 2-7D summarize the distribution of schools in the SASS/CCD frame by sampling stratum, where the sampling strata are defined by level, type of locale, and enrollment size category. Within each stratum, the distribution of schools by percentage minority enrollment is also shown. The corresponding distributions by region are shown in tables 2-7E through 2-7H. Though not used for sample allocation purposes, minority status and region were used as implicit stratifiers in the sampling process.

Initially, the target sample size of 3,000 responding schools was allocated to four instructional level categories as follows: 750 elementary schools, 1,000 middle schools, 1,000 secondary schools, and 250 combined schools (see table 2-1 for definition of the four

[11] Heaviside, S., Rowand, C., Williams, C., and Farris, E. Project Officers Burns, S., and McArthur, E. (1998). *Violence and Discipline Problems in U.S. Public Schools*: *1996–97* (NCES 98-030). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Table 2-5. Estimated percentage of schools in FRSS reporting selected types of crimes, by school characteristics: 1996–97 school year

School characteristic	Number of schools in population	Sample size	Type of incident reported to law enforcement				
			Physical attacks or fights with weapons (%)	Robbery (%)	Physical attacks or fights without weapons (%)	Theft or larceny (%)	Vandal-ism (%)
All (regular) public schools	77,752	1,234	6	3	28	31	38
Instructional level							
Elementary	48,100	354	2	1	12	19	30
Middle	14,008	439	12	5	51	44	47
Secondary	15,644	441	13	8	55	55	52
*Enrollment size**							
Less than 300	20,280	169	2	#	17	18	23
300-999	50,071	745	6	2	26	30	40
1,000+	7,402	320	20	16	67	68	62
Type of locale							
City	17,990	406	10	8	30	34	41
Urban fringe	19,017	279	6	3	28	29	37
Town	19,656	296	3	1	32	36	44
Rural	21,089	253	5	1	21	24	30
Percentage minority enrollment							
Less than 5 percent	24,208	309	3	1	22	24	29
5 to 19 percent	17,555	297	6	2	27	28	40
20 to 49 percent	17,747	290	7	3	32	31	38
50 percent or more	17,425	328	9	7	32	41	47
Region							
Northeast	14,997	229	4	3	23	26	37
Southeast	16,949	296	5	2	29	32	36
Central	22,500	323	6	3	26	26	30
West	23,203	386	8	4	32	28	47

Rounds to zero.

*Enrollment sizes used in FRSS report. See Heaviside, et al. (1998).

SOURCE: U.S. Department of Education. National Center for Education Statistics. *Violence and Discipline Problems in U.S. Public Schools: 1996–97*. NCES 98-030, by S. Heaviside, C. Rowand, C. Williams, and E. Farris. Project Officers, S. Burns and E. McArthur. Washington, DC: 1998.

Table 2-6. Standard errors of percentage of schools in FRSS reporting selected types of crimes, by school characteristics: 1996–97 school year

School characteristic	Number of schools in population	Sample size	Type of incident reported to law enforcement				
			Physical attacks or fights with weapons	Robbery	Physical attacks or fights without weapons	Theft or larceny	Vandal-ism
All (regular) public schools	77,752	1,234	0.5	0.4	1.1	1.5	1.6
Instructional level							
Elementary	48,100	354	0.8	0.4	1.7	2.0	2.3
Middle	14,008	439	1.2	0.9	2.3	2.5	2.6
Secondary	15,644	441	1.5	1.0	2.9	3.1	2.7
*Enrollment size**							
Less than 300	20,280	169	0.9	#	2.9	3.1	3.9
300-999	50,071	745	0.8	0.5	1.4	1.8	1.7
1,000+	7,402	320	2.0	1.8	3.0	3.2	3.3
Type of locale							
City	17,990	406	1.7	1.3	2.4	2.7	3.2
Urban fringe	19,017	279	1.2	0.7	2.6	2.4	2.9
Town	19,656	296	0.9	0.4	2.3	3.5	3.5
Rural	21,089	253	1.2	0.6	2.6	3.3	3.7
Percentage minority enrollment							
Less than 5 percent	24,208	309	0.9	0.5	2.4	2.8	3.3
5 to 19 percent	17,555	297	1.2	0.8	3.0	3.1	3.0
20 to 49 percent	17,747	290	1.6	0.8	2.9	3.3	3.5
50 percent or more	17,425	328	1.3	1.1	3.1	3.5	3.3
Region							
Northeast	14,997	229	1.0	0.6	2.3	3.3	3.6
Southeast	16,949	296	1.3	0.6	2.7	3.6	3.1
Central	22,500	323	1.4	0.8	2.6	2.7	3.6
West	23,203	386	1.2	0.8	3.0	3.0	3.5

Rounds to zero.
*Enrollment sizes used in FRSS report. See Heaviside, et al. (1998).
SOURCE: U.S. Department of Education. National Center for Education Statistics. *Violence and Discipline Problems in U.S. Public Schools*: *1996–97*. NCES 98-030, by S. Heaviside, C. Rowand, C. Williams, and E. Farris. Project Officers, S. Burns and E. McArthur. Washington, DC: 1998.

Table 2-7A. Number of elementary schools in SASS/CCD frame, by type of locale, enrollment size, and minority status: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Number of schools in frame (row total)	Percentage minority enrollment*			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Elementary	City	Less than 300	2,104	204	471	561	868
		300 to 499	5,243	310	979	1,415	2,539
		500 to 999	6,851	182	850	1,695	4,124
		1,000+	760	3	34	90	633
	Urban fringe	Less than 300	2,467	821	865	472	309
		300 to 499	6,124	1,362	2,259	1,480	1,023
		500 to 999	7,952	1,191	2,600	2,059	2,102
		1,000+	508	44	108	128	228
	Town	Less than 300	2,199	877	763	328	231
		300 to 499	2,479	674	791	555	459
		500 to 999	1,651	414	385	437	415
		1,000+	68	11	4	24	29
	Rural	Less than 300	6,530	4,054	1,321	640	515
		300 to 499	2,965	1,541	711	417	296
		500 to 999	1,750	776	480	314	180
		1,000+	40	10	11	13	6
Total			49,691	12,474	12,632	10,628	13,957

*Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1990–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-7B. Number of middle schools in SASS/CCD frame, by type of locale, enrollment size, and minority status: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Number of schools in frame (row total)	Percentage minority enrollment*			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Middle	City	Less than 300	281	29	32	50	170
		300 to 499	478	39	81	112	246
		500 to 999	2,210	104	417	629	1,060
		1,000+	843	11	109	235	488
	Urban fringe	Less than 300	502	162	140	116	84
		300 to 499	1,033	337	353	207	136
		500 to 999	3,049	581	1,190	752	526
		1,000+	920	90	305	295	230
	Town	Less than 300	597	229	161	101	106
		300 to 499	909	272	260	207	170
		500 to 999	1,096	264	330	276	226
		1,000+	83	17	26	22	18
	Rural	Less than 300	1,863	988	434	247	194
		300 to 499	771	399	176	104	92
		500 to 999	529	236	163	88	42
		1,000+	40	11	14	14	1
Total			15,204	3,769	4,191	3,455	3,789

*Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-7C. Number of secondary schools in SASS/CCD frame, by type of locale, enrollment size, and minority status: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Number of schools in frame (row total)	Percentage minority enrollment*			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Secondary	City	Less than 300	218	23	28	47	120
		300 to 499	83	14	15	14	40
		500 to 999	353	42	48	67	196
		1,000+	1,787	68	315	554	850
	Urban fringe	Less than 300	212	62	60	42	48
		300 to 499	273	126	74	52	21
		500 to 999	1,083	405	405	178	95
		1,000+	2,134	282	793	638	421
	Town	Less than 300	221	62	60	57	42
		300 to 499	455	188	107	98	62
		500 to 999	890	356	253	166	115
		1,000+	509	119	178	122	90
	Rural	Less than 300	1,736	890	408	275	163
		300 to 499	752	441	157	97	57
		500 to 999	627	350	139	85	53
		1,000+	178	61	64	41	12
Total			11,511	3,489	3,104	2,533	2,385

*Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-7D. Number of combined schools in SASS/CCD frame, by type of locale, enrollment size, and minority status: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Number of schools in frame (row total)	Percentage minority enrollment*			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Combined	City	Less than 300	223	18	28	55	122
		300 to 499	51	10	7	9	25
		500 to 999	104	8	23	23	50
		1,000+	144	8	17	35	84
	Urban fringe	Less than 300	188	58	40	39	51
		300 to 499	116	57	37	12	10
		500 to 999	277	117	68	67	25
		1,000+	152	37	49	39	27
	Town	Less than 300	134	52	31	27	24
		300 to 499	116	52	29	22	13
		500 to 999	195	73	46	44	32
		1,000+	59	12	13	21	13
	Rural	Less than 300	1,825	996	359	198	272
		300 to 499	774	475	143	80	76
		500 to 999	591	329	110	101	51
		1,000+	50	20	18	7	5
Total			4,999	2,322	1,018	779	880

*Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-7E. Number of elementary schools in SASS/CCD frame, by type of locale, enrollment size, and region: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Number of schools in frame (row total)	Region*			
				North-east	South-east	Central	West
Elementary	City	Less than 300	2,104	371	308	821	604
		300 to 499	5,243	841	1,024	1,664	1,714
		500 to 999	6,851	1,170	1,483	1,246	2,952
		1,000+	760	225	126	85	324
	Urban fringe	Less than 300	2,467	867	302	753	545
		300 to 499	6,124	2,043	751	1,866	1,464
		500 to 999	7,952	1,758	1,688	1,349	3,157
		1,000+	508	66	204	22	216
	Town	Less than 300	2,199	210	285	1,080	624
		300 to 499	2,479	197	647	789	846
		500 to 999	1,651	114	754	298	485
		1,000+	68	5	49	5	9
	Rural	Less than 300	6,530	852	1,021	2,937	1,720
		300 to 499	2,965	505	877	1,001	582
		500 to 999	1,750	398	715	361	276
		1,000+	40	11	23	2	4
Total			49,691	9,633	10,257	14,279	15,522

*Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.
SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-7F. Number of middle schools in SASS/CCD frame, by type of locale, enrollment size, and region: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Number of schools in frame (row total)	Region*			
				North-east	South-east	Central	West
Middle	City	Less than 300	281	56	46	90	89
		300 to 499	478	89	100	185	104
		500 to 999	2,210	346	494	555	815
		1,000+	843	161	208	57	417
	Urban fringe	Less than 300	502	137	45	179	141
		300 to 499	1,033	324	120	349	240
		500 to 999	3,049	870	521	755	903
		1,000+	920	175	295	101	349
	Town	Less than 300	597	28	91	236	242
		300 to 499	909	69	279	265	296
		500 to 999	1,096	89	406	269	332
		1,000+	83	8	35	20	20
	Rural	Less than 300	1,863	98	147	868	750
		300 to 499	771	134	242	249	146
		500 to 999	529	146	194	100	89
		1,000+	40	18	15	4	3
Total			15,204	2,748	3,238	4,282	4,936

*Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-7G. Number of secondary schools in SASS/CCD frame, by type of locale, enrollment size, and region: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Number of schools in frame (row total)	Region*			
				North-east	South-east	Central	West
Secondary	City	Less than 300	218	35	36	80	67
		300 to 499	83	24	19	21	19
		500 to 999	353	86	87	111	69
		1,000+	1,787	271	420	403	693
	Urban fringe	Less than 300	212	32	22	68	90
		300 to 499	273	76	31	106	60
		500 to 999	1,083	445	153	321	164
		1,000+	2,134	488	451	468	727
	Town	Less than 300	221	4	37	66	114
		300 to 499	455	31	101	180	143
		500 to 999	890	94	279	288	229
		1,000+	509	32	180	128	169
	Rural	Less than 300	1,736	48	59	789	840
		300 to 499	752	92	157	325	178
		500 to 999	627	137	230	183	77
		1,000+	178	52	80	17	29
Total			11,511	1,947	2,342	3,554	3,668

*Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-7H. Number of combined schools in SASS/CCD frame, by type of locale, enrollment size, and region: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Number of schools in frame (row total)	Region*			
				North-east	South-east	Central	West
Combined	City	Less than 300	223	25	43	66	89
		300 to 499	51	12	6	14	19
		500 to 999	104	21	29	35	19
		1,000+	144	20	62	39	23
	Urban fringe	Less than 300	188	22	33	61	72
		300 to 499	116	47	22	34	13
		500 to 999	277	141	67	50	19
		1,000+	152	43	61	23	25
	Town	Less than 300	134	2	48	38	46
		300 to 499	116	5	48	45	18
		500 to 999	195	26	104	50	15
		1,000+	59	9	39	10	1
	Rural	Less than 300	1,825	71	227	860	667
		300 to 499	774	136	242	305	91
		500 to 999	591	171	272	117	31
		1,000+	50	11	31	6	2
Total			4,999	762	1,334	1,753	1,150

*Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

instructional level categories). Such an allocation was designed to permit a relatively detailed analysis of the three major instructional levels (elementary, middle, secondary), but was also expected be reasonably efficient for overall estimates. Combined schools (schools with both elementary and secondary grades) were placed in a separate stratum and were sampled independently of the secondary schools. This had virtually no impact on subsequent sample weighting procedures. However, since the sample of combined schools was expected to be relatively small, estimates for subcategories based on grade span (e.g., combined schools with mostly high grades vs. those with mostly low grades) will be subject to very large sampling errors.

Next, within each of the four instructional level categories defined in tables 2-7A through 2-7H, the sample size was further allocated to subgroups (substrata) defined by type of locale and enrollment size class in rough proportion to the aggregate square root of the enrollment of schools in the subgroup. Tables 2-8A through 2-8D summarize the aggregate square root of enrollment by sampling stratum and percentage minority enrollment. The corresponding tables showing breakouts by region within stratum are shown in tables 2-8E through 2-8H. The use of the square root of enrollment to determine the sample allocation was expected to be reasonably efficient for estimating both school-level characteristics (e.g., the number or percentage of schools that reported a certain type of crime) and quantitative measures correlated with enrollment (e.g., the number of incidents or the number of students in schools that reported a certain type of crime). The resulting allocation has the effect of varying the sampling rates by type of locale and size class within each instructional level. In particular, large schools generally had higher probabilities of selection than small schools under this allocation.

Tables 2-9A through 2-9D summarize the allocation of the sample (in terms of the target numbers of responding schools) by sampling stratum (i.e., the groups defined by level, type of locale, and enrollment size class) and percentage minority enrollment. The sample sizes were obtained by dividing the measure of size for a stratum (the aggregate square root of the enrollment corresponding to a row in tables 2-8A through 2-8D) by the total measure of size, and multiplying the result by the sample size for that instructional level. For example, elementary schools with an enrollment size less than 300 and that were located in cities accounted for 3 percent of the measure of size for elementary schools (31,305 divided by 1,029,790); 3 percent of the intended sample size of 750 elementary schools is 23. The sample size for the stratum (row) was then distributed to the four minority status groups in proportion to the numbers of schools in the group. Tables 2-9E through 2-9H also summarize the allocation of the sample by region. Region was not a sampling variable, but the use of sorting had the effect of implicitly stratifying the sample by region.

To compensate for losses due to nonresponse, a somewhat larger sample of about 3,300 schools was selected for the study. In the 1996–97 FRSS survey on school violence, response rates were found to vary by type of locale, enrollment size, and minority status. The speculated response rates given in tables 2-10A and 2-10B are based roughly on the FRSS survey on school violence experience. Using the response rates in tables 2-10A and 2-10B, the numbers of schools to be selected for SSOCS:2000 were derived as shown in tables 2-11A through 2-11H. However, the actual numbers of sampled schools differed slightly from the target numbers because of the use of overlap minimization procedures described in the following section.

Table 2-8A. Aggregate measure of size of elementary schools in SASS/CCD frame, by type of locale, enrollment size, and minority status: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Measure of size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Elementary	City	Less than 300	31,305	2,858	7,045	8,428	12,974
		300 to 499	105,237	6,163	19,611	28,440	51,022
		500 to 999	177,864	4,548	21,565	43,709	108,042
		1,000+	26,364	107	1,155	3,066	22,035
	Urban fringe	Less than 300	35,729	11,736	12,712	6,805	4,476
		300 to 499	123,411	27,247	45,512	29,963	20,689
		500 to 999	204,484	30,161	66,177	53,206	54,940
		1,000+	17,302	1,495	3,642	4,324	7,842
	Town	Less than 300	30,164	11,263	10,814	4,748	3,339
		300 to 499	49,171	13,278	15,614	11,080	9,199
		500 to 999	41,681	10,420	9,618	11,067	10,576
		1,000+	2,341	366	135	829	1,011
	Rural	Less than 300	80,696	49,089	16,615	8,218	6,774
		300 to 499	58,560	30,318	14,055	8,300	5,886
		500 to 999	44,105	19,374	12,182	7,989	4,559
		1,000+	1,376	343	383	441	208
Total			1,029,790	218,766	256,836	230,615	323,573

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the 1997–98 CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.
[2]Aggregate measure of size is equal to the sum of the square root of the enrollment of the schools in each type-of-locale, enrollment size, and minority status cell.
SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-8B. Aggregate measure of size of middle schools in SASS/CCD frame, by type of locale, enrollment size, and minority status: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Measure of size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Middle	City	Less than 300	3,527	384	417	618	2,108
		300 to 499	9,713	782	1,646	2,276	5,010
		500 to 999	60,193	2,792	11,347	17,094	28,961
		1,000+	30,051	371	3,769	8,297	17,614
	Urban fringe	Less than 300	7,133	2,258	2,043	1,684	1,148
		300 to 499	20,878	6,797	7,140	4,198	2,742
		500 to 999	82,029	15,337	32,096	20,339	14,257
		1,000+	32,430	3,122	10,703	10,359	8,245
	Town	Less than 300	8,565	3,225	2,322	1,481	1,537
		300 to 499	18,144	5,417	5,210	4,156	3,361
		500 to 999	28,515	6,873	8,531	7,225	5,886
		1,000+	2,811	572	895	732	612
	Rural	Less than 300	22,580	11,683	5,400	3,176	2,321
		300 to 499	15,186	7,851	3,457	2,072	1,806
		500 to 999	13,635	6,015	4,249	2,304	1,068
		1,000+	1,348	368	473	471	36
Total			356,738	73,848	99,698	86,481	96,711

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.
[2]Aggregate measure of size is equal to the sum of the square root of the enrollment of the schools in each type-of-locale, enrollment size, and minority status cell.
SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-8C. Aggregate measure of size of secondary schools in SASS/CCD frame, by type of locale, enrollment size, and minority status: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Measure of size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Secondary	City	Less than 300	2,448	214	277	542	1,414
		300 to 499	1,679	285	306	279	809
		500 to 999	9,902	1,132	1,317	1,904	5,549
		1,000+	75,407	2,608	12,684	23,009	37,107
	Urban fringe	Less than 300	2,468	729	699	535	505
		300 to 499	5,507	2,564	1,479	1,041	423
		500 to 999	29,831	11,112	11,236	4,896	2,586
		1,000+	86,854	10,500	31,426	26,551	18,376
	Town	Less than 300	2,956	813	821	769	553
		300 to 499	9,140	3,773	2,168	1,957	1,242
		500 to 999	23,889	9,494	6,873	4,425	3,097
		1,000+	18,509	4,206	6,440	4,510	3,353
	Rural	Less than 300	21,647	11,221	5,020	3,432	1,975
		300 to 499	14,868	8,714	3,114	1,925	1,115
		500 to 999	16,403	9,137	3,694	2,219	1,353
		1,000+	6,425	2,119	2,346	1,512	449
Total			327,933	78,620	89,900	79,508	79,905

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.
[2]Aggregate measure of size is equal to the sum of the square root of the enrollment of the schools in each type-of-locale, enrollment size, and minority status cell.
SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-8D. Aggregate measure of size of combined schools in SASS/CCD frame, by type of locale, enrollment size, and minority status: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Measure of size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Combined	City	Less than 300	2,188	189	294	529	1,177
		300 to 499	1,026	196	148	179	504
		500 to 999	2,827	207	617	639	1,363
		1,000+	5,754	307	682	1,406	3,359
	Urban fringe	Less than 300	1,950	660	445	399	445
		300 to 499	2,321	1,138	741	250	192
		500 to 999	7,419	3,099	1,859	1,784	677
		1,000+	6,308	1,678	1,865	1,680	1,085
	Town	Less than 300	1,398	585	300	293	220
		300 to 499	2,332	1,049	586	441	256
		500 to 999	5,107	1,893	1,194	1,162	858
		1,000+	2,123	417	460	778	468
	Rural	Less than 300	23,257	12,967	4,695	2,530	3,064
		300 to 499	15,239	9,345	2,811	1,582	1,501
		500 to 999	15,195	8,362	2,856	2,673	1,304
		1,000+	1,768	676	661	253	177
Total			96,212	42,769	20,214	16,579	16,649

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.
[2]Aggregate measure of size is equal to the sum of the square root of the enrollment of the schools in each type-of-locale, enrollment size, and minority status cell.
SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-8E. Aggregate measure of size of elementary schools in SASS/CCD frame, by type of locale, enrollment size, and region: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Measure of size (row total)[2]	Region[1]			
				North-east	South-east	Central	West
Elementary	City	Less than 300	31,305	5,511	4,581	12,202	9,012
		300 to 499	105,237	16,879	20,565	33,376	34,417
		500 to 999	177,864	30,429	38,455	32,271	76,710
		1,000+	26,364	7,858	4,329	2,965	11,212
	Urban fringe	Less than 300	35,729	12,561	4,366	10,895	7,907
		300 to 499	123,411	41,147	15,149	37,544	29,571
		500 to 999	204,484	45,027	43,422	34,407	81,628
		1,000+	17,302	2,240	6,952	749	7,360
	Town	Less than 300	30,164	2,809	3,931	14,656	8,769
		300 to 499	49,171	3,892	12,869	15,591	16,819
		500 to 999	41,681	2,866	19,062	7,488	12,265
		1,000+	2,341	167	1,696	167	311
	Rural	Less than 300	80,696	10,376	12,704	35,947	21,669
		300 to 499	58,560	9,952	17,351	19,724	11,532
		500 to 999	44,105	9,992	18,076	9,042	6,995
		1,000+	1,376	387	782	69	138
Total			1,029,790	202,092	224,290	267,093	336,316

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.
[2]Aggregate measure of size is equal to the sum of the square root of the enrollment of the schools in each type-of-locale, enrollment size, and region subgroup.
SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-8F. Aggregate measure of size of middle schools in SASS/CCD frame, by type of locale, enrollment size, and region: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Measure of size (row total)[2]	Region[1]			
				North-east	South-east	Central	West
Middle	City	Less than 300	3,527	696	581	1,132	1,118
		300 to 499	9,713	1,809	2,033	3,756	2,115
		500 to 999	60,193	9,426	13,458	15,097	22,213
		1,000+	30,051	5,718	7,468	1,962	14,904
	Urban fringe	Less than 300	7,133	1,955	640	2,530	2,007
		300 to 499	20,878	6,547	2,427	7,050	4,853
		500 to 999	82,029	23,357	14,043	20,223	24,406
		1,000+	32,430	6,072	10,493	3,521	12,343
	Town	Less than 300	8,565	397	1,314	3,354	3,499
		300 to 499	18,144	1,377	5,564	5,292	5,911
		500 to 999	28,515	2,314	10,577	6,985	8,639
		1,000+	2,811	270	1,174	690	676
	Rural	Less than 300	22,580	1,169	1,796	10,393	9,222
		300 to 499	15,186	2,637	4,771	4,900	2,878
		500 to 999	13,635	3,748	5,014	2,563	2,310
		1,000+	1,348	609	498	134	107
Total			356,738	68,100	81,852	89,582	117,203

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.
[2]Aggregate measure of size is equal to the sum of the square root of the enrollment of the schools in each type-of-locale, enrollment size, and region subgroup.
SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-8G. Aggregate measure of size of secondary schools in SASS/CCD frame, by type of locale, enrollment size, and region: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Measure of size (row total)[2]	Region[1]			
				North-east	South-east	Central	West
Secondary	City	Less than 300	2,448	403	403	896	746
		300 to 499	1,679	485	384	425	384
		500 to 999	9,902	2,408	2,456	3,092	1,946
		1,000+	75,407	11,395	17,200	16,254	30,559
	Urban fringe	Less than 300	2,468	368	258	801	1,041
		300 to 499	5,507	1,533	624	2,145	1,204
		500 to 999	29,831	12,266	4,215	8,839	4,510
		1,000+	86,854	18,917	18,314	18,564	31,059
	Town	Less than 300	2,956	54	494	878	1,530
		300 to 499	9,140	623	2,028	3,620	2,870
		500 to 999	23,889	2,516	7,490	7,726	6,157
		1,000+	18,509	1,141	6,475	4,607	6,287
	Rural	Less than 300	21,647	604	734	9,901	10,408
		300 to 499	14,868	1,819	3,102	6,425	3,522
		500 to 999	16,403	3,586	6,002	4,789	2,026
		1,000+	6,425	1,867	2,891	599	1,069
Total			327,933	59,984	73,070	89,562	105,318

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.

[2]Aggregate measure of size is equal to the sum of the square root of the enrollment of the schools in each type-of-locale, enrollment size, and region subgroup.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-8H. Aggregate measure of size of combined schools in SASS/CCD frame, by type of locale, enrollment size, and region: 1997–1998

Instructional level	Type of locale	Enrollment size of school	Measure of size (row total)[2]	Region[1]			
				North-east	South-east	Central	West
Combined	City	Less than 300	2,188	244	419	651	873
		300 to 499	1,026	241	121	281	383
		500 to 999	2,827	570	788	951	518
		1,000+	5,754	793	2,417	1,612	932
	Urban fringe	Less than 300	1,950	228	344	652	726
		300 to 499	2,321	942	437	680	263
		500 to 999	7,419	3,781	1,797	1,332	509
		1,000+	6,308	1,621	2,436	1,056	1,195
	Town	Less than 300	1,398	23	502	409	464
		300 to 499	2,332	101	963	908	360
		500 to 999	5,107	675	2,738	1,298	396
		1,000+	2,123	305	1,410	365	43
	Rural	Less than 300	23,257	921	2,868	11,174	8,294
		300 to 499	15,239	2,675	4,770	6,001	1,793
		500 to 999	15,195	4,357	7,052	2,986	801
		1,000+	1,768	370	1,113	208	78
Total			96,212	17,846	30,174	30,565	17,626

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.

[2]Aggregate measure of size is equal to the sum of the square root of the enrollment of the schools in each type-of-locale, enrollment size, and region subgroup.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-9A. Allocation of the elementary school sample, by type of locale, enrollment size, and minority status: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Elementary	City	Less than 300	23	2	5	6	9
		300 to 499	77	4	14	21	37
		500 to 999	130	3	16	32	79
		1,000+	19	0	1	2	16
	Urban fringe	Less than 300	26	9	9	5	3
		300 to 499	90	20	33	22	15
		500 to 999	149	22	48	39	40
		1,000+	13	1	3	3	6
	Town	Less than 300	22	8	8	3	2
		300 to 499	36	10	11	8	7
		500 to 999	30	8	7	8	8
		1,000+	2	0	0	1	1
	Rural	Less than 300	59	36	12	6	5
		300 to 499	43	22	10	6	4
		500 to 999	32	14	9	6	3
		1,000+	1	0	0	0	0
Total			750	159	187	168	236

[1] Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.

[2] Sample sizes are the expected numbers of respondents. Entries in this table were obtained by allocating the total sample size of 750 elementary schools to strata defined by type of locale and enrollment size class (rows) in proportion to the aggregate square root of the enrollment in the stratum (see table 2-8A). That is, the measure of size for a stratum (row) was divided by the total measure of size, and the result was multiplied by 750. The sample size for the stratum was then distributed to the four percent minority groups in proportion to the number of schools in the group.

NOTE: Detail may not add to totals because of rounding.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-9B. Allocation of the middle school sample, by type of locale, enrollment size, and minority status: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Middle	City	Less than 300	10	1	1	2	6
		300 to 499	27	2	5	6	14
		500 to 999	169	8	32	48	81
		1,000+	84	1	11	23	49
	Urban fringe	Less than 300	20	6	6	5	3
		300 to 499	59	19	20	12	8
		500 to 999	230	43	90	57	40
		1,000+	91	9	30	29	23
	Town	Less than 300	24	9	7	4	4
		300 to 499	51	15	15	12	9
		500 to 999	80	19	24	20	17
		1,000+	8	2	3	2	2
	Rural	Less than 300	63	33	15	9	7
		300 to 499	43	22	10	6	5
		500 to 999	38	17	12	6	3
		1,000+	4	1	1	1	0
Total			1,000	207	279	242	271

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.

[2]Sample sizes are the expected numbers of respondents. Entries in this table were obtained by allocating the total sample size of 1,000 middle schools to strata defined by type of locale and enrollment size class (rows) in proportion to the aggregate square root of the enrollment in the stratum (see table 2-8B). That is, the measure of size for a stratum (row) was divided by the total measure of size, and the result was multiplied by 1,000. The sample size for thc stratum was then distributed to the four percent minority groups in proportion to the number of schools in the group.

NOTE: Detail may not add to totals because of rounding.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-9C. Allocation of the secondary school sample, by type of locale, enrollment size, and minority status: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Secondary	City	Less than 300	7	1	1	2	4
		300 to 499	5	1	1	1	2
		500 to 999	30	3	4	6	17
		1,000+	230	8	39	70	113
	Urban fringe	Less than 300	8	2	2	2	2
		300 to 499	17	8	5	3	1
		500 to 999	91	34	34	15	8
		1,000+	265	32	96	81	56
	Town	Less than 300	9	2	3	2	2
		300 to 499	28	12	7	6	4
		500 to 999	73	29	21	13	9
		1,000+	56	13	20	14	10
	Rural	Less than 300	66	34	15	10	6
		300 to 499	45	27	9	6	3
		500 to 999	50	28	11	7	4
		1,000+	20	6	7	5	1
Total			1,000	240	274	242	244

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.

[2]Sample sizes are the expected numbers of respondents. Entries in this table were obtained by allocating the total sample size of 1,000 secondary schools to strata defined by type of locale and enrollment size class (rows) in proportion to the aggregate square root of the enrollment in the stratum (see table 2-8C). That is, the measure of size for a stratum (row) was divided by the total measure of size, and the result was multiplied by 1,000. The sample size for thc stratum was then distributed to the four percent minority groups in proportion to the number of schools in the group.

NOTE: Detail may not add to totals because of rounding.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-9D. Allocation of the combined school sample, by type of locale, enrollment size, and minority status: 2000

Instructional Level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Combined	City	Less than 300	6	0	1	1	3
		300 to 499	3	1	0	0	1
		500 to 999	7	1	2	2	4
		1,000+	15	1	2	4	9
	Urban fringe	Less than 300	5	2	1	1	1
		300 to 499	6	3	2	1	0
		500 to 999	19	8	5	5	2
		1,000+	16	4	5	4	3
	Town	Less than 300	4	2	1	1	1
		300 to 499	6	3	2	1	1
		500 to 999	13	5	3	3	2
		1,000+	6	1	1	2	1
	Rural	Less than 300	60	34	12	7	8
		300 to 499	40	24	7	4	4
		500 to 999	39	22	7	7	3
		1,000+	5	2	2	1	0
Total			250	111	53	43	43

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.

[2]Sample sizes are the expected numbers of respondents. Entries in this table were obtained by allocating the total sample size of 250 combined schools to strata defined by type of locale and enrollment size class (rows) in proportion to the aggregate square root of the enrollment in the stratum (see table 2-8D). That is, the measure of size for a stratum (row) was divided by the total measure of size, and the result was multiplied by 250. The sample size for thc stratum was then distributed to the four percent minority groups in proportion to the number of schools in the group.

NOTE: Detail may not add to totals because of rounding.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-9E. Allocation of the elementary school sample, by type of locale, enrollment size, and region: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Region[1]			
				North-east	South-east	Central	West
Elementary	City	Less than 300	23	4	3	9	7
		300 to 499	77	12	15	24	25
		500 to 999	130	22	28	24	56
		1,000+	19	6	3	2	8
	Urban fringe	Less than 300	26	9	3	8	6
		300 to 499	90	30	11	27	21
		500 to 999	149	33	32	25	59
		1,000+	13	2	5	1	5
	Town	Less than 300	22	2	3	11	6
		300 to 499	36	3	9	11	12
		500 to 999	30	2	14	5	9
		1,000+	2	0	1	0	0
	Rural	Less than 300	59	8	9	26	15
		300 to 499	43	7	13	14	8
		500 to 999	32	7	13	7	5
		1,000+	1	0	1	0	0
Total			750	147	163	195	244

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.

[2]Sample sizes are the expected numbers of respondents. Entries in this table were obtained by allocating the total sample size of 750 elementary schools to strata defined by type of locale and enrollment size class (rows) in proportion to the aggregate square root of the enrollment in the stratum (see table 2-8E). That is, the measure of size for a stratum (row) was divided by the total measure of size, and the result was multiplied by 750. The sample size for the stratum was then distributed to the four regions in proportion to the number of schools in the region.

NOTE: Detail may not add to totals because of rounding.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-9F. Allocation of the middle school sample, by type of locale, enrollment size, and region: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Region[1]			
				North-east	South-east	Central	West
Middle	City	Less than 300	10	2	2	3	3
		300 to 499	27	5	6	11	6
		500 to 999	169	26	38	42	62
		1,000+	84	16	21	6	42
	Urban fringe	Less than 300	20	5	2	7	6
		300 to 499	59	18	7	20	14
		500 to 999	230	66	39	57	68
		1,000+	91	17	29	10	34
	Town	Less than 300	24	1	4	9	10
		300 to 499	51	4	16	15	17
		500 to 999	80	6	30	20	24
		1,000+	8	1	3	2	2
	Rural	Less than 300	63	3	5	29	25
		300 to 499	43	7	13	14	8
		500 to 999	38	11	14	7	6
		1,000+	4	2	1	0	0
Total			1,000	191	229	252	327

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.

[2]Sample sizes are the expected numbers of respondents. Entries in this table were obtained by allocating the total sample size of 1,000 middle schools to strata defined by type of locale and enrollment size class (rows) in proportion to the aggregate square root of the enrollment in the stratum (see table 2-8F). That is, the measure of size for a stratum (row) was divided by the total measure of size, and the result was multiplied by 1,000. The sample size for the stratum was then distributed to the four regions in proportion to the number of schools in the region.

NOTE: Detail may not add to totals because of rounding.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-9G. Allocation of the secondary school sample, by type of locale, enrollment size, and region: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Region[1]			
				North-east	South-east	Central	West
Secondary	City	Less than 300	7	1	1	3	2
		300 to 499	5	1	1	1	1
		500 to 999	30	7	7	9	6
		1,000+	230	35	54	52	89
	Urban fringe	Less than 300	8	1	1	2	3
		300 to 499	17	5	2	7	4
		500 to 999	91	37	13	27	14
		1,000+	265	61	56	58	90
	Town	Less than 300	9	0	2	3	5
		300 to 499	28	2	6	11	9
		500 to 999	73	8	23	24	19
		1,000+	56	4	20	14	19
	Rural	Less than 300	66	2	2	30	32
		300 to 499	45	6	9	20	11
		500 to 999	50	11	18	15	6
		1,000+	20	6	9	2	3
Total			1,000	186	225	277	312

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.

[2]Sample sizes are the expected numbers of respondents. Entries in this table were obtained by allocating the total sample size of 1,000 secondary schools to strata defined by type of locale and enrollment size class (rows) in proportion to the aggregate square root of the enrollment in the stratum (see table 2-8G). That is, the measure of size for a stratum (row) was divided by the total measure of size, and the result was multiplied by 1,000. The sample size for the stratum was then distributed to the four regions in proportion to the number of schools in the region.

NOTE: Detail may not add to totals because of rounding.

SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-9H. Allocation of the combined school sample, by type of locale, enrollment size, and region: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Region[1]			
				North-east	South-east	Central	West
Combined	City	Less than 300	6	1	1	2	2
		300 to 499	3	1	0	1	1
		500 to 999	7	1	2	2	1
		1,000+	15	2	6	4	2
	Urban fringe	Less than 300	5	1	1	2	2
		300 to 499	6	2	1	2	1
		500 to 999	19	10	5	3	1
		1,000+	16	5	7	2	3
	Town	Less than 300	4	0	1	1	1
		300 to 499	6	0	3	2	1
		500 to 999	13	2	7	3	1
		1,000+	6	1	4	1	0
	Rural	Less than 300	60	2	8	28	22
		300 to 499	40	7	12	16	5
		500 to 999	39	11	18	8	2
		1,000+	5	1	3	1	0
Total			250	47	79	78	46

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.
[2]Sample sizes are the expected numbers of respondents. Entries in this table were obtained by allocating the total sample size of 250 combined schools to strata defined by type of locale and enrollment size class (rows) in proportion to the aggregate square root of the enrollment in the stratum (see table 2-8H). That is, the measure of size for a stratum (row) was divided by the total measure of size, and the result was multiplied by 250. The sample size for the stratum was then distributed to the four regions in proportion to the number of schools in the region.
NOTE: Detail may not add to totals because of rounding.
SOURCE: Special tabulations from the sampling frame for the U.S. Department of Education, National Center for Education Statistics, School and Staffing Survey, 1999–2000, which was based on the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1997–98 data file.

Table 2-10A. Speculated response rates, by type of locale, enrollment size, and minority status

Instructional Level	Type of locale	Enrollment size of school	Percentage minority enrollment*			
			Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
All levels	City	Less than 300	95	92	86	83
		300 to 499	95	92	86	83
		500 to 999	95	92	86	83
		1,000+	90	85	80	80
	Urban fringe	Less than 300	96	95	91	87
		300 to 499	95	93	90	87
		500 to 999	95	93	90	86
		1,000+	92	90	87	85
	Town	Less than 300	96	95	91	87
		300 to 499	95	93	89	86
		500 to 999	95	93	89	86
		1,000+	92	92	90	85
	Rural	Less than 300	96	95	93	90
		300 to 499	96	95	93	90
		500 to 999	96	95	93	90
		1,000+	96	95	93	90

*Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.
SOURCE: Speculated response rates are rough estimates based on response rates achieved in 1996–97 FRSS survey on school violence. These estimates are made for the U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-10B. Speculated response rates, by type of locale, enrollment size, and region

Instructional Level	Type of locale	Enrollment size of school	Region*			
			North-East (%)	South-East (%)	Central (%)	West (%)
All levels	City	Less than 300	85	87	87	85
		300 to 499	85	87	87	84
		500 to 999	84	89	89	83
		1,000+	80	83	83	80
	Urban fringe	Less than 300	92	95	95	91
		300 to 499	91	93	93	89
		500 to 999	91	93	93	89
		1,000+	89	89	90	86
	Town	Less than 300	92	94	95	92
		300 to 499	90	93	93	89
		500 to 999	90	91	92	89
		1,000+	88	90	90	87
	Rural	Less than 300	94	96	97	93
		300 to 499	93	95	95	92
		500 to 999	94	95	95	92
		1,000+	92	94	94	91

*Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.

NOTE: Detail may not add to totals because of rounding.

SOURCE: Speculated response rates are rough estimates based on response rates achieved in 1996–97 FRSS survey on school violence. These estimates are made for the U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-11A. Number of elementary schools to be sampled including allowance for nonresponse, by type of locale, enrollment size, and minority status: 2000

Instructional level	Type of locale	Enrollment size of school	Sample size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Elementary	City	Less than 300	26	2	6	7	11
		300 to 499	89	5	16	24	45
		500 to 999	152	3	17	37	95
		1,000+	24	0	1	3	20
	Urban fringe	Less than 300	28	9	10	5	4
		300 to 499	98	21	36	24	17
		500 to 999	165	23	52	43	47
		1,000+	14	1	3	4	7
	Town	Less than 300	23	9	8	4	3
		300 to 499	39	10	12	9	8
		500 to 999	34	8	8	9	9
		1,000+	2	0	0	1	1
	Rural	Less than 300	62	37	13	6	5
		300 to 499	45	23	11	7	5
		500 to 999	34	15	9	6	4
		1,000+	1	0	0	0	0
Total			837	167	200	189	280

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.
[2]Sample sizes are the numbers of schools to be selected. Entries in this table were obtained by dividing the target sample sizes in table 2-9A by the corresponding response rate in table 2-10A.
NOTE: Detail may not add to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-11B. Number of middle schools to be sampled including allowance for nonresponse, by type of locale, enrollment size, and minority status: 2000

Instructional level	Type of locale	Enrollment size of school	Sample size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Middle	City	Less than 300	12	1	1	2	7
		300 to 499	32	2	5	7	17
		500 to 999	196	8	35	55	98
		1,000+	104	1	12	29	62
	Urban fringe	Less than 300	21	7	6	5	4
		300 to 499	64	20	21	13	9
		500 to 999	252	45	97	63	47
		1,000+	103	10	33	33	27
	Town	Less than 300	26	9	7	5	5
		300 to 499	56	16	16	13	11
		500 to 999	88	20	26	23	19
		1,000+	9	2	3	2	2
	Rural	Less than 300	67	34	16	10	7
		300 to 499	45	23	10	6	6
		500 to 999	40	18	13	7	3
		1,000+	4	1	1	1	0
Total			1,119	218	302	276	323

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.
[2]Sample sizes are the numbers of schools to be selected. Entries in this table were obtained by dividing the target sample sizes in table 2-9B by the corresponding response rate in table 2-10A.
NOTE: Detail may not add to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-11C. Number of secondary schools to be sampled including allowance for nonresponse, by type of locale, enrollment size, and minority status: 2000

Instructional level	Type of locale	Enrollment size of school	Sample size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Secondary	City	Less than 300	9	1	1	2	5
		300 to 499	6	1	1	1	3
		500 to 999	35	4	4	7	20
		1,000+	283	9	46	88	141
	Urban fringe	Less than 300	8	2	2	2	2
		300 to 499	18	8	5	4	1
		500 to 999	98	36	37	17	9
		1,000+	300	35	106	93	66
	Town	Less than 300	10	3	3	3	2
		300 to 499	30	12	7	7	4
		500 to 999	79	31	22	15	11
		1,000+	63	14	21	15	12
	Rural	Less than 300	70	36	16	11	7
		300 to 499	48	28	10	6	4
		500 to 999	53	29	12	7	5
		1,000+	21	7	8	5	2
Total			1,131	254	301	282	294

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.
[2]Sample sizes are the numbers of schools to be selected. Entries in this table were obtained by dividing the target sample sizes in table 2-9C by the corresponding response rate in table 2-10A.
NOTE: Detail may not add to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-11D. Number of combined schools to be sampled including allowance for nonresponse, by type of locale, enrollment size, and minority status: 2000

Instructional level	Type of locale	Enrollment size of school	Sample size (row total)[2]	Percentage minority enrollment[1]			
				Less than 5 percent	5 to 19 percent	20 to 49 percent	50 percent or more
Combined	City	Less than 300	7	1	1	2	4
		300 to 499	3	1	0	1	2
		500 to 999	8	1	2	2	4
		1,000+	18	1	2	5	11
	Urban fringe	Less than 300	5	2	1	1	1
		300 to 499	6	3	2	1	1
		500 to 999	21	9	5	5	2
		1,000+	18	5	5	5	3
	Town	Less than 300	4	2	1	1	1
		300 to 499	7	3	2	1	1
		500 to 999	15	5	3	3	3
		1,000+	6	1	1	2	1
	Rural	Less than 300	64	35	13	7	9
		300 to 499	42	25	8	4	4
		500 to 999	42	23	8	7	4
		1,000+	5	2	2	1	1
Total			271	116	56	48	51

[1]Categories are based on information in the 1997–98 CCD file and are intended to illustrate the variation in percentage minority enrollment. Schools for which minority enrollment is missing in the CCD file are included in the "less than 5 percent" category. For analysis purposes, categories based on reported minority enrollment should be used and need not coincide with those given here.
[2]Sample sizes are the numbers of schools to be selected. Entries in this table were obtained by dividing the target sample sizes in table 2-9D by the corresponding response rate in table 2-10A.
NOTE: Detail may not add to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-11E. Number of elementary schools to be sampled including allowance for nonresponse, by type of locale, enrollment size, and region: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Region[1]			
				North-East	South-east	Central	West
Elementary	City	Less than 300	26	5	4	10	8
		300 to 499	89	14	17	28	30
		500 to 999	152	26	32	26	67
		1,000+	24	7	4	3	10
	Urban fringe	Less than 300	28	10	3	8	6
		300 to 499	98	33	12	29	24
		500 to 999	165	36	34	27	67
		1,000+	14	2	6	1	6
	Town	Less than 300	23	2	3	11	7
		300 to 499	39	3	10	12	14
		500 to 999	34	2	15	6	10
		1,000+	2	0	1	0	0
	Rural	Less than 300	62	8	10	27	17
		300 to 499	45	8	13	15	9
		500 to 999	34	8	14	7	6
		1,000+	1	0	1	0	0
Total			836	165	178	212	281

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.
[2]Sample sizes are the numbers of schools to be selected. Entries in this table were obtained by dividing the target sample sizes in table 2-9E by the corresponding response rate in table 2-10B.
NOTE: Detail may not add to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-11F. Number of middle schools to be sampled including allowance for nonresponse, by type of locale, enrollment size, and region: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Region[1]			
				North-East	South-east	Central	West
Middle	City	Less than 300	12	2	2	4	4
		300 to 499	32	6	7	12	7
		500 to 999	196	31	42	48	75
		1,000+	104	20	25	7	52
	Urban fringe	Less than 300	21	6	2	8	6
		300 to 499	64	20	7	21	15
		500 to 999	252	72	42	61	77
		1,000+	103	19	33	11	40
	Town	Less than 300	26	1	4	10	11
		300 to 499	56	4	17	16	19
		500 to 999	88	7	32	21	27
		1,000+	9	1	4	2	2
	Rural	Less than 300	67	4	5	30	27
		300 to 499	45	8	14	14	9
		500 to 999	40	11	15	8	7
		1,000+	4	2	2	0	0
Total			1,119	215	252	274	378

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.
[2]Sample sizes are the numbers of schools to be selected. Entries in this table were obtained by dividing the target sample sizes in table 2-9F by the corresponding response rate in table 2-10B.
NOTE: Detail may not add to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-11G. Number of secondary schools to be sampled including allowance for nonresponse, by type of locale, enrollment size, and region: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Region[1]			
				North-east	South-east	Central	West
Secondary	City	Less than 300	9	1	1	3	3
		300 to 499	6	2	1	1	1
		500 to 999	35	9	8	11	7
		1,000+	283	44	65	62	111
	Urban fringe	Less than 300	8	1	1	3	4
		300 to 499	18	5	2	7	4
		500 to 999	98	41	14	29	15
		1,000+	300	68	63	64	105
	Town	Less than 300	10	0	2	3	5
		300 to 499	30	2	7	12	10
		500 to 999	79	8	25	25	21
		1,000+	63	4	22	16	21
	Rural	Less than 300	70	2	2	31	34
		300 to 499	48	6	10	21	12
		500 to 999	53	12	19	15	7
		1,000+	21	6	9	2	4
Total			1,131	211	252	305	363

[1]Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.
[2]Sample sizes are the numbers of schools to be selected. Entries in this table were obtained by dividing the target sample sizes in table 2-9G by the corresponding response rate in table 2-10B.
NOTE: Detail may not add to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-11H. Number of combined schools to be sampled including allowance for nonresponse, by type of locale, enrollment size, and region: 2000

Instructional level	Type of locale	Enrollment size of school	Target sample size (row total)[2]	Region[1]			
				North-East	South-east	Central	West
Combined	City	Less than 300	7	1	1	2	3
		300 to 499	3	1	0	1	1
		500 to 999	8	2	2	3	2
		1,000+	18	3	8	5	3
	Urban fringe	Less than 300	5	1	1	2	2
		300 to 499	6	3	1	2	1
		500 to 999	21	11	5	4	1
		1,000+	18	5	7	3	3
	Town	Less than 300	4	0	1	1	1
		300 to 499	7	0	3	3	1
		500 to 999	15	2	8	4	1
		1,000+	6	1	4	1	0
	Rural	Less than 300	64	3	8	29	24
		300 to 499	42	7	13	16	5
		500 to 999	42	13	19	8	2
		1,000+	5	1	3	1	0
Total			271	52	85	83	51

[1] Regions are the four regions defined for the National Assessment of Educational Progress (NAEP). The northeast region consists of Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The central region consists of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin. The west region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming.
[2] Sample sizes are the numbers of schools to be selected. Entries in this table were obtained by dividing the target sample sizes in table 2-9H by the corresponding response rate in table 2-10B.
NOTE: Detail may not add to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Selection of the Sample

The period of data collection for the SSOCS:2000 coincided or overlapped with a number of other ongoing NCES studies, e.g., the 1999-2000 Schools and Staffing Survey (SASS), the Early Childhood Longitudinal Study-Kindergarten (ECLS-K), the National Assessment of Educational Progress (NAEP), and a Fast Response Survey System (FRSS) survey on teacher quality issues. NCES was concerned that the large sample sizes for these surveys might result in substantial overlap if the samples were drawn independently, potentially placing excessive burden on some schools. To reduce respondent burden, NCES wished to avoid selecting the same school for more than one of these studies to the extent feasible. This section describes the procedures developed for SSOCS:2000 to minimize the sample overlap with the other NCES surveys. In the discussion that follows, the NAEP and ECLS-K samples are treated as a single sample because they were selected independently.

The approach used to minimize overlap between SSOCS:2000 and the other NCES surveys is an extension of procedures developed and currently used by NCES to minimize overlap between SASS and NAEP/ECLS-K. To illustrate the general approach used in the SASS, let $P_i(s_1)$ denote the probability that school i is selected for NAEP/ECLS-K (say), and let $P_i(s_2)$ denote the corresponding desired probability of selecting the school for SASS. Similarly, let $P_i(\overline{s}_1)$ denote the probability that school i is not selected for NAEP/ECLS-K and let $P_i(\overline{s}_2)$ denote the probability that school i is not selected for SASS. The goal is to select the SASS sample in a way that minimizes overlap with NAEP/ECLS-K. This is accomplished by assigning schools a "conditional" probability of selection $P_i(s_2 \mid NAEP/ECLS\text{-}K)$ according to the following rules:

If $P_i(s_1) + P_i(s_2) <= 1$, then set $P_i(s_2 \mid NAEP/ECLS\text{-}K) = \dfrac{P_i(s_2)}{P_i(\overline{s}_1)}$ if school i was *not* selected for NAEP/ECLS-K; otherwise, set $P_i(s_2 \mid NAEP/ECLS\text{-}K) = 0$.

If $P_i(s_1) + P_i(s_2) > 1$, then set $P_i(s_2 \mid NAEP/ECLS\text{-}K) = 1$ if school i was *not* selected for NAEP/ECLS-K; otherwise, set $P_i(s_2 \mid NAEP/ECLS\text{-}K) = \dfrac{P_i(s_2) - P_i(\overline{s}_1)}{1 - P_i(\overline{s}_1)}$.

It can be shown that when the schools are selected with the conditional probabilities described above, the resulting sample retains the desired probabilities of selection specified for SASS, while achieving minimal overlap with the other NCES samples. For the SSOCS:2000, this procedure was extended to minimize overlap with the SASS, NAEP/ECLS-K, and the FRSS survey on teacher quality

samples. Implementation of the overlap minimization procedure required that appropriate conditional selection probabilities (using formulas analogous to those described above) be calculated and assigned to each school in the sampling frame. The use of these conditional selection probabilities ensured that the number of overlapping schools would be kept to a minimum, while achieving the desired probabilities of selection for SSOCS:2000. Details of the overlap minimization procedures developed for SSOCS:2000 sample selection are given in appendix A.

To select the sample, schools in the frame were sorted into sampling strata defined by level, type of locale, and enrollment size class, and then by minority status and region within stratum. Sorting by minority status and region within the sampling strata induced additional implicit stratification[12] of the frame, and was desirable because estimates of the prevalence and numbers of some crime incidents varied by these characteristics in the FRSS survey on school violence. For example, in tables 2-5 and 2-6, it can be seen that the percentage of low-minority schools (those with less than 5 percent minority enrollment) reporting physical attacks with weapons, robbery, physical attacks without weapons, thefts, and vandalism is significantly lower than the corresponding percentages of high minority schools (50 percent or more minority enrollment). Similarly, there is significant regional variation in the percentage of schools reporting certain types of incidents such as physical attacks with weapons, physical attacks without weapons, and vandalism, with generally higher rates of these types of incidents in the west than in other regions. The main function of the sorting was to ensure that the various minority status groups and regions were appropriately represented in the sample. Within each sampling sub-stratum, schools were then selected systematically and with probabilities proportionate to the conditional probabilities computed earlier using the formulas in appendix A. The resulting sample is summarized in table 2-12 by instructional level, type of locale and enrollment size class.

Summary of Sample Yields

Table 2-13 summarizes the number of responding schools and corresponding response rates by selected school-level characteristics. A total of 2,270 schools completed the survey for an overall (weighted) response rate of 70 percent. As indicated in table 2-13, response rates in the SSOCS:2000 varied by level (with somewhat lower response rates for elementary and middle schools than for secondary and combined schools), enrollment size of school (with generally lower response rates for large

[12] That is, when the schools are sorted by a variable such as region, then the proportion of schools selected per region will be roughly equivalent to that region's proportion of all schools. This is called implicit stratification because, although region is not a stratification variable, the geographic distribution of the sample will be similar to one obtained using explicit stratification. For example, selecting every tenth school will not necessarily result in an appropriate distribution across regions if the schools are not sorted by region, but it generally will if the schools are sorted.

schools than for smaller ones), and type of locale (lower response rates in city and urban fringe settings than in towns or rural areas). Response rates also varied by minority status (lower response rates in high minority schools than in others) and region (lower response rates in northeast and west than in the southeast and central regions). All of the differences cited above were tested using approximate design effects and procedures described in Approximate Sampling Errors in chapter 8, and were found to be statistically significant at the 95 percent confidence level. Additional details about the response rates achieved in the SSOCS:2000 are given in the report: Impact of Nonresponse on Estimates from the 2000 School Survey on Crime and Safety (SSOCS:2000) in appendix H.

Table 2-12. Number of schools selected for the SSOCS, by sampling strata defined by level, type of locale, and enrollment size class: 2000

Instructional level	Type of locale	Number of sample schools	Enrollment size class			
			Less than 300	300 to 499	500 to 999	1,000 or more
Elementary	City	292	27	89	152	24
	Urban fringe	306	28	98	165	15
	Town	99	24	39	34	2
	Rural	144	61	48	34	1
Middle	City	343	12	30	201	100
	Urban fringe	451	21	66	260	104
	Town	178	24	56	91	7
	Rural	157	67	48	37	5
Secondary	City	331	9	6	35	281
	Urban fringe	423	8	18	97	300
	Town	180	10	29	80	61
	Rural	191	70	48	52	21
Combined	City	35	7	4	9	15
	Urban fringe	48	4	7	21	16
	Town	30	4	7	14	5
	Rural	154	63	44	41	6
Total		3,362	439	637	1,323	963

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 2-13. Distribution of sample schools by response status and corresponding response rates, by selected school characteristics: 2000

Characteristic	Total	Resp-ondent	Nonresp-ondent	Ineligible	Unweighted response rate (percent)	Weighted response rate (percent)
Total	3,366*	2,270	1,044	52	68.5	70.0
Instructional level						
Elementary	841	565	266	10	68.0	69.0
Middle	1,131	749	368	14	67.1	69.7
Secondary	1,125	757	350	18	68.4	71.0
Combined	269	199	60	10	76.8	79.6
Enrollment size						
Less than 300	439	315	91	33	77.6	76.3
300 to 499	639	466	166	7	73.7	70.9
500 to 999	1,325	905	413	7	68.7	67.5
1,000 or more	963	584	374	5	61.0	61.1
Type of locale						
City	1,003	603	380	20	61.3	63.6
Urban fringe	1,228	810	407	11	66.6	67.5
Town	487	365	113	9	76.4	75.4
Rural	648	492	144	12	77.4	77.0
Percentage minority						
Less than 5 percent/miss	780	597	167	16	78.1	77.8
5 to 19 percent	885	624	253	8	71.2	71.3
20 to 49 percent	793	506	278	9	64.5	65.4
50 percent or more	908	543	346	19	61.1	64.6
Region						
Northeast	647	397	247	3	61.6	64.1
Southeast	772	548	212	12	72.1	74.0
Central	904	668	218	18	75.4	77.1
West	1,043	657	367	19	64.2	64.3

*Four of the originally sampled "schools" included separately administered elementary and secondary schools. The four "extra" schools were added to the sample.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

3. QUESTIONNAIRE DESIGN

Because SSOCS is intended to be a recurring survey, an extensive effort was devoted to developing the baseline questionnaire. The first part of the process was a literature review to determine both the substantive and methodological issues that were relevant to the survey design, along with a review of extant surveys on school crime to determine gaps in existing data, issues related to questionnaire construction, and potential items for use in SSOCS:2000. Based on these reviews, and on consultations with the Safe and Drug-Free Schools program within the U.S. Department of Justice and the Office of Special Education Programs within the U.S. Department of Education, a list of research objectives was developed along with a draft questionnaire. Both the research objectives and the questionnaire were reviewed by a Technical Review Panel consisting of researchers on school crime, educators, policymakers, and representatives of relevant education-related organizations. Also, in seeking input and subsequent endorsements for the survey, a large number of education organizations were contacted and provided drafts of the questionnaire. Several changes were made to the questionnaire based on these organizations' reviews.

The development of the questionnaire was an iterative process, with regular internal reviews and updates, external reviews by the TRP and governmental units, pretesting of the survey instrument with 14 schools (as described below), and review for clearance by the Office of Management and Budget and the Education Information Advisory Committee (EIAC) of the Council of Chief State School Officers.

Pretesting

After multiple revisions to the questionnaire (see appendix B for pretest materials. Note that the two versions of the questionnaire in this appendix differ from the final version of the questionnaire which is presented in appendix C), an initial pretest of the SSOCS:2000 was conducted. The purpose of the pretest was to determine that respondents understood all of the questionnaire items, that data were available, and that the level of burden was acceptable.

Sites for pretesting were chosen to provide diversity in instructional level, size, urbanicity, and region. Nine sites were chosen for the first pretest; however, because the TRP greatly changed the questionnaire while the pretest was in progress and there was a need to pretest those changes, the first pretest was conducted with five sites. Principals who were selected for the pretest were called on the

telephone, given a short description of the survey, and asked to participate in the pretest. If they agreed they were asked to complete the questionnaire and fill out a commentary guide indicating completion time, problem questions, undefined terms, and other comments about the questionnaire (such as the content, format, and appearance). After the questionnaire responses and comments were reviewed, an attempt was made to interview the pretest respondents by telephone to obtain further information about their comments, to determine the reasons for any problems that were identified in the questionnaire, and to answer a scripted set of questions that had been identified as issues to be resolved through the pretest. Three of the five who completed the questionnaire and the commentary guide went on to complete the follow-up (see first pretest telephone follow-up in appendix B).

The pretest led to comments about specific items as well as general comments about the questionnaire. The pretest respondents indicated that the survey was comprehensive and that it provided a good picture of the situation at their individual schools. Respondents were able to provide the data requested, except in a few instances where the schedule of the pretest did not allow them to contact key individuals for some information.

The amount of time required to complete the questionnaire averaged 89 minutes per person. The length of the questionnaire was greatly reduced after the initial pretest in order to reduce burden and to control survey administration costs. The reduction in length was based in part on the research priorities identified for the study and in part on the pretests (when respondents indicated that certain data were hard to provide or not meaningful for their schools).

Many of the comments regarding specific questionnaire items concerned clarifying question wording or making instructions more specific. For example, one respondent had trouble with question 1a ("During the 1999-2000 school year, did your school require visitors to sign or check in?") because the policy was currently being changed. Signs had been ordered to notify visitors to sign in, but they had not yet arrived. The instruction "If your school changed its practices in the middle of the school year, please answer regarding your most recent practice" was then added.

One of the substantive changes resulting from the pretest related to collecting data about zero tolerance policies. The pretest found there was little variation among schools in whether they had zero tolerance policies, but sometimes substantial variation in what those policies meant. Thus, the questions were not providing useful information in the judgment of the TRP. The TRP recommended dropping the questions and instead using a question (presently question 21) about disciplinary actions to determine the degree to which certain disciplinary actions were implemented automatically.

Following revisions, a second pretest was conducted. Again nine sites were selected for the pretest; one was replaced after it was unable to complete the questionnaire over the requested time interval. Completed questionnaires were received from eight of the sites and interviews were conducted with seven of the eight sites. The sites were chosen to have diversity in instructional level, size, urbanicity, and region.

In the second pretest telephone follow-up (see appendix B), the issues that were raised regarding specific questions were minor. The amount of time required to complete the questionnaire averaged 87-94 minutes per respondent. The second pretest, therefore, failed to show a reduction in the amount of time required to complete the survey compared with the first pretest (89 minutes), despite the fact that most of the problems with instructions and clarity appeared to have been fixed after the first pretest. Thus, the time required to complete the questionnaire appeared to reflect the type and quantity of data collected rather than problems in questionnaire construction. For this reason, substantial reductions were made in questionnaire length. In the second pretest, 7 of the 8 who completed the questionnaire and the commentary guide also completed the phone follow-up.

There were no questions that pretest respondents refused to answer. However, some respondents did ask for assurances of confidentiality and for information regarding the purpose of the study. Once that information was provided the respondents were willing to provide the requested information. Several comments indicated that it was important to have the right personnel available because a single person may not have all of the data requested. Some were unable to provide answers for that reason, and others found it helpful to gather a group together to answer the questionnaire.

Questionnaire Content

This section presents the content of the SSOCS:2000 questionnaire in detail.

Characteristics of School Policies

This section collects data about the nature of current school policies relating to crime and discipline. These data are important in order to help schools to know where they stand in relation to other schools, and to help policymakers to know which actions are already being taken and which actions might be encouraged in the future. Potentially, the data can also be used by researchers interested in evaluating the success of school policies. That is, though this study was not designed as an evaluation, the presence

of school policies can be correlated with the rates of crime provided elsewhere on the questionnaire, with appropriate controls for school characteristics. (Without controls for school characteristics, the data might lead to misleading results. For example, the schools that face the greatest problems may have the strongest policies. One might incorrectly infer that the strong policies are responsible for the high crime rates, when the more likely explanation may be that they are in response to the high crime rates. Adjusting for school characteristics helps to avoid such false findings.)

Question 1 asks about several kinds of school policies and practices:

- Items 1a through 1g ask about access to the school grounds. The ability of students and outsiders to enter and leave the campus throughout the school day affects the amount of control that administrators have over the school environment, and the potential for bringing weapons or drugs onto the campus.

- Items 1d, 1f, 1h-1j, 1o-1q, and 1s ask about ways that students are monitored to prevent crime. Such actions can directly affect crime because students may be more reluctant to engage in inappropriate activities for fear of being caught. The school climate also may be affected because students may feel more secure knowing that violators of school policies are likely to be caught.

- Items 1m and 1n ask how school policies regarding student conduct are communicated to students and parents. Adequately communicating the policies is a necessary first step in gaining compliance with the policies.

- Item 1n also asks about one aspect of parent involvement (i.e., whether and how parents are informed of school policies, with the expectation that parents will support those policies). Many believe that involving parents is a key way to prevent school crime. Communicating policies to parents is a necessary first step.

- Item 1r provides information about the school environment (e.g., are students and outsiders able to identify staff who might help with a problem?) and about the school's ability to monitor the grounds and identify outsiders.

- Item 1t asks about the availability of telephones in most classrooms. The availability of telephones affects teachers' ability to obtain help without leaving the classroom, and affects the administration's ability to communicate with teachers.

- Item 1u provides information about schools' compliance with federal laws on tobacco use. It also provides some information about the degree of discipline enforced in the school environment.

Question 2 asks about the existence of written plans for dealing with crises. When crises occur, there may not be time or an appropriate environment for making critical decisions, and key school leaders may not be available to immediately provide guidance. Thus, having a written plan is considered important in preparing schools to deal with crises effectively.

School Violence Prevention Programs and Practices

This section asks what programs schools may have to prevent or reduce violence. The presence of such programs is a sign that schools are being proactive by seeking to prevent violence before it occurs rather than reacting to it. The specific elements that are listed have been identified through past research as being the most prevalent. As with the first section, the data may be used by schools to know how they compare with other schools, and by policymakers who wish to know what programs are already in place.

Question 3 is a general question designed to provide an initial measure of the extent of school programs, while allowing schools that lack programs to skip irrelevant parts of the questionnaire.

Question 4 is based on research identifying which types of violence prevention programs directed toward students are most prevalent in schools.

Questions 5 and 6 examine school activities that are directed toward teachers or the environment (rather than students) to prevent or reduce violence. The items are taken from past research on school crime or from recommendations of the Technical Review Panel.

Questions 7 through 9 ask about the use of paid law enforcement or security services on the school grounds or at school events. The goal of one federal initiative is to help fund the presence of such police, so determining the frequency of using such personnel may help in guiding federal policy. Besides directly affecting school crime, the use of paid law enforcement personnel also affects the school environment; it may help to prevent illegal actions and to create a feeling of security among students. It also may affect (in either a positive or negative way) the feeling of freedom on school grounds. Thus, the times the law enforcement personnel are present, their visibility, and their carrying of weapons are all important.

Questions 10 and 11 ask about schools' actions to train teachers to identify potentially violent students. Schools now can obtain early warning signs to identify such potentially violent students, and their use of such profiles may affect both general levels of discipline and the potential for crises (such as multiple shootings). The involvement of teachers is important because teachers collectively spend the most time with students and observe students closely.

Question 12 asks for principals' perceptions of the factors that limit their efforts to reduce or prevent crime. Though principals are not trained evaluators, they are the people who are most

knowledgeable about the situations at their individual schools, and they know whether their own actions have been constrained by the factors listed. The pretest examined whether items l and m could be combined; it found that federal policies concerning disabled students are one of the most widely mentioned factors, but that other federal policies also are mentioned. Thus, the items were kept separate.

Violent Deaths at School and Elsewhere

Questions 13 and 14 ask about violent deaths. Violent deaths get substantial attention by the media but are actually relatively rare, and there is evidence that (in general) schools are much safer than students' neighboring communities. These questions help to verify the relative frequency of violent deaths at school and at other locations. Because violent deaths are rare, a skip pattern is used to simplify the questionnaire for most respondents.

The Frequency of Other Incidents at Schools

This section asks the frequency of various kinds of crime at school (other than violent deaths). The data can be used directly as an indicator of the degree of safety in U.S. public schools, and indirectly to rank schools in terms of the number of problems they face.

Question 15 asks about one of the areas where NCES most often receives questions: the number of shootings at schools. Previously no quantifiable statistics have been available.

Question 16 asks about the frequency of a number of crimes. By asking for both the total number of incidents and the number reported to police, it also provides information on how schools respond to crime. Hate crimes appear to be relatively rare, but are an important priority to the federal government. Gangs appear to be a growing problem in schools, so information is obtained about gang-related crimes as well.

Question 17 provides information about the degree to which crime changes from one year to another. If crimes are largely random events (in terms of which schools experience them), then policymakers may need different policies than if the crimes consistently occur only at some types of schools. Question 17 also may be used to adjust the responses to question 16 for schools that complete the questionnaire before the school year is completed, and thus have potentially provided incomplete data. The adjustments could be made by performing regression analysis to compare changes over time among

schools that reported for the whole year and those that did not, and thus to estimate the additional number of crimes that would be reported if the entire school year were included.

Question 18 asks about one aspect of school-wide costs of crime. Actions such as bomb threats not only affect student safety, but they affect the school environment as well. There is anecdotal evidence that these crimes may be increasing.

Disciplinary Problems and Actions

There is evidence that schools' ability to control crime is associated with their control of lesser violations. That is, lesser violations are an indication of the state of discipline in the school, so that when these violations are controlled, students do not progress to more serious disciplinary problems. This section asks about the degree to which schools face such disciplinary problems, and the way that they respond to them.

Question 19 asks about the frequency of seven different kinds of disciplinary problems. It provides a general measure of the degree to which there are disciplinary problems at each school.

Question 20 asks what kinds of disciplinary actions were available to each school, and whether they were actually used. It is not intended to be comprehensive, but rather focuses on some of the most important strategies. The data will help policymakers to know what options and what constraints principals face; for example, if an action is available in principle but not in practice, then policymakers would need to act in a different way than if the action is available but not used.

Question 21 asks about the number of various types of offenses committed by students, and the resulting disciplinary actions. This provides valuable information about how school policies are actually implemented (rather than simply what policies are in place). For example, many schools claim to have zero tolerance policies, but some schools have extremely strong policies while other zero tolerance policies allow so many options that there is little or no constraint on what disciplinary action is imposed. Question 21 provides a way of examining this issue by providing information on how many different kinds of actions are taken with regard to a particular offense, and how many times no action is taken.

Question 22 looks more specifically at the constraints potentially placed on schools' disciplinary actions by restrictions associated with the Individuals with Disabilities Education Act

(IDEA). It will help policymakers and researchers to know how often the restrictions actually result in different disciplinary actions than would have occurred otherwise.

School Characteristics

This section asks for a variety of types of information about the characteristics of the schools responding to the survey. This information is necessary in order to be able to understand the degree to which different schools face different situations. For example, one school might have highly effective programs and policies yet still have high crime rates because of the school's location in a high crime neighborhood; another school might appear to have effective policies based on its crime rates but actually have higher crime rates than similar schools. Note that the information requested in this section will be supplemented by data from the 1998-1999 Common Core of Data (CCD) — namely, by data on enrollment, race/ethnicity, the grade levels served, and the metropolitan status. Also, some items from question 1 (on school policies) will also provide information on the school disciplinary environment, so they also may be considered as providing school characteristics.

Question 23 asks for the total enrollment. A CCD measure of enrollment was used to draw the sample, but an updated measure is important because the level of school crime has been related to school size. The updated measure will also help to provide a more accurate measure of student-teacher ratios.

Question 24 provides information on the percentage of students receiving free or reduced-price lunches (24a, a measure of poverty), with limited English proficiency (24b, a measure of the cultural environment), in special education (24c, a measure of the academic environment), who are male (24d; most crimes are committed by males, so the percentage who are male can affect the overall crime rate), and with various levels of academic proficiency and interest (24e through 24g). All of these factors have been associated with crime rates.

Question 25 asks for the number of classroom changes made in a typical day. This is important because it affects schools' ability to control the student environment. When students are in hallways, there are more opportunities for problems. Also, a school with fewer classroom changes is likely to be more personal and to have closer relationships between the students and teachers.

Question 26 asks for the total number of paid staff in three categories. This can be used in combination with enrollment data to compute the student/faculty ratio (which is part of the academic

environment), and in combination with question 11 to compute the percentage of teachers involved in training to recognize early warning signs of potentially violent students. Counselors and special education teachers are especially likely to deal with "problem" students, so counts of these staff will help in knowing the resources that schools have for dealing with such students.

Question 27 provides information on the degree to which a school might be expected to have problems with crime based on the community where it is located. It thus provides a non-school-based way of comparing schools in similar situations.

Question 28 asks for the school type. Schools that target particular groups of students (such as magnet schools) have more control over who is in the student body, and may have better motivated students (because the students have chosen a particular program). Charter schools have more freedom than regular schools in their school policies, may have more control over who is admitted into the student body, and may have better motivated students (because the students chose the school).

Question 29 asks for the school's average unexcused absence rate. This is a measure of truancy and thus a measure of the level of disciplinary problems at the school. It also is a measure of the academic environment.

Question 30 asks for the number of transfers. When students transfer after the school year has started, schools have less control over whether and how the students are acculturated to the school. These students are likely to have less attachment to the school and to the other students, thus increasing the risk of disciplinary problems.

Questions 31 and 32 are used to examine whether schools that respond to the survey before the school year is completed report fewer crimes than schools reporting for the entire year. If so, then adjusted crime rates can be calculated (using multiple regression in combination with data from questions 14, 15, 16, 17, 21, and selected school characteristics). The final report presents both sets of numbers to establish upper and lower bounds for the number of crimes occurring at schools. Researchers who are evaluating the effectiveness of school policies and programs will probably prefer to use the adjusted measures so that their measures of crime rates will be more consistently defined across all schools.

Research Questions

The specific research questions addressed by the questionnaire are listed below.

1. What is the frequency of crime at public schools?
    - ♦ What is the number of incidents, by type of crime and location?
        - What violent deaths have occurred, and where?
        - How many crimes were reported to police?
    - ♦ How has the frequency of selected crimes changed over time?
    - ♦ What is the impact of crime on school activities?
    - ♦ What percentage of violent crime had a gang-related component?
    - ♦ What percentage of violent crime had a hate crime-related component?
2. What is the frequency of various types of disciplinary problems?
3. How many disciplinary actions have been taken, by type of action and type of offense?
    - ♦ How consistently are disciplinary actions performed?
4. What policies do schools have to prevent and respond to crime?

♦ How is access controlled to the campus?

♦ How are students monitored to prevent crime?

♦ Does the school have a code of conduct, and if so, how is it communicated to students and parents?

♦ Is there a crisis management plan, and if so, what events does it cover?

♦ What zero tolerance policies are in place? (by implication only)

♦ Are telephones available in classrooms?

5. What are the characteristics of school programs and practices to prevent or control crimes?
    - ♦ Do schools have formal programs to prevent or reduce violence?
    - ♦ What components are included in the school programs and practices?
        - Are teachers trained to identify potentially violent students?
    - ♦ What is the level of participation by teachers?
    - ♦ How many and what types of security personnel are used by the schools?
    - ♦ What factors limit the effectiveness of school programs?
6. What is the relationship of special education students to school crime?

- What disciplinary actions are taken with regard to special education students?
- What procedures are followed with regard to disciplining special education students?

School Characteristics

7. What characteristics of schools are correlated with crime and are needed to put the other answers in context?

    ♦ What are the school/student demographic characteristics?

- What is the total enrollment?
- What is the racial/ethnic composition?
- What are parents' economic resources (students' eligibility for free and reduced-price lunch)?
- What percentage of students do not speak English as their primary language?
- What grade levels are served?
- What is the metropolitan status of the area served by the school?
- How do schools describe the crime level of the surrounding community?

    ♦ What is the school environment like?

- What is the disciplinary environment like in terms of the level of expectations regarding students' behavior and students' general compliance with school rules and discipline?
- What is the academic environment like, in terms of the academic orientation and strength of the school and its students?
- Does the school target certain students?

4. DATA COLLECTION

This section presents an overview of the data collection procedures for the School Survey on Crime and Safety. Included are descriptions of survey mailout activities, receipt control, nonresponse follow-up, and interviewer training.

Mailout Activities

Data collection for SSOCS:2000 began on March 27, 2000. Survey packets were mailed to 3,362 elementary, middle, junior high, secondary, and combined public schools. (Four schools were added at a later date for a total of 3,366 schools.[13]) Each packet (see appendix C) contained a letter of introduction, the School Survey on Crime and Safety questionnaire, a brochure explaining the survey, a flyer identifying survey endorsements, and a Westat business reply envelope for returning the completed questionnaire.

Mailing labels were computer generated from a file containing identifying information including a unique 4-digit school ID assigned for tracking purposes. The letter to the principal did not contain the principal's name since this might have made the survey seem less confidential. For similar reasons, the mailing label on the envelope did not have the principal's name, but rather was addressed to Principal followed by the school name, Westat school ID number (in the upper right corner), and school address. This ID was also included on the questionnaire label. Mailout activities were conducted under the direction of the SSOCS:2000 operations supervisor, whose role was to ensure efficiency and quality control.

The letter of introduction, addressed generically to "Dear Principal", clearly explained the study and its purpose, stressed its importance, and stated that the data would be held in strict confidence. It asked that the questionnaire be returned by April 17, 2000. The letter indicated that some questions might require access to school records and that the person most knowledgeable about the school's disciplinary actions should complete the questionnaire. The principal was asked to complete questions 12 and 20, regardless of whether he/she was the individual responding to the other questions.

[13] One of the sampled schools was actually three schools housed in one building. Questionnaire packets were mailed to the two additional schools. Another sampled school was actually three schools in separate buildings located at the same address. Questionnaire packages were mailed to the two additional schools.

In addition to the mailout to sampled schools, there were two subsequent mailouts. In early April, all school superintendents who had at least one school sampled in their district were sent the same packet by Federal Express that was sent to all school principals. A cover letter (see appendix D) accompanying the packet described the SSOCS:2000 study and asked superintendents to encourage their schools to participate if they asked for authorization. The letter also provided reasons for maintaining confidentiality and not providing districts with names of the specific schools sampled. The second packet, also mailed in April, was addressed to the Chief State School Officer and was sent to all fifty states and the District of Columbia. A cover letter (see appendix D) for this packet contained similar information to that contained in the superintendent letter, and asked that the Officer encourage schools to participate in the study if his/her opinion was sought. Both letters included the name and toll-free telephone number of the SSOCS:2000 Project Director in case there were questions about the study.

All Postmaster returns were handled by a Westat employee experienced in tracing. For each case that was returned to Westat, a corrected or alternate address label was made and the packet was remailed.

Receipt Control

Westat's automated receipt control and status monitoring system tracked the flow of processing for each case in the study sample. The receipt control file was updated daily by one of two field room staff members as questionnaires were received during data collection. The receipt control file contained the following variables: NCES identifier, Westat ID, sampling strata needed for status reports, date(s) of questionnaire mailout, date and status of nonresponse follow-up, date of (and status codes reflecting) mail, fax, or telephone data collection, date of data editing, batch number for keyed data, and date and status of data retrieval for each case. (See next chapter for details about data preparation procedures.)

During the period of data collection, weekly status reports were prepared from the receipt system. These reports covered cases sampled and completed, ineligible cases, initial and final refusals, follow-up status, and other nonresponse.

Nonresponse Follow-Up

Nonresponse follow-up was required for approximately 2,914 cases (87%). Telephone follow-up for nonresponse began about 3 weeks after the questionnaires were mailed to the schools (i.e., the week of April 17). Each interviewer was given batches of nonresponse cases that included a Respondent Information Sheet (RIS) (see appendix E) and a corresponding call record (see appendix E) for each case. Each batch also contained a transmittal listing all ID's for that specific batch to aid the interviewer in keeping track of his/her caseload.

The top of the RIS indicated the school district name, principal's name, school name, address, and telephone number, and ID number assigned to the case. Interviewers used this section to identify respondents and conduct nonresponse follow-up calls for the SSOCS:2000 study. The lower section of the RIS contained a prepared script that interviewers were instructed to follow. Using this script, interviewers were prompted to ask for the principal, and were then led through a series of questions to determine the status of the questionnaire; that is, whether it had been mailed back and when, or whether the principal was still working on the survey. These follow-up calls were made to each school that had not returned the questionnaire. Interviewers recorded their information in spaces provided on the RIS.

If the school principal indicated that either he/she had not received the packet, or that it had been misplaced, the interviewer completed a remail request form. Remail requests were given to field room personnel who prepared the packets and sent them via Federal Express. As the end of data collection approached, packets were sent by fax instead, to insure prompt delivery. The school's mailing address was verified, and the respondent was asked for the best days and times to be contacted in the next few days to confirm receipt of the packet. In addition, interviewers tried to ascertain a date when the completed questionnaire would be mailed back. If the principal desired guidance on when the questionnaire could be returned, the interviewer asked if it could be completed within the next 2-3 weeks. The respondent was then recontacted if the questionnaire was not received by the agreed-upon date.

Initial instructions to interviewers were to make a maximum of 5 phone calls to a school. However, when time permitted, additional calls were attempted in an effort to boost the response rate. When a phone call went unanswered, but the school had voice mail, the interviewer left a message regarding the survey and provided the SSOCS:2000 toll-free 800 number. If the interviewer did not hear back from the school within 2-3 days after leaving the voice mail message, he/she attempted another phone call to the school. (If the school's answering machine message indicated the school was closed until a specific date, the interviewer called back on the date specified.)

The history of all telephone attempts and contacts was recorded on the corresponding call record for each nonresponse case. The result of each call was indicated by a specific interim or final status code.

Interviewers were monitored throughout the data collection process. Both group and individual meetings with interviewers and the SSOCS:2000 operations supervisor were held frequently. The purpose of these meetings was to check on caseload progress, offer suggestions, ask questions, and to provide feedback in general. When an interviewer encountered a problem in conducting a telephone prompt, he/she completed a problem sheet or spoke directly to the supervisor. In cases where the supervisor needed guidance, the Project Director was consulted.

Periodically throughout data collection, additional interviewers were added as needed. After signing a confidentiality form, each received thorough training from the SSOCS:2000 operations supervisor. (See Interviewer Training Procedures below.)

Data collection for SSOCS:2000 was originally scheduled to close on June 30, 2000 but was extended twice, ultimately to August 15, to boost the response rate. A number of questionnaires were received after the end of data collection, and some were later added in the process of editing the data.

A meeting to debrief interviewing staff was held on November 7, 2000, to review the telephone prompting and data retrieval processes. It was led by the SSOCS:2000 Project Director and attended by eight Westat interviewers, the SSOCS:2000 operations supervisor, and an NCES representative. This meeting focused on obtaining interviewers' observations about respondents' comments regarding the questionnaire, discussing respondent facility or difficulty answering specific questions, and any additional feedback the interviewers may have had.

Interviewer Training Procedures

The School Survey on Crime and Safety was staffed with interviewers selected from Westat's Education Area Telephone Operations Group (EATOG). Initial interviewer training was conducted at Westat's Rockville, Maryland offices on April 14, 2000. Training was led by Westat's Education Studies Operations Supervisor and the SSOCS:2000 Project Director and was also attended by NCES representatives.

Interviewers received project-specific training for SSOCS:2000 at a 4 hour training session. Training included the following: an overview of the study, a discussion regarding confidentiality, general procedures to use when contacting schools, recording information about assigned cases, successfully prompting schools to submit a completed questionnaire, and strategies for refusal avoidance. In addition, interviewers practiced answering likely respondent questions. Each interviewer was provided with a manual that followed the training agenda and which served as a reference during the follow-up operation.

5. DATA PREPARATION

This section presents an overview of data preparation procedures for the School Survey on Crime and Safety. Included are descriptions of the coding and editing specifications, range specifications, logic edits, frequency and cross-tabulation review, and frequency review of text items.

Coding and Data Retrieval

After questionnaires were received and entered into the receipt control system, they were manually coded. All coders were provided with code books and given project-specific training by the SSOCS:2000 operations supervisor. After the initial coding, each case was turned over to one of two other experienced coders who verified the coding for accuracy and consistency.

As part of the coding and editing process, questionnaires were reviewed for item nonresponse. The results were used for two purposes: to determine whether the questionnaires should be assigned to data retrieval to get more complete answers, and ultimately at the end of data collection, to determine whether sufficient data were collected to constitute a satisfactory questionnaire. Questionnaires were sent to data retrieval if any key item was missing data, or if more than 50 percent of the total number of items had missing data. In the final questionnaire approved by OMB, there were 259 items; thus, at least 130 items had to be completed for the questionnaire not to be sent to data retrieval. The actual number of items that were applicable varied from one school to another depending on the skip patterns involved. To simplify the computation of the percentage of items that were completed, legitimate skips were treated as valid responses (i.e., the respondent completed all of the items that were applicable). This avoided changing the denominator from one school to another. The practical effect was that questionnaires were sent to data retrieval if the number of missing responses (i.e., items for which there was neither a response nor a legitimate skip) exceeded 50 percent of the total (i.e., exceeded 129.) In the majority of cases in which questionnaires were missing greater than 50 percent of the total responses, either large blocks of questions were blank, or entire pages of the questionnaire had been purposefully or accidentally left blank. In these cases an attempt was made by interviewers to obtain responses to all missing questions. If a respondent did not have time to answer all questions, the interviewer's efforts were concentrated on obtaining answers to key items, demographic data, and enough of the remaining questions to be able to include the questionnaire in the database.

Following is a list of the key items. Any of these items that had missing data or data that conflicted with other responses and that could not be imputed through logical imputation were sent to data retrieval. (See Item Response and Imputation chapter for rules for logical imputation of key items.) A total of 123 items are listed as key.

Question number	Number of items
2	5
3	1
9a	1
10	1
14	12
15	3
16 (columns 2-4)	43
19	7
21 (columns 1-3 for all, and columns 4-5 for a,b,c,d)	41
24	7
28	1
29	1

If questionnaires were sent to data retrieval, then at a minimum an attempt was made to obtain responses to all key items and to all of the questions collecting demographic data (i.e., questions 23 through 31). Coders prepared detailed instructions for the telephone interviewers on which items should be discussed with the respondents, and what the problems were (e.g., missing data, or the response conflicted with another response elsewhere on the questionnaire). Cases were not removed from data retrieval until either the respondent had been reached or the period allotted for data retrieval had ended. After data retrieval was completed, a questionnaire had to have at least 50 percent of all items and at least 75 percent of all key items completed in order to be considered valid for inclusion in the data set. Responses of "don't know" were not considered as valid responses when counting the number of items completed.

As part of the coding process, special codes were assigned to indicate the reasons for missing data. The following codes were used:

(blank) = Legitimate skip
7 = Refusal
8 = Don't know
9 = Missing

The special codes were adjusted so they did not conflict with legitimate in-range responses. For example, a refusal code of 97 or 997 was used if 7 or 97 were legitimate responses.

Interviewers assigned to data retrieval were trained by the SSOCS:2000 operations supervisor. They were provided with copies of the questionnaire and project-specific training manuals, including question by question specifications (see appendix F). Training emphasized protecting respondent confidentiality, reading the questions verbatim, keeping one's tone and comments neutral so as not to lead the respondent, and proper recording of responses. The majority of data retrieval was conducted over the telephone. However, in some cases, respondents were faxed pages of the questionnaire needing responses or clarifications. This was done in cases where the problems were extensive or where the respondent specifically requested a copy of his/her original responses.

Data Editing

Data editing (correcting interviewer, respondent, and program errors) was performed both throughout and following the data collection. These procedures included confirming that data were within the defined range of values for each item; performing logic and structural edits; reviewing cross-tabulations between data items; and reviewing frequency distributions for individual data items to ensure that skip patterns were followed appropriately. After the imputation of missing values was completed, these procedures were repeated to ensure that no errors were introduced during imputation.

Range Specifications

The ranges of most of the items were determined by the codes available for the responses, since most were close-ended. For open-ended items that required an entry by the respondent (for example the number of incidents of an offense) ranges were defined to check whether the responses were reasonable.

Range checks (see appendix G) included both soft- and hard-range edits. A "soft-range" is one that represents the reasonable expected range of values but does not include all possible values. For example, the range for total school enrollment is 25-9996. Any number less than 25 would fall outside the expected range. For key items, responses outside the soft-range were confirmed with the respondent during data retrieval phone calls. If a respondent could not be reached, or if the item was not a key item, the response was accepted "as is". "Hard ranges" are those that have a finite set of parameters for an item. For example, a respondent may have indicated 3/1/00 as the date he/she completed the questionnaire. This value is out of range because the questionnaire was not mailed to the respondent until 3/27/00. Similarly, for items within question 24 (for example, percentage of male students) responses

greater than 100 percent were not accepted. For key items, respondents were called in order to re-ask the question. If a respondent insisted that a response outside the hard range was correct, or if the respondent could not be reached, or if the item was not a key item, the out of range response was not accepted and the response was coded as missing.

Consistency Checks (Logic Edits)

Consistency or logic checks (see appendix G) examine the relationships between responses to ensure that they do not conflict with one another or that the response to one item does not make the response to another unlikely. For example, if a respondent indicated in question 21 that some students were removed with no continuing services for at least 1 year, then question 20a should have the response "available and used."

Several procedures were followed when inconsistencies were identified. In some cases, the appropriate answer was clear from the context of the other questionnaire responses. If the item was a key item, respondents were called to resolve the inconsistency. Some respondents provided corrected responses, while others insisted that their responses were correct, and others could not be reached. If the item was not a key item, respondents were not called to resolve the inconsistencies.

If the inconsistencies were not resolved by contacts with the respondent, then the specific action taken depended on the particular item. Some responses were assigned missing values, while others were recoded based on other responses in the questionnaire. The last section of this chapter, "Data Anomalies," contains a description of inconsistencies in the data that were not corrected.

Frequency and Cross-Tabulation Review

The frequencies of responses to all data items were reviewed to ensure that appropriate skip patterns were followed. Members of the data preparation team checked each item to make sure the correct number of responses were represented. If a discrepancy was discovered, the problem case was identified and reviewed to determine the appropriate response. If the respondent's information was missing, the item was coded as "not ascertained" and any key items were later imputed.

Frequency Review of Text Items

The "other, specify" open-ended text responses (questions 8e and 28) were reviewed to determine if they should be coded into one of the existing code categories. When a respondent selected an "other" response, it was reviewed by the data preparation staff and, where appropriate, coded into one of the existing response categories. For question 8e, the remaining open-ended responses (i.e., those not recoded into existing categories) were coded into four subsequently created categories: Drug Awareness Resistance Education (D.A.R.E.)/education programs, special circumstances/events, random/as needed basis, and no information. Three responses were deleted as they did not answer the question. Question 28 provided few open-ended responses that could not be recoded into existing categories. The remaining responses were therefore kept in the original "other, specify" category.

Data Anomalies

The remainder of this section lists some inconsistencies that were identified but not corrected. In these cases, because there might be disagreement about the best interpretation of the data, the responses were left unchanged so analysts could have control over what adjustments were made.

Question 11a in some cases shows a greater number of teachers than the sum of questions 26a1 and 26a2. Some respondents may have interpreted the question as referring to the number of teachers in the district, rather than the number of teachers at that school.

Some respondents reported a greater number of incidents for question 16 than were reported for question 21 for the same type of offense. It is possible for the number to be greater (e.g., if nonstudents were involved), but generally one would expect question 21 to have the greater number (because it counted each student separately, while question 16 counted the number of incidents regardless of the number of offenders).

In some cases, responses to question 21 indicated that specific disciplinary actions were taken in 1999-2000, while the responses to question 20 indicated that those same actions were not available. Most likely, the specific list of offenses in question 21 may have reminded respondents of disciplinary actions that they did not remember when responding to question 20, which was much more general. In other cases, question 20 indicates that certain disciplinary actions were available and used by the school, but question 21 indicates that they were not used. A possible explanation for this

inconsistency is that responses in question 20 mistakenly reflect more than just the 1999-2000 school year.

In question 22 some respondents reported more offenses involving drugs and weapons than they reported for the total number of offenses. It is possible that schools excluded offenses involving drugs or weapons from the "total" column since they were covered in the drugs and weapons column.

6. UNIT RESPONSE

Definition of Response Rate

A response rate is the ratio of the number of completed questionnaires to the number of cases sampled and eligible to complete the survey. This rate can be either unweighted or weighted. The unweighted rate, computed using the raw number of cases, provides a useful description of the success of the operational aspects of the study. The weighted rate, computed by summing the weights for both the numerator and the denominator, gives a better description of the impact of nonresponse on weighted estimates developed from the survey.

The survey responses were monitored through an automated receipt control system. Approximately three weeks after the initial mailout, Westat interviewers began calling nonrespondents to verify that they received the questionnaire and to prompt the individuals to respond. Additional telephone prompts were made as the data collection progressed.

Several other steps were taken to maximize the response rate. The package containing the questionnaire also included a specially designed brochure describing the purpose of the study along with a page of study endorsements (see appendix C). The mailed questionnaire was accompanied by a postage-paid return reply envelope. A toll-free 800 number was also provided so that people could call to resolve questions about the survey. Remails were sent by Federal Express or faxed in order to assure prompt receipt of the questionnaire, and to give the survey greater importance in the eyes of the potential respondents. All questionnaires that were received were reviewed for consistency and completeness; if a questionnaire had too few items completed to be counted as a response (or if it had missing or conflicting data on key items), telephone interviewers called to obtain more complete responses. All telephone interviews were conducted by interviewers who had received both general training in telephone interviewing techniques, and project-specific training for SSOCS:2000.

Refusal conversion efforts were used to obtain responses from principals who had initially refused to complete the questionnaire. Whenever a refusal occurred, the interviewer recorded the respondent's reasons for refusing to participate. Interviewers also rated the strength of the refusal as mild, firm, or hostile. Standard refusal conversion procedure was to examine the reason(s) for refusal and call back any mild or firm refusal cases and attempt to gain the respondent's cooperation. All cases that were rated by interviewers as hostile were reviewed by the Project Director who assessed the respondent's verbatim reason for nonparticipation, made the decision whether the interviewer's rating was appropriate,

and then decided whether refusal conversion should be attempted. Cases determined to be truly hostile were not released for conversion. For this study, all refusal conversion attempts were conducted by a single experienced interviewer specifically trained in refusal conversion techniques. For most of the data collection period, at least a two-week hold was placed on initial refusals before a conversion attempt was made. This period was decreased near the end of data collection to facilitate survey closeout while maximizing the response rate. A case was coded as a final refusal if a second refusal was obtained when a refusal conversion attempt was made. Altogether there were 357 initial refusals for this study (11 percent of 3,314 eligible cases). Of that total, there were 55 cases (15 percent of initial refusals) in which refusal conversion efforts were successful and completed questionnaires were received. Of the remaining 302 refusals, 11 (4 percent of remaining cases) were final refusals (i.e., cases in which the principal, when recontacted, reiterated his/her refusal to participate in the survey) and 291 (96 percent of total refusals) were cases in which the interviewer was not able to recontact the principal to attempt refusal conversion.

There were 45 cases (1 percent of 3,314 eligible cases) in which no initial contact was ever established. In these cases, the interviewer was unable to reach any individual who knew whether the questionnaire packet had been received, or could indicate what the status was.

After data retrieval was completed, questionnaires had to have at least 50 percent of all items and at least 75 percent of all key items completed in order to be considered valid for inclusion in the data set. Responses of "don't know" were not considered valid responses when counting the number of items completed.

All of the response rates were weighted to account for different probabilities of selection. The weighting gives a more accurate representation of the proportion of the population that responded than unweighted response rates. Schools that were determined to be ineligible to participate in the survey (i.e., they were not regular schools, they were ungraded, or the highest grade was kindergarten or lower) were not included in the calculation of response rates. Overall, the weighted response rate was approximately 70 percent. The final number of respondents was 2,270.

Table 6-1 shows the characteristics of the schools that were selected and also of those that responded. Some categories of schools were more likely to respond than others; for example, schools were more likely to respond if they were in rural areas or towns, had low enrollment, were combined schools, or had a low percentage of students who were in minority racial/ethnic groups. To adjust for these differences, the final weight includes an adjustment for unit nonresponse.

Table 6-1. Response status and response rate of the SSOCS sample, by school characteristics: 2000

Category	Completed surveys	Non-response	Out of scope	Refusal	Incomplete data	Total	Unweighted response rate
Total	2,270	631	52	302	111	3,366	0.68
Instructional level							
Elementary	565	171	10	74	21	841	0.68
Middle	749	223	14	103	42	1,131	0.67
Secondary	757	197	18	111	42	1,125	0.68
Combined	199	40	10	14	6	269	0.77
Type of locale							
City	603	234	20	97	49	1,003	0.61
Urban fringe	810	234	11	133	40	1,228	0.67
Town	365	76	9	31	6	487	0.76
Rural	492	87	12	41	16	648	0.77
Enrollment size							
Under 300	315	61	33	23	7	439	0.77
300-999	1,371	363	14	164	52	1,964	0.70
1,000 or more	584	207	5	115	52	963	0.61
Percentage minority							
Less than 5 percent and missing.	597	104	16	52	11	780	0.78
5 to 19 percent	624	153	8	81	19	885	0.71
20 to 49 percent	506	163	9	77	38	793	0.65
50 percent or more	543	211	19	92	43	908	0.61

NOTE: School counts in this table are based on the original sample, which was drawn from the 1997-1998 CCD frame. They do not correspond directly with numbers on the data file, which has slightly different categories and updated values from the 1998-1999 CCD. In addition, the numbers for Enrollment size in this table are from the 1997-1998 CCD, while those on the data file are from the questionnaire responses.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Summary of Report on Impact of Nonresponse on Estimates from the 2000 School Survey on Crime and Safety (SSOCS:2000)

The overall (weighted) response rate for the SSOCS:2000 was 70 percent, which is lower than the current NCES target of 85 percent for cross-sectional sample surveys (Flemming, 1992, NCES Standard I-02-92).[14] NCES requires that the representativeness of the sample be evaluated by a nonresponse bias study whenever the total nonresponse (including both unit nonresponse and item nonresponse) is lower than 70 percent (NCES Standard III-05-92), as occurs for most items in the

[14] Flemming, E. (1992). *NCES Statistical Standards* (NCES 92-021). Statistical Standards and Methodology Division, U.S. Department of Education, National Center for Education Statistics.

SSOCS:2000. This section summarizes the results of a report prepared in response to that requirement; the full text of that report is provided in appendix H.

Bias, in this context, is the expected difference between the estimate from the survey and the actual population value. Nonresponse may result in bias if the nonresponding schools differ in some systematic way from the schools that did respond. For example, if schools with high crime rates were less likely to respond, then the survey might understate the extent of school crime. Nonresponse adjustments to the weights are used to adjust for systematic differences in response rates to reduce the opportunity for bias. The purpose of the nonresponse study was to examine the adequacy of the nonresponse adjustments, and if appropriate, to modify them.

Survey nonresponse was examined by reviewing the response rates by selected school characteristics, performing a Chi-square automatic interaction detector (CHAID) analysis to identify the significant predicators of response propensity, and performing regression analyses to identify variables that are correlated with selected survey items. Generally, the characteristics that are related to nonresponse in the SSOCS:2000 are also correlated with many of the variables collected in the survey. These characteristics include instructional level, type of locale, enrollment size of school, region, pupil-to-teacher ratio, minority status, and others. This suggests that the type of nonresponse adjustments to be used to weight the SSOCS:2000 data may be effective in reducing nonresponse biases (Kalton, 1983)[15].

The analyses were used to develop a revised set of nonresponse adjustments to the weights. Comparison of weighted estimates using "initial" and "final" weights revealed virtually no significant differences. This suggests that much of the variation in response rates was captured in the original sampling strata (which were defined by instructional level, type of locale, and enrollment size of school). Inclusion of additional variables to form weighting classes (e.g., region, pupil-to-teacher ratio, minority status, and others) did not have an appreciable effect on the weighted estimates for the 22 survey variables examined. Nonetheless, the revised weights were retained based on theoretical considerations suggesting that the weighting classes derived from the CHAID analysis would be effective in attenuating nonresponse biases for a broad range of statistics.

The analysis also included a comparison of the weighted estimates with those estimates that would have been obtained if data collection were stopped when the response rate reached 50 percent. In the vast majority of cases, the differences that appeared between the two sets of estimates were not

[15] Kalton, G. (1983). *Compensating for Missing Survey Data,* Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan.

statistically significant. One cannot be sure that the same finding would occur if the comparison were between a 70 percent response rate and a higher response rate, but this finding provides some encouragement that the extent of such bias might be tolerably small.

Implications for Future Surveys

The information in this analysis can be used when planning any future SSOCS surveys. One conclusion is that the choice of stratification variables for this survey appears to have been very effective, since the stratification variables were often related to the analysis variables, and little improvement in relative bias occurred when comparing the adjusted weights with the initial weights. Thus, the sample design appears promising for later surveys as well. Second, the data also suggest that there is a reasonable prospect for improving response rates in later years, particularly if the results from this survey are used to plan the later surveys. It is encouraging that there was little opposition to the survey as such, and that nonresponse primarily was due to factors such as the schedule of the survey, difficulty in contacting the principals, and the busyness of many principals. Since one of the primary difficulties was contacting school principals during the summer, there may be substantial potential for improving response rates by modifying the schedule for the survey, moving either to earlier in the spring (allowing more time before schools close for the summer) or to the fall of the following academic year.

Some key changes to the questionnaire also may have substantial potential for improving the response rate. Questions 16 and 21 were clearly the most difficult sections of the questionnaire, and the low response rates to them were directly responsible for dropping many schools from the data file. Also, given the difficulty that people had with these questions, it is likely that these questions also increased the perception of burden and complexity regarding the questionnaire, and they may have led to some questionnaire nonresponse as well as to incompletely filling out the questionnaires. Thus, simplifying these questions by dropping some columns and rows might both allow more cases to be allowed in the data file and result in higher response rates from other schools.

7. ITEM RESPONSE AND IMPUTATION

Item nonresponse refers to missing data items in an otherwise completed questionnaire. The items may be missing because the respondent was careless, refused to provide an answer, or could not obtain the requested information. In SSOCS:2000, response rate ranges for individual items within the questionnaire, ranged from 35 percent to 100 percent (after rounding). Generally item response rates were quite high. The only questions receiving lower than a 90 percent response rate were questions 9 (with response rates of 80-81 percent, not counting question 9a, which was a key item), 17 (71-77 percent), 21 (95-99 percent for the key items, but 35-55 percent for the remaining items), 22 (58–73 percent), and 24 (89-99 percent). As noted, if over 50 percent of all items or over 25 percent of key items were not completed, the questionnaire was not included in the database.

Imputation was used to adjust for item nonresponse on items specified as key by NCES. Because more extensive follow-up was conducted when nonresponse appeared on key items, item response rates were often higher for these items than on regular questionnaire items.[16] Table 7-1 presents the range of frequencies of missing values and response rates for the key survey questions. Items with low response rates are indicated on the restricted-use file only. (Refer to appendix I for detailed response rates and appendix J for more information on item response rates for all items.)

[16]Also, sometimes the items that were not key items were more difficult to respond to. For example, part of question 21 asked for the number of "other" disciplinary actions taken besides removals, transfers to specialized schools, or out-of-school suspensions. Schools found it easier to provide the number of the most serious disciplinary actions than to provide the number of all other actions, especially for some of the less serious offenses which may have been associated with a wide range of actions.

Table 7-1. Frequency of imputation and response rate for key data items in SSOCS:2000

Questionnaire item	Total number of items	Number of missing values for items		Percent response rates for items	
		Minimum	Maximum	Minimum	Maximum
2: School has written plan for specified crises	5	0	2	99.91	100.00
3: School had programs to prevent/reduce violence	1	0	0	100.00	100.00
9A: Hours a paid security person was on duty	1	37	37	97.69	97.69
10: School trained teachers to recognize violent students	1	2	2	99.91	99.91
14: Count of deaths occurring at school by type	12	2	3	99.47	99.91
15: Number of incidents involving shooting	3	1	1	99.97	99.97
16: (columns 2-4): Number of incidents by category	43	0	32	98.59	100.00
19: Frequency of problems at school by type	7	1	5	99.78	99.96
21: (columns 1-3 for all, 4-5 for a,b,c,d): Number of students involved in offenses by type	41	3	71	99.87	96.87
24: Percentage of students fitting selected criteria	7	18	254	88.81	99.21
28: Type of school	1	4	4	99.82	99.82
29: Unexcused absentee rate	1	8	8	99.65	99.65

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

All key data items with missing values were imputed using well-known procedures. Depending on the type of data to be imputed and the extent of missing values, logical imputation, poststratum means, or "hot-deck" imputation methods were employed. For three data items, imputation was done using information from the 1998–99 CCD file. Table 7-2 presents a summary of the imputation procedures used for the key data items. The table also presents the classification variables used for forming imputation classes (cells) for each question.

Table 7-2. Summary of imputation methods used for key data items

Questionnaire item	Imputation method	Classification variables used in forming imputation classes		Comment
		Hard boundary	Soft boundary	
2: School has written plan for specified crises	Hot-deck	Instructional Level, Type of Locale	Crime level in the area (Q27), Enrollment Size, Region	Imputed as a block
3: School had programs to prevent/reduce violence		No imputation was required as there were no missing values after initial data review		
9A: Hours a paid security person was on duty	Logical	Using question 8 (Q8)		
	Mean	Instructional Level, Type of Locale, Crime level in the area (Q27), Enrollment Size		Minimum cell size restriction of 10 was used
10: School trained teachers to recognize violent students	Hot-deck	Instructional Level, Type of Locale	Crime level in the area (Q27), District Enrollment Size, Region	Imputed as a block
14: Count of deaths occurring at school by type	Logical	Using question 13 (Q13)		
	Hot-deck	Q13, Instructional Level, Type of Locale	Crime level in the area (Q27), Enrollment Size, Region	Q14 and Q15 were imputed as one block in hot-deck
15: Number of incidents involving shooting	Hot-deck	Q13, Instructional Level, Type of Locale	Crime level in the area (Q27), Enrollment Size, Region	
16: (columns 2-4): Number of incidents by category	Hot-deck	Instructional Level, Type of Locale	Crime level in the area (Q27), total number of incidents (if available)	All items within each main item were imputed as a block
	Hot-deck	Instructional Level, Type of Locale	Crime level in the area (Q27), Enrollment Size, Region	
19: Frequency of problems at school by type	Hot-deck	Instructional Level, Type of Locale	Crime level in the area (Q27), Enrollment Size, Region	All 7 items were imputed as a block

See footnote at end of table.

Table 7-2. Summary of imputation methods used for key data items (continued)

Questionnaire item	Imputation method	Classification variables used in forming imputation classes		Comment
21A-J: (columns 1-3 for all, 4-5 for a,b,c,d): Number of students involved in offenses by type	Logical	For those items which are related to Q20 and if relevant items in Q20 are not missing		Otherwise, hot-deck was used
	Hot-deck	Instructional Level, Type of Locale	Crime level in the area (Q27), Enrollment Size, Region	All items within each main item were imputed as a block
21K: Total number of students who committed offenses by disciplinary action taken.	Logical	Totals of Q21A to Q21J		
24A: Percentage of students eligible for free or reduced price lunch	CCD (1998–99)	If the information is available on CCD		Mean imputation was used only if the item was missing in CCD
	Mean	Instructional Level, Type of Locale, Free Lunch Category		
24B: Percentage of students with limited English proficiency	Hot-deck	Instructional Level, Type of Locale	Percentage of Asian or Hispanic students, Region	
24C: Percentage of students with special education	Hot-deck	Instructional Level, Type of Locale	Enrollment Size, Region	
24D: Percentage of male students	CCD (1998–99)	If the information is available on CCD		Mean imputation was used only if the item was missing in CCD
	Mean	Instructional Level, Type of Locale, Free Lunch Category		
24E-G: Percentage of students fitting other selected criteria	Mean	Instructional Level, Type of Locale, Crime level in the area (Q27), Enrollment Size		Minimum cell size restriction of 10 was used
28: Type of school	CCD (1998–99)	Using information from CCD		
29: Unexcused absentee rate	Mean	Instructional Level, Type of Locale, Crime level in the area (Q27), Region		Minimum cell size restriction of 10 was used

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Methods Used

Logical imputation, rather than data retrieval, was used for some key items in order to complete the responses. Logical imputation is the assignment of data values based on other informaiton in the data record, as indicated in the bullets below. The following are the rules for logical imputation used in these situations:

- If question 3 was marked "no" or left blank but some of question 4 was marked "yes," then the response to question 3 was changed to "yes."
- If question 10 was marked "no" or left blank but question 11 was answered with non-zero responses, then the response to question 10 was changed to "yes."
- If question 13 was marked "no" or left blank, but question 14 was answered with non-zero responses, then the response to question 13 was changed to "yes."
- If question 15 was left blank, but questions 15a and 15b were marked with zeros, then the response to question 15 was assumed to be zero.
- If the total number of incidents for an item in question 16 was marked with a zero, then the remaining responses to the right of that total (i.e., the number reported to police, the number of hate crimes, and the number that were gang-related) were assumed to be zero.
- If the completed items in columns one through three of question 21 summed to the total in 21k, then blanks in the same columns were assumed to be zero.

Besides the logical imputation done during coding and editing of data, logical imputation was also applied in situations where a missing response could be inferred with certainty (or high degree of probability) from other information in the data record. For example, question 21 of the SSOCS:2000 questionnaire asks for frequency of disciplinary actions for specific crimes but question 20 asks if those disciplinary actions were available and applied in the school. If the school did not apply those disciplinary actions (i.e., if question 20 indicated the actions were not used) then the frequency of disciplinary actions (in question 21) was logically imputed as zero.

Poststratum Mean Imputation

In the poststratum means method, a record with missing data was assigned the mean value of those cases in the same "poststratum" for which information on the item was available. The poststrata or "imputation classes" were defined on the basis of variables that were correlated with the item being imputed. Preliminary exploratory analyses (e.g., using chi-square tests of association, correlation analysis, and regression analysis) were carried out to identify the relevant classification variables. The strength of association of the variables in combination with subjective assessment was used to prioritize the

importance of the variables in forming the imputation classes. Table 7-2 lists the variables (in order of importance) used in forming the imputation classes. In the case of mean imputation, a minimum cell size restriction of 10 valid observations was applied and the imputation was initially applied to those poststrata for which the minimum cell size restriction was met. For the remaining poststrata, where the cell sizes were less than the specified minimum, the classification variables (as listed in table 7-2) were dropped sequentially in reverse order of importance until the minimum cell size requirement was met. However, the need for such collapsing to increase the cell size was not very frequent, and in most cases the final cell sizes were several times larger than the specified minimum.

A disadvantage of the poststratum mean imputation is that all missing-value cases in the same cell receive the same imputed value. This results in some "clumping" of the data at the imputed value, which is undesirable for analyses in which the distribution of values is of interest. To ensure that the mean imputation did not distort the overall distribution of the relevant data item, the frequency distributions of the observations before and after imputation were always examined. As the number of imputation cells formed were very large compared to the number of missing observations for each of the mean imputed data item, only in rare cases did more than one missing value appear in an imputation cell and were imputed by the same mean value. Consequently, in none of the cases was an alternative imputation method required because of the clumping of imputed values. For example in the case of question 9A, as 192 imputation cells were formed for 33 missing observations none of the imputation cells had more than one missing observation, and hence all imputed values were different.

"Hot-deck" Imputation

In the "hot-deck" technique, cases with missing items were assigned the corresponding value of a "similar" respondent in the same "poststratum." Similar to the poststratum means approach, preliminary exploratory analyses were carried out to identify the relevant classification variables to be used to define the poststrata. The classification variables were separated into two groups – "hard" and "soft" boundary variables, as shown in table 7-2. The hard boundary variables were considered to be so important that the imputation classes were always formed within those boundaries. The boundaries formed by the soft boundary variables were crossed, if necessary, to form the imputation class. For example in imputing question 2, instructional level and type of locale were used as hard boundary variables but question 27, enrollment size, and region were used as soft boundary variables, as shown in table 7-2. For imputing the missing question 2 value of a school, the process at the first attempt searched for schools with nonmissing question 2 values by matching all the hard and soft boundary variables. If no school was available at that stage, the last soft boundary variable, i.e., region was dropped from the

matching process. That means, if a school was found from a different region but with the same values of the other matching variables then the value of question 2 for that school was used for imputing. However, if there was no school that could be matched even after dropping region variable then the next soft boundary variable, i.e., enrollment size, was dropped from the matching process. In this way, the process was continued by dropping all the soft boundary variables one by one until a matching school was found. However, the hard boundary variables were never dropped from the matching process. The record supplying the imputed value is referred to as a "donor," and was randomly selected (usually without replacement to avoid multiple uses of a single donor) when more than one school could be matched. There was no restriction on the minimum cell size. As long as there was a donor, it was used for imputing.

In the hot-deck imputation, related items within the questionnaire were often imputed as a "block." For example, the items 16A_2, 16A_3, 16A_4 under question 16A were considered as one block, and the values of a single donor were imputed to all missing items within the block. This was done to ensure consistency among the responses to related items within a main item. In some cases, nonmissing values were used as one of the soft boundary matching values to select donors. That also helped ensure consistency between nonmissing and imputed values. For example in imputing the various parts of question 16 (i.e., questions 16A to 16L), all items within each part (i.e., questions 16A_2, 16A_3, etc.) were imputed as a single block. Moreover, where available total number of incidents for an item (e.g., question 16A_1) was used as a matching variable to impute the corresponding missing values of the other items (e.g., the number of incidents reported to police (question 16A_2), number that were hate crimes (question 16A_3), etc.). Similarly, all items within each major part of question 21 were imputed as a block. In some cases, all items within a whole question (e.g., question 19) were imputed as a block; in other cases, all items associated with two or more related questions (e.g., questions 14 and 15) were imputed as one block. Where meaningful, responses to related questions were used as matching variables for imputation; for example, responses to question 17 were used as matching values to impute some items in question 16.

Trimming and Imputation Flags

The numbers of missing values imputed using the procedures described above are documented in appendix I for each key survey item. The distributions of the values before and after imputation were examined to ensure that the imputation (a) did not distort the distribution and (b) had no undue impact on the survey estimates. In only one case, an imputed value was "trimmed" to prevent it from having undue influence on the final estimates. Trimming is the process of replacing any unusually

large imputed value by the nearest value that is not considered to be an outlier. It was applied when the difference in estimates before and after imputation was large for a particular imputation cell. The difference in estimates before and after imputation was considered to be large if it was statistically significant, i.e., greater than the margin of sampling error. In other words, an imputed value was identified as unusually large by considering the overall distribution of the data points.

Imputation flags were created for all imputed items to enable users to identify imputed values. Users can use the imputation flag to delete the imputed values, use alternative imputation procedures, or account for the impact of imputation in the analysis of the data. For example, some users might wish to analyze the data with the missing values rather than the imputed values, or some users might wish to replace the imputed value by using some alternative imputation approach.

The codes that are used for imputation flags are as follows: 1 = Hot-deck imputation, 2 = Hot-deck imputation with collapsed imputation cell (i.e., a cell formed by dropping one or more of the soft boundary variables presented in table 7-2), 3 = Logical imputation, 5 = Mean imputation. For questions 21K1, 21K2, and 21K3, which are the sums of pertinent items in the same column of the reporting grid, the imputation codes indicate the number of items among the ten items contributing to the total that were imputed. For example, a code of 3 would indicate that three of the ten items contributing to the total were imputed.

8. WEIGHTING AND VARIANCE ESTIMATIONS

Weighting Methodology

As described earlier in chapter 2, a stratified random sample design was used to select schools for the SSOCS:2000. Over 3,000 schools were selected at rates that varied by sampling stratum; i.e., the classes formed by crossing instructional level (elementary, middle, secondary, combined), type of locale (city, urban fringe, town, rural), and enrollment size class (less than 300, 300-499, 500-999, 1,000+). Since the schools were selected with unequal probabilities, sampling weights are required for analysis to inflate the survey responses to population levels. Weighting is also used to reduce the potential bias resulting from nonresponse and possible undercoverage of the sampling frame. The following sections summarize the procedures used to develop sampling weights for analysis of the SSOCS:2000 data. Also described are the procedures used to develop replicate weights for variance estimation.

Base Weights

The essential component of the sampling weight is the "base weight." The base weight is equal to the reciprocal of the probability of selecting a school for the sample, and will produce unbiased (or consistent) estimates of population totals and ratios if there is no survey nonresponse. For the stratified sample design used to select the SSOCS:2000 sample, the selection probability for the *i*th sampled school in stratum *h* is simply

$$p_{hi} = n_h/N_h \ , \qquad (8\text{-}1)$$

where N_h is the total number of schools in the population (frame) in stratum *h*; and n_h is the number of sample schools in stratum *h*.

The corresponding base weight for the *i*th sampled school in stratum *h* is defined to be

$$w_{hi} = 1/p_{hi} \ . \qquad (8\text{-}2)$$

Note that the weighted count of the sampled schools (using the base weight) equals the number of schools in the sampling frame. Moreover, the base weights are said to be unbiased because, for any set of reported sample values, $y_1, y_2, ..., y_n$, the weighted sum

$$\hat{y} = \sum_{i=1}^{n} w_i y_i , \qquad (8\text{-}3)$$

provides an unbiased estimate of the corresponding population total.

The base weights developed for the SSOCS:2000 sample are documented in table 8-1. As stated above, the base weight shown in the table is simply the ratio of the number of schools in the population in a given instructional level, type of locale, and enrollment size class (i.e., sampling stratum) to the corresponding number of sampled schools. The population and sample counts used to calculate the base weights can be obtained from tables 2-7A through 7-D and 2-12 in chapter 2. Note that the sampling rates for SSOCS:2000 depended only on instructional level, type of locale, and enrollment size class. Thus, the base weights depend only on instructional level, type of locale, and enrollment size class. (Although minority status and region were used as implicit stratifiers in the selection process, the sampling rates within a stratum did not vary by these characteristics.)

Table 8-1. Base weights for the SSOCS school sample, by instructional level, type of locale, and enrollment size class (sampling strata): 2000

Type of locale	Enrollment size of school	Instructional level			
		Ele-mentary	Middle	Secondary	Combined
City	Less than 300	77.93	23.42	24.22	31.86
	300 to 499	58.91	15.93	13.83	12.75
	500 to 999	45.07	11.00	10.09	11.56
	1,000+	31.67	8.43	6.36	9.60
Urban fringe	Less than 300	88.11	23.90	26.50	47.00
	300 to 499	62.49	15.65	15.17	16.57
	500 to 999	48.19	11.73	11.16	13.19
	1,000+	33.87	8.85	7.11	9.50
Town	Less than 300	91.63	24.88	22.10	33.50
	300 to 499	63.56	16.23	15.69	16.57
	500 to 999	48.56	12.04	11.13	13.93
	1,000+	34.00	11.86	8.34	11.80
Rural	Less than 300	107.05	27.81	24.80	28.97
	300 to 499	61.77	16.06	15.67	17.59
	500 to 999	51.47	14.30	12.06	14.41
	1,000+	40.00	8.00	8.48	8.33

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Adjustments for Nonresponse

The base weights are theoretically unbiased if there is no nonresponse. Since roughly 30 percent of the schools in the SSOCS:2000 sample did not complete the survey, adjustments were made to compensate for the missing survey data. The type of nonresponse addressed in this section is referred to as "unit nonresponse." Unit nonresponse (or "whole questionnaire nonresponse") occurs when there is no information for an eligible sampled school because of a principal's refusal to participate in the survey, interviewers' inability to contact the principal, or for other reasons. The unit nonresponse adjustment procedures used in the SSOCS:2000 are described below.

To compensate for unit nonresponse, adjustment factors were calculated within selected weighting classes, and these factors were applied to the base weights of the responding schools. The

weighting classes were determined using a statistical algorithm known as CHAID (chi-square automatic interaction detector). The CHAID algorithm was used to partition the SSOCS:2000 sample into subsets that were homogeneous in terms of response propensity. Separate CHAID analyses were applied to the 12 major instructional level and type of locale categories listed in table 8-2. The variables that were treated as the "independent" variables (i.e., potential predictors of response propensity) in the CHAID analysis were derived from the 1997–98 CCD file and included:

- Instructional level (1 = elementary, 2 = middle, 3 = secondary; 4 = combined)
- Type of locale (1 = city, 2 = urban fringe; 3 = town; 4 = rural)
- Region (1 = Northeast; 2 = Southeast; 3 = Central; 4 = West)
- Enrollment size of school (1 = less than 300; 2 = 300 to 499; 3 = 500 to 999; 4 = 1,000 or more)
- Minority status (1 = under 5 percent minority enrollment or missing in the 1997–98 CCD; 2 = 5 to 19.9 percent minority; 3 = 20 to 49.9 percent minority; 4 = 50 percent or more minority)
- Percentage of students eligible for free/reduced-price lunch (1 = missing in the 1997–98 CCD; 2 = less than 35 percent; 3 = 35 to 49 percent; 4 = 50 to 74 percent; 5 = 75 percent or more)
- Pupil-to-teacher ratio (1 = missing in the 1997–98 CCD; 2 = less than 15 pupils per teacher; 3 = 15 to 17.9 pupils per teacher; 4 = 18 to 20.9 pupils per teacher; 5 = 21 pupils per teacher or more)
- District enrollment size class (1 = less than 2,500; 2 = 2,500 to 9,999; 3 = 10,000 to 24,999; 4 = 25,000 to 99,999; 5 = 100,000 or more)
- Ratio of guidance counselors to teaching staff in district (1 = missing in the 1997–98 CCD; 2 = less than 2.5 counselors per teacher; 3 = 2.5 to 3.49 counselors per teacher; 4 = 3.5 to 3.99 counselors per teacher; 5 = 4 or more counselors per teacher)
- Ratio of graduates to drop outs in district (1 = missing or not applicable in the 1997–98 CCD; 2 = less than 12 graduates per drop out; 3 = 12 to 21.9 graduates per drop out; 4 = 22 to 44.9 graduates per drop out; 5 = 45 or more graduates per drop out)

Starting with the classification variables listed above, the CHAID algorithm identifies the variables that are the most significant predictors of response propensity and then uses this information to successively partition the sample into subsets. The formation of subsets is accomplished by splitting an

existing cell into "subcells" that are internally homogeneous with respect to response propensity. An example of the output from the CHAID analysis is shown in the figure 8-1. Each terminal branch of the tree diagram in the figure represents a "final" subset or cell within which schools have the same expected response propensity. The variables that are used to form these cells are the significant predictors of nonresponse. For example, in figure 8-1, the significant predictors are pupil-to-teacher ratio, enrollment size class, and region. For the purpose of constructing nonresponse weighting adjustment cells, the CHAID analysis is efficient and economical. Additional information about the computational methods used in the CHAID analysis is given in Magidson (1993).[17]

Table 8-2 summarizes the results of the CHAID analysis as applied to the 12 major groups of schools defined by level and type of locale. The analysis was applied separately to these 12 groups because level and type of locale were expected to define the primary subgroups for analysis. Although enrollment size class was used to define the sampling strata (see Chapter 2), it was not used to define the initial subgroups for the CHAID analysis. Instead, enrollment size class (along with minority status, region, and the other 1997–98 CCD variables listed previously) was used a predictor variable in the CHAID analysis to account for possible variation in response propensity by size of school. As can be seen in the last column of table 8-2, region and the school enrollment size were identified as significant predictors of response propensity for 8 and 6 of the 12 major groups, respectively. Minority status categories and the district level counselors-to-teacher ratio were significant predictors for 4 of the 12 groups. District enrollment size, the ratio of graduates to dropouts, and percentage of students eligible for free/reduced-price lunch appeared significant for 3 groups. The school-level pupil-to-teacher ratio also appeared significant for 2 groups.

The definitions of the weight adjustment classes determined by the CHAID analysis are summarized in table 8-3. Although a total of 51 cells are listed in the table, cells with small sample sizes were collapsed with an adjacent cell. Thus, for nonresponse adjustment purposes, 49 weighting cells were used.

[17] Magidson, J. (1993). SPSS® for Windows™ CHAID™, Release 6.0, SPSS Inc.

To illustrate the approach used to calculate the nonresponse adjustments, let w_{gi} denote the base weight for the *i*th sampled school in adjustment class *g*. Further, let

$$N_{Rg} = \sum_{i=1}^{n_{Rg}} w_{gi} \tag{8-4}$$

denote the sum of the base weights of the eligible *responding* schools in class *g*, and let

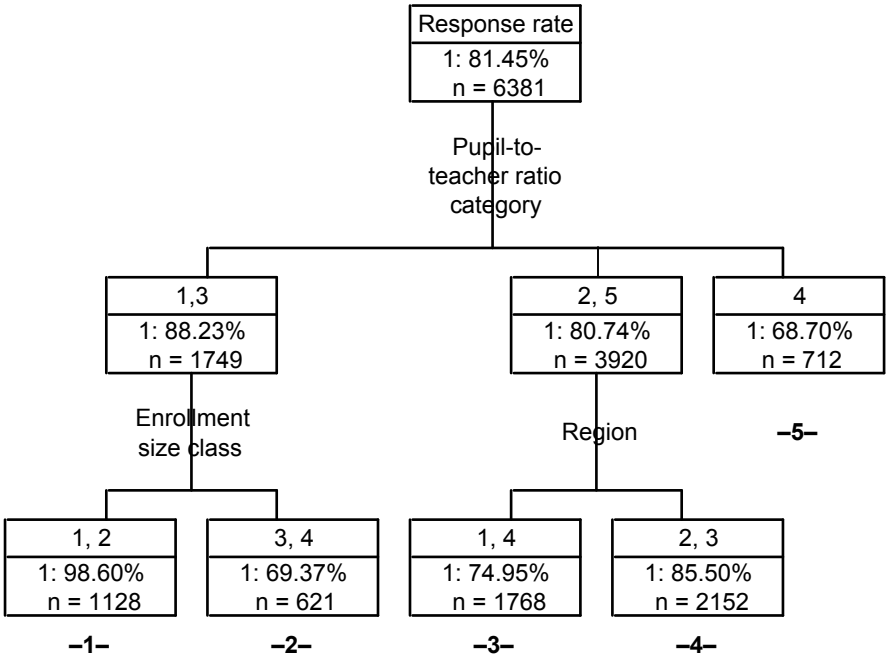
$$N_{Ng} = \sum_{i=1}^{n_{Ng}} w_{gi} \tag{8-5}$$

denote the corresponding sum of the base weights of the *nonresponding* schools in class *g*. The nonresponse-adjusted weight, $w_{gi}^{(a)}$, for the *i*th responding school in class *g* was then computed as

$$w_{gi}^{(a)} = w_{gi}\left(\frac{N_{Rg}+N_{Ng}}{N_{Rg}}\right) \tag{8-6}$$

The above formula shows that the nonresponse-adjusted weight equals the base weight times an inflation factor equal to the total weight of the eligible sampled schools divided by the total weight of the responding schools. The reciprocal of the adjustment factor is equal to the (weighted) response rate. The inflation factors used to weight the SSOCS:2000 sample, also referred to as nonresponse adjustment factors, are summarized in the last column of table 8-3. The adjustment has the effect of distributing the weight of the nonresponding schools among the responding schools in the same adjustment class *g*. The nonresponse-adjusted weights, $w_{gi}^{(a)}$, have the property that the weighted count of the responding schools using the nonresponse weights equals the corresponding weighted count of the eligible sampled schools using the base weights. Because the variables used to define the weighting classes are correlated with both response propensity and characteristics collected in the survey, the nonresponse-adjusted weights given by formula (8-6) are expected to be effective for reducing nonresponse bias.

Figure 8-1. Results of CHAID analysis for secondary/combined school in rural locales: 2000



NOTE: The percentages shown in the figure are weighted response rates. The n's are (base) weighted counts of schools in the cell. The text given below a box describes the variable used to subdivide the cell. For example, "pupil-to-teacher ratio category" refers to the five-level variable defined at the beginning of this section. All of the other variables used in the CHAID analysis are also defined at the beginning of this section. The five terminal cells denoted by the symbols 1, 2, ..., 5 are those determined by the CHAID analysis to be internally homogeneous with respect to response propensity. For example, CHAID cell 1 includes schools in pupil-to-teacher ratio categories 1 or 3, and enrollment size class 1 or 2. On the other hand, CHAID cell 3 includes schools in pupil-to-teacher ratio categories 2 or 5, and regions 1 or 4, and so on.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 8-2. Variables identified in CHAID analysis to be significant predictors of response propensity within 12 broad design strata, defined by instructional level and type of locale: 2000

Instructional level	Type of locale	Number of schools included in analysis	Variables identified in CHAID as significant predictors of response propensity
Elementary	City	290	Region; ratio of graduates to dropouts; free lunch category; pupil-to-teacher ratio
	Urban fringe	303	Ratio of counselors to teaching staff; region; minority status
	Town	95	School enrollment size class
	Rural	143	Ratio of counselors to teaching staff; region
Middle	City	339	Ratio of graduates to dropouts; minority status; school enrollment size class; free lunch category
	Urban fringe	447	School enrollment size class; district enrollment size class; region
	Town	177	Ratio of counselors to teaching staff
	Rural	154	Region
Secondary/ Combined	City	354	District enrollment size class; free lunch category; minority status
	Urban fringe	467	Region; school enrollment size class; district enrollment size class; minority status; ratio of counselors to teaching staff; ratio of graduates to dropouts
	Town	206	Region; school enrollment size class
	Rural	339	Pupil-to-teacher ratio; school enrollment size class; region

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 8-3. Definition of nonresponse adjustment classes and corresponding adjustment factors: 2000

Level/ type of locale	Final CHAID cell	Categorical variables used to define adjustment classes									Adjustment factor
		Level	Region	School size class	Minority status	Free lunch category	Pupil-teacher ratio category	District size category	Counselors-teacher ratio	Graduates-dropout ratio	
1. Elem/City	1	1	1, 4	All	All	All	1, 3, 5	All	All	1, 3, 5	1.51
	2	1	1, 4	All	All	All	2, 4	All	All	1, 3, 5	2.44
	3	1	1, 4	All	All	All	All	All	All	2, 4	1.44
	4	1	2, 3	All	All	1, 2	All	All	All	All	1.22
	5	1	2, 3	All	All	3, 4, 5	All	All	All	All	1.57
2. Elem/Urban fringe	1[1]	1	1, 2	All	All	All	All	All	1, 2	All	1.80
	2[1]	1	3, 4	All	All	All	All	All	1, 2	All	1.80
	3	1	All	All	1	All	All	All	3, 4, 5	All	1.19
	4	1	All	All	2, 4	All	All	All	3, 4, 5	All	1.35
	5	1	All	All	3	All	All	All	3, 4, 5	All	1.63
3. Elem/Town	1	1	All	1, 3, 4	All	All	All	All	All	All	1.23
	2	1	All	2	All	All	All	All	All	All	1.66
4. Elem/Rural	1	1	1, 2, 4	All	All	All	All	All	1, 2, 5	All	1.27
	2	1	3	All	All	All	All	All	1, 2, 5	All	1.06
	3	1	All	All	All	All	All	All	3, 4	All	1.77
5. Middle/City	1[2]	2	All	All	1, 2, 3	All	All	All	All	1	2.05
	2[2]	2	All	All	4	All	All	All	All	1	2.05
	3	2	All	1, 2, 3	All	1-3, 5	All	All	All	2, 3, 5	1.62
	4	2	All	1, 2, 3	All	4	All	All	All	2, 3, 5	1.13
	5	2	All	4	All	All	All	All	All	2, 3, 5	2.13
	6	2	All	All	All	All	All	All	All	4	1.27
6. Middle/Urban fringe	1	2	All	1, 2	All	All	All	All	All	All	1.29
	2	2	1	3	All	All	All	1, 2	All	All	1.63
	3	2	2, 3, 4	3	All	All	All	1, 2	All	All	1.32
	4	2	All	3	All	All	All	3, 4, 5	All	All	1.79
	5	2	All	4	All	All	All	All	All	All	1.96
7. Middle/Town	1	2	All	All	All	All	All	All	1, 3	All	1.42
	2	2	All	All	All	All	All	All	2, 4, 5	All	1.19
8. Middle/Rural	1	2	1, 4	All	All	All	All	All	All	All	1.44
	2	2	2, 3	All	All	All	All	All	All	All	1.17

See footnote at end of table.

Table 8-3. Definition of nonresponse adjustment classes and corresponding adjustment factors: 2000 (continued)

	Categorical variables used to define adjustment classes										
Level/ type of locale	Final CHAID cell	Level	Adjust-ment factor	School size class	Minor-ity status	Free lunch category	Pupil-teacher ratio category	District size category	Couns-elors-teacher ratio	Grad-uates-dropout ratio	Adjust-ment factor
9. Sec-comb/ City	1	3, 4	All	All	All	1, 3	All	1–4	All	All	1.87
	2	3, 4	All	All	1, 2	2, 4, 5	All	1–4	All	All	1.22
	3	3, 4	All	All	3	2, 4, 5	All	1–4	All	All	1.44
	4	3, 4	All	All	4	2, 4, 5	All	1–4	All	All	1.70
	5	3, 4	All	All	All	All	All	5	All	All	2.62
10. Sec-comb/ Urban fringe	1	3, 4	1	All	All	All	All	1	All	All	1.27
	2	3, 4	1	All	All	All	All	2–5	All	1, 2	1.50
	3	3, 4	1	All	All	All	All	2–5	All	3, 4, 5	2.33
	4	3, 4	2, 3	1, 2, 3	All	All	All	All	All	All	1.21
	5	3, 4	2, 3	4	1, 2	All	All	All	All	All	1.28
	6	3, 4	2, 3	4	3, 4	All	All	All	All	All	1.56
	7	3, 4	4	All	All	All	All	All	1–3	All	1.74
	8	3, 4	4	All	All	All	All	All	4, 5	All	1.37
11. Sec-comb/ Town	1	3, 4	1, 4	1, 3	All	All	All	All	All	All	1.51
	2	3, 4	2, 3	2, 4	All	All	All	All	All	All	1.13
	3	3, 4	2, 3	All	All	All	All	All	All	All	1.37
12. Sec-comb/ Rural	1	3, 4	All	1, 2	All	All	1, 3	All	All	All	1.01
	2	3, 4	All	3, 4	All	All	1, 3	All	All	All	1.45
	3	3, 4	1, 4	All	All	All	2, 5	All	All	All	1.34
	4	3, 4	2, 3	All	All	All	2, 5	All	All	All	1.17
	5	3, 4	All	All	All	All	4	All	All	All	1.44

[1]Due to small sample sizes, cells were collapsed for nonresponse adjustment purposes. The adjustment factor of 1.80 is for the collapsed group.

[2]Due to small sample sizes, cells were collapsed for nonresponse adjustment purposes. The adjustment factor of 2.05 is for the collapsed group.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Poststratification Adjustments

To partially account for changes in the universe of public schools, the final step in the weighting process was to calibrate (i.e., poststratify) the nonresponse-adjusted weights to current population counts derived from the 1998–99 NCES CCD public school universe file. Since the SSOCS:2000 sampling frame was based on the earlier 1997–98 CCD file (see chapter 2), this adjustment had the effect of partially offsetting losses in the sample due to school closure or reorganization. The adjustments were made using a ratio raking algorithm within broad classes (poststrata) defined by level, enrollment size class, and type of locale. The ratio raking algorithm described below was used rather than direct poststratification to avoid problems associated with small cell sizes.

Tables 8-4 and 8-5 summarize the 1998–99 CCD population counts (control totals) used to poststratify the weights. Table 8-4 gives the population counts by instructional level and size class (the first weighting variable or "raking dimension") while table 8-5 gives the corresponding counts by level and type of locale (the second "raking dimension"). Note that both raking dimensions are "bivariate," i.e., are defined on the basis of two classification variables. Thus, three variables (level, size, locale) are used for raking within two dimensions. To illustrate the ratio raking procedure, let *DIM*1 and *DIM*2 denote the two "raking dimensions." Then, for each of the 12 levels defined by *DIM*1, an adjustment factor, $F^{(1)}_{DIM1}$, was computed as

$$F^{(1)}_{DIM1} = \frac{N_{DIM1}}{\sum_{i=1}^{n_1} w_i^{(a)}} \tag{8-7}$$

where N_{DIM1} is the population count for the given level of *DIM*1, $w_i^{(a)}$ is the nonresponse-adjusted weight, and the denominator of $F^{(1)}_{DIM1}$ extends over the responding schools in the given cell (level) of *DIM*1. For example, as indicated in table 8-4, *DIM*1 refers to the cross classification of instructional level by enrollment size class and thus consists of 12 levels. An intermediate *DIM*1-adjusted weight for each level was then calculated as

$$w_i^{DIM1} = F^{(1)}_{DIM1}\ w_i^{(a)} \tag{8-8}$$

The adjustment given by (8-7) and (8-8) will force the weighted sample counts to equal the corresponding population counts for each level of *DIM*1 (i.e., by level and size class), but will not guarantee that the weighted counts for each of the 12 levels of *DIM*2 (defined by instructional level and type of locale) agree with the respective *DIM*2-population counts in table 8-5. Thus, the next step was to use the *DIM*1-adjusted weights, w_i^{DIM1}, to calculate adjustment factors within each level of *DIM*2 as follows:

$$F_{DIM2}^{(1)} = \frac{N_{DIM2}}{\sum_{i=1}^{n_2} w_i^{DIM1}} \tag{8-9}$$

where N_{DIM2} is the population count for the given level of *DIM*2, and where the denominator of $F_{DIM2}^{(1)}$ extends over the responding schools in the given cell (level) of *DIM*2. An intermediate *DIM*2-adjusted weight was then calculated as

$$w_i^{DIM2} = F_{DIM2}^{(1)} \; w_i^{DIM1} \tag{8-10}$$

After implementing (8-9) and (8-10), the resulting weighted counts for the 12 levels of *DIM*2 will agree with the corresponding control totals in table 8-5. However, the weighted counts for the 12 levels of *DIM*1 may no longer agree with the corresponding control totals in table 8-4. Thus, the procedure was repeated (i.e., iterated) starting with *DIM*1 and continuing through *DIM*2 until the difference between the calculated weighted sums and the corresponding population counts was negligible for *all* levels of each raking dimension. Specifically, the raking iterations continued until the estimated totals for every level of each raking dimension were all within 1 of the corresponding control totals. Tables 8-6 and 8-7 summarize the weighted counts of the sample before and after poststratification for each of the raking dimensions.

Table 8-4. Definition of poststratification cells and population counts (control totals) for first raking dimension, level and size class (*DIM*1): 2000

Instructional level	Enrollment size class	Level of *DIM*1	Number of eligible schools in 1998–99 CCD file*
1 Elementary	1 <300	1	13,406
	2 300-499	2	17,005
	3 500-999	3	18,081
	4 1000+	4	1,404
			49,896
2 Middle	1 <300	5	3,232
	2 300-499	6	3,185
	3 500-999	7	7,017
	4 1000+	8	1,957
			15,391
3 Secondary or	1 <300	9	5,033
4 Combined	2 300-499	10	2,578
	3 500-999	11	4,055
	4 1000+	12	5,018
			16,684
Total			81,971

*Counts exclude schools in the outlying U.S. territories, nonregular schools such as special education, vocational, alternative/other schools, ungraded schools, and schools with a high grade of kindergarten or lower. See table 2-1 in chapter 2 for definition of instructional levels used in this table.

SOURCE: Special tabulations from the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1998–99 data file for the U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 8-5. Definition of poststratification cells and population counts (control totals) for second raking dimension, level and type of locale (*DIM*2): 2000

Instructional level	Type of Locale[1]	Level of *DIM*2	Number of eligible schools in 1998–99 CCD file[2]
1 Elementary	1. City	1	13,263
	2. Urban fringe	2	16,679
	3. Town	3	5,570
	4. Rural	4	14,384
			49,896
2 Middle	1. City	5	3,413
	2. Urban fringe	6	5,402
	3. Town	7	2,467
	4. Rural	8	4,109
			15,391
3 Secondary or 4 Combined	1. City	9	2,693
	2. Urban fringe	10	4,286
	3. Town	11	2,399
	4. Rural	12	7,306
			16,684
Total			81,971

[1]Type of locale categories reflect the new coding system introduced in the 1998–99 CCD file. "City" includes schools in central cities of a CMSA or MSA, "urban fringe" includes schools in an incorporated place, Census-designated place, or non-place territory within a CMSA or MSA, "town" includes schools in an incorporated place or Census designated place outside a CMSA or MSA, and "rural" includes all remaining schools.

[2]Counts exclude schools in the outlying U.S. territories, nonregular schools such as special education, vocational, alternative/other schools, ungraded schools, and schools with a high grade of kindergarten or lower. See table 2-1 in chapter 2 for the definition of instructional levels used in this table.

SOURCE: Special tabulations from the U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1998–99 data file for the U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 8-6. Weighted counts of sample before and after poststratification, by first raking dimension, level and size class (*DIM1*): 2000

Instructional level	Enrollment size class	Level of *DIM*1	Sample Size	Weighted count using nonresponse-adjusted weights[1]	Weighted count using poststratified weights[2] (final raked)
1 Elementary	1 <300	1	113	13,799.51	13,406.00
	2 300-499	2	194	17,608.60	17,005.00
	3 500-999	3	243	16,606.50	18,081.00
	4 1000+	4	27	1,365.44	1,404.00
			577	49,380.05	49,896.00
2 Middle	1 <300	5	96	3,078.77	3,232.00
	2 300-499	6	142	3,011.76	3,185.00
	3 500-999	7	384	6,858.06	7,017.00
	4 1000+	8	122	2,004.47	1,957.00
			744	14,953.06	15,391.00
3 Secondary or	1 <300	9	126	4,150.13	5,032.55
4 Combined	2 300-499	10	135	2,690.94	2,577.88
	3 500-999	11	247	3,879.94	4,055.01
	4 1000+	12	441	4,933.94	5,018.56
			949	15,654.95	16,684.00
Total			2,270	79,988.07	81,971.00

[1]Weighted counts differ from corresponding 1997–98 CCD counts in tables 2-2A and 2-2B for two reasons: (a) closed and other out-of-scope schools are not included in the weighted counts, and (b) a small number of schools that were not listed separately in the 1997–98 CCD frame but were added to the sample during data collection (see footnote in table 2-13) are included in the weighted counts. Weighted counts also differ slightly from those in table H2-1B of appendix H because the variables used to classify schools for poststratification are based on the 1998–99 CCD data, whereas the variables used to classify schools in table H2-1B of appendix H were based on the 1997–98 CCD data.

[2]Weights are final (fully raked) weights. Compare with tables 8-4 and 8-5.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Table 8-7. Weighted counts of sample before and after poststratification, by second raking dimension, level and type of locale (*DIM*2): 2000

Instructional level	Type of locale	Level of *DIM*2	Sample Size	Weighted count using nonresponse-adjusted weights[1]	Weighted count using poststratified weights[2] (final raked)
1 Elementary	1. City	1	165	13,161.73	13,263.00
	2. Urban fringe	2	205	16,314.36	16,679.00
	3. Town	3	59	5,211.80	5,570.00
	4. Rural	4	148	14,692.16	14,384.00
			577	49,380.05	49,896.00
2 Middle	1. City	5	185	3,462.42	3,413.00
	2. Urban fringe	6	267	5,114.35	5,402.00
	3. Town	7	131	2,513.12	2,467.00
	4. Rural	8	161	3,863.17	4,109.00
			744	14,953.06	15,391.00
3 Secondary or 4 Combined	1. City	9	177	2,290.13	2,693.00
	2. Urban fringe	10	323	4,234.09	4,286.00
	3. Town	11	138	2,254.42	2,399.00
	4. Rural	12	311	6,876.31	7,306.00
			949	15,654.95	16,684.00
Total			2,270	79,988.07	81,971.00

[1]Weighted counts differ from corresponding 1997–98 CCD counts in tables 2-2A and 2-2B for two reasons: (a) closed and other out-of-scope schools are not included in the weighted counts, and (b) a small number of schools that were not listed separately in the 1997–98 CCD frame but were added to the sample during data collection (see footnote in table 2-13) are included in the weighted counts. Weighted counts also differ slightly from those in table H2-1B of appendix H because the variables used to classify schools for poststratification are based on the 1998–99 CCD data, whereas the variables used to classify schools in table H2-1B of appendix H were based on the 1997–98 CCD data.

[2]Weights are final (fully raked) weights. Compare with tables 8-4 and 8-5.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

Methods for Computing Sampling Errors

In surveys with complex sample designs such as the SSOCS:2000, estimates of variance (e.g., standard errors) that are based on simple random sampling assumptions are generally inappropriate. Note that the terms "estimates of variance," "standard errors," and "sampling errors" are all used to refer to the variability (and, hence, precision) of a sample-based estimate. As discussed previously, the SSOCS:2000 sample design employed extensive stratification, and the weighting procedures included both nonresponse and poststratification adjustments. To accommodate these features of the SSOCS:2000 sample design, either of the methods described below (Replication or Taylor Series) can be employed.

Replication Sampling Errors

One method of computing sampling errors to reflect various aspects of the sample design and estimation procedures is the replication method. Under replication methods, a specified number of subsamples of the full sample (called "replicates") are created. The survey estimates can then be computed for each of the replicates by creating replicate weights that mimic the actual sample design and estimation procedures used in the full sample. The variability of the estimates computed from the replicate weights is then used to estimate the sampling errors of the estimates from the full sample. An important advantage of the replication methods is that they preclude the need to specify cumbersome variance formulas that are typically needed for complex sample designs (McCarthy, 1966).[18] Another advantage is that they can readily be adapted to reflect the variance resulting from nonresponse (and other weight) adjustment procedures.

The two most prevalent replication methods are balanced repeated replication (BRR) and jackknife replication. The two methods differ in the manner in which the replicates are constructed. For the SSOCS:2000, a variant of jackknife replication was used to develop replicate weights for variance estimation because the jackknife method is believed to perform somewhat better than BRR for estimates of moderately rare events (e.g., number of schools in which a serious crime was committed). Under the jackknife method, the replicates are formed by deleting specified subsets of units from the full sample. The jackknife method provides a relatively simple way of creating the replicates for variance estimation and has been used extensively in NCES surveys (e.g., it has been used in the National Household

[18] McCarthy, P. (1966). *Replication: An Approach to the Analysis of Data from Complex Surveys*. Vital and Health Statistics, Series 2, No. 14. Washington, DC: U.S. Department of Health, Education and Welfare.

Education Survey (NHES) and numerous surveys conducted under the Fast Response Survey System (FRSS)).

For the SSOCS:2000, 50 jackknife replicates were defined as follows:

1. The 3,362 schools selected for the sample were sorted in the same order as that used in sample selection. Thus, the schools were sorted initially by sampling stratum (instructional level, type of locale, and enrollment size class), and then by minority status and region within each sampling stratum (see Selection of the Sample in chapter 2).

2. Next, 50 "variance estimation units" were formed by assigning the first school on the sorted list and every 50th school thereafter to "variance unit 1," the second school on the list and every 50th school thereafter to "variance unit 2," the third school on the list and every 50th school thereafter to "variance unit 3", and so on up to "variance unit 50." Note that each variance unit contains about 1/50th of the schools in the full sample, and together the 50 variance units are mutually exclusive and exhaustive.

3. In a few cases, a sampled school-record actually consisted of more than one school (see footnote in table 2-13). For variance estimation purposes, the added schools were assigned to the variance unit to which the original school-record was assigned.

4. Fifty jackknife replicates were then defined by deleting each variance unit in turn from the full sample. In other words, "jackknife replicate 1" consisted of the sampled schools that remained after deleting variance unit 1. "Jackknife replicate 2" consisted of the sampled schools that remained after deleting variance unit 2. "Jackknife replicate 3" consisted of the sampled schools that remained after deleting variance unit 3, and so on up to "jackknife replicate 50."

5. Corresponding to each jackknife replicate, a replicate base weight was calculated and assigned to each school. For example, corresponding to jackknife replicate 1, the replicate 1 base weight for sampled school i was 0 if the school was deleted from replicate 1, and set to $(50/49)w_i$ otherwise, where w_i is the full-sample base weight for school i. Similarly, the replicate 2 base weight for sampled school i was 0 if the school was deleted from replicate 2, and set to $(50/49)w_i$ otherwise. The assignment of replicate base weights continued in this fashion until each school had a series of 50 replicate base weights. The method used to create the replicate base weights is referred to as "JK1" (see Westat, 2000).[19]

6. Using the procedures described in the previous sections, the nonresponse and poststratification adjustments developed for the full sample were applied separately to each of the 50 sets of replicate base weights. In other words, the entire weighting process was redone 50 times. Note that for this purpose, 1997–98 CCD data were used to develop the nonresponse adjustments for the full sample and for each of the 50 replicates, whereas 1998–99 CCD data were used to develop the poststratified (raked) weights for the full sample and for each of the 50 replicates (see previous discussion

[19] Westat. (2000). *WESVAR 4.0 User's Guid*e. Rockville, MD: Westat.

under Adjustments for Nonresponse and Poststratification Adjustments). At the end of this process, each school in the final data set had 51 weights, one full-sample weight and 50 replicate weights. These weights are defined in the analytic files as FWT (for the full sample) and FWT1 through FWT50 (for the fifty replicate weights), respectively.

7. There was no difference in the methods used for the full-sample and replicate weights with respect to the convergence criterion in the raking process. For the full sample and all 50 replicates, the raking iterations were stopped when the respective weights converged to within 1 of the corresponding control totals.

To illustrate how sampling errors are calculated under the jackknife replication approach, let $\hat{r}$ denote a weighted survey estimate (e.g., the number of schools reporting a particular type of crime/incident, or the ratio of the number of occurrences of an incident to enrollment). Further, let $\hat{r}_j$ be the corresponding estimate for a given jackknife replicate j. The estimated variance of $\hat{r}$ can be computed from the formula

$$var(\hat{r}) = F \sum_{j=1}^{K} (\hat{r}_j - \hat{r})^2 \qquad (8\text{-}11)$$

where the summation extends over all $K = 50$ jackknife replicates, and $F = (K-1)/K = 49/50$. Note that the variance given by formula (8-11) provides a measure of the replicate-to-replicate variability of the sample-based estimate, $\hat{r}$. The standard error of the estimate is simply the square root of $var(\hat{r})$, which in turn can be used to construct confidence limits around the "true" value being estimated.

The computation of the sampling errors using the replicate weights described previously can be done easily using the Windows-based software package WesVar Software; the replication method should be specified as JK1. The current version of WesVar is available from Westat (wesvar@westat.com). A previous (unsupported) version of WesVarPC (version 2.12) is also available free of charge.

Taylor Series Approximation

Another approach to the valid estimation of sampling errors for complex sample designs is to use a Taylor series approximation to compute sampling errors. To produce standard errors using a Taylor series program, such as SUDAAN (www.rti.org),[20] two variables are required to identify the stratum and the primary sampling unit (PSU). For the SSOCS:2000, the stratum-level variable is the indicator of the sampling stratum from which the school was selected. The PSU indicates the first-stage sampling within the stratum (i.e., the sampled school).

The required PSU and stratum variables appear on each school record in the analysis file as WESID and STR_SOCS, respectively. These variables can be used in SUDAAN to produce standard errors by specifying that the design is a "with replacement" sample (DESIGN = WR) and that the sampling levels are given by the appropriate stratum and PSU variables (i.e., use STR_SOCS WESID in the NEST statement).

Stata (www.stata.com),[21] another software package that uses Taylor series methods, also uses the PSU and stratum variables to define the units needed for standard error computation. To specify the stratum, PSU and weight variables in Stata use the svyset strata, svyset psu, and svyset pweight commands. For example, use the following commands to specify these design parameters:

- svyset strata str_socs
- svyset psu wesid
- svyset pweight fwt

Data users should be aware that the use of different approaches or software packages in the calculation of standard errors may result in slightly different standard errors. Estimates of standard errors computed using the replication method and the Taylor series method are nearly always very similar, but not identical.

---

[20] Shah, B., Barnwell, B., Bieler, G. (1995). *SUDAAN User's Manual, Software for Analysis of Correlated Data*, Release 6.40. Research Triangle Park, NC: Research Triangle Institute.

[21] Stata Corporation. (2001). *Stata User's Guide*, Release 7, College Station, TX: Stata Corp.

Calculation of Confidence Intervals

As mentioned previously, the standard error of an estimate is defined to be the square root of the variance of the estimate. The standard error can be used to construct numerical limits within which the value being estimated can be expected to fall with a high degree of confidence. For example, a confidence interval for a proportion (e.g., the proportion of schools that report the occurrence of a particular type of crime) can be computed as

$$\hat{p} \pm z\, se(\hat{p}) \qquad (8\text{-}12)$$

where $\hat{p}$ is the sample-based estimated proportion, $se(\hat{p})$ is the standard error of $\hat{p}$, and z is an appropriate percentage point from a standard normal distribution. The value of z to use in formula (8-12) depends on the desired level of confidence; e.g., for 95 percent confidence limits, $z = 1.96$; for 99 percent confidence limits, $z = 2.58$.

To illustrate the use of formula (8-12), suppose that $\hat{p} = 0.30$ and $se(\hat{p}) = 0.023$. Then, 95 percent confidence limits around the true proportion are given by

$$lower\ limit = 0.30 - 1.96(0.023) = 0.255,$$

$$upper\ limit = 0.30 + 1.96(0.023) = 0.345.$$

Thus, while the "point" estimate of $\hat{p} = 0.30$ appears to indicate that 30 percent of the schools have the specified characteristic, the lower and upper confidence limits put the estimate in proper perspective: With 95 percent confidence, the true percentage having the specified characteristic is likely to be anywhere between 25.5 percent to 34.5 percent, a difference of 9 percentage points. Formula (8-12), which is referred to as the "normal approximation," is appropriate as long as $\hat{p}$ is not close to 0 or 1.

Approximate Sampling Errors

Although calculating the sampling errors using the methods described above is recommended for many applications, simple approximations of the sampling errors may be valuable for some purposes. A discussion of one such approximation follows.

Most statistical software packages (e.g., SPSS, SAS) compute standard errors of the estimates based upon simple random sampling assumptions. The standard error from this type of statistical software can be adjusted for the complexity of the sample design to approximate the standard error of the estimate under the actual sample design used in the survey. For example, the variance of an estimated proportion in a simple random sample (SRS) is the estimated proportion (p) times its complement ($1-p$) divided by the sample size (n). The standard error is the square root of this quantity. That is, under simple random sampling, the standard error of the estimated proportion p is: $SE_{SRS}(p) = \sqrt{\frac{p(1-p)}{n}}$. This standard error can be adjusted to more closely approximate the standard error for the estimates from the SSOCS:2000 using the method described below.

A simple approximation of the impact of the sample design on the standard errors of the estimates that has proved useful in many surveys is to adjust the simple random sample standard error estimate by the root design effect (DEFT). Since estimated DEFTs are highly unstable (e.g., see discussion in Skinner, Holt, and Smith, 1989, page 46), an average DEFT computed over many related statistics is usually used to approximate the standard error for an estimate.[22] The DEFT is the ratio of the standard error of the estimate computed using the replication method discussed above to the standard error of the estimate under the assumptions of simple random sampling. For example, for an estimated proportion, p, $DEFT = \sqrt{\frac{var(p)}{p(1-p)/n}}$, where $var(p)$ is the variance of the estimated proportion using the replication or Taylor series methods described earlier, and n is the sample size. An average DEFT is obtained by computing the DEFT for a number of estimates and then averaging the DEFTs. A standard error for an estimate can then be approximated by multiplying the simple random sample standard error estimate by the average DEFT. For example, an approximate standard error of p that reflects the complex sample design features is given by:

$$SE(p) = DEFT\sqrt{\frac{p(1-p)}{n}},$$

where *DEFT* is the average *DEFT* computed earlier. Note that other sources (e.g., see Kish, 1965) use the term DEFF to refer to the design effect (i.e., the ration of the *variance* of an estimate to the corresponding variance that would have been obtained with a simple random sample of the same size). The DEFT defined above is simply the square root of the DEFF.

[22] Skinner, C., Holt, D., and Smith, T. (eds). (1989). *Analysis of Complex Surveys.* New York: J. Wiley and Sons.

In complex sample designs like the SSOCS:2000, the DEFT is often greater than 1.0 due to the differential weights attached to the sampled schools. In the SSOCS:2000, differential weights result from the disproportionate allocation of sample to strata and use of nonresponse and poststratification weighting adjustments. Both of these factors can result in an average DEFT greater than 1.

As stated above, the average DEFT can be used to approximate the standard error for an estimate. Note that the appropriate value of DEFT depends on the particular domain being analyzed (e.g., the DEFT for elementary schools is different from that of secondary schools). An example of how to approximate the standard error for a percent estimate is as follows. If a weighted estimate of 46 percent is obtained for some characteristic (e.g., suppose that 46% control access to school grounds during school hours), then an approximate standard error can be developed in a few steps. First, obtain the simple random sample standard error of the estimate as:

$$\sqrt{\frac{\hat{p}(100-\hat{p}}{n}}$$

where $\hat{p}$ is the weighted estimate (percentage) and n is the unweighted sample size on which the percentage is based. Since the full SSOCS:2000 sample is being used for this estimate n = 2,270. Then, the corresponding simple random sample standard error is $\sqrt{46(54)/2{,}270} = 1.05$. In this example, the approximate standard error of the estimate is 1.05 times DEFT, where DEFT is the appropriate *root* design effect. If 1.4 is chosen as a conservative estimate of the DEFT, the estimated standard error would be 1.47 (i.e., 1.4 times 1.05).

The approximate standard error for a mean can be developed using a related procedure. First, the mean is estimated using the full sample weight and a standard statistical package like SAS or SPSS. Second, the simple random sample standard error is obtained through a similar, but unweighted analysis. Third, the standard error from the unweighted analysis is multiplied by the mean DEFT. For example, suppose that the estimated (weighted) mean number of hours per week that 1 paid law enforcement person was on duty at school was 10, and the simple random sampling error (unweighted) was 8 hours. Then the approximate standard error for the estimate would be 11.2 hours (8 hours x 1.4).

Users who wish to adjust the standard errors for estimates of parameters in regression models should follow a procedure similar to that discussed for means, above. Specifically, the estimates of the parameter in the model can be estimated using a weighted analysis in a standard statistical software package such as SAS or SPSS. A similar analysis using the same statistical model, but unweighted, will

provide the simple random sample standard errors for these parameter estimates. The standard errors can then be multiplied by the DEFT to arrive at the adjusted standard error.

Alternatively, the final weight can be adjusted to reflect the DEFT before the parameter estimates are calculated in a statistical software package such as SAS or SPSS. To do this, first sum the values of the final weights for the cases being examined (usually this would be the total sample size of 2,270, but one might look at fewer cases because of missing data, or because one is only interested in schools with particular characteristics). For example, for an analysis of total incidents of vandalism, sum the final weights, W_i for all 2,270 cases on file. Next, divide this sum by the number of cases to generate an average final weight. That is, the average final weight is given by: $\bar{w} = \frac{\sum_{i=1}^{n} w_i}{n}$, where n is the sample size. Multiply the average final weight by the square of the DEFT for the population of interest to obtain the adjusted average weight, i.e., $\bar{w}_{adj} = DEFT^2 \bar{w}$. Divide the final weight by the adjusted average weight and save the quotient as a new final weight, $w_i^{new} = \frac{w_i}{\bar{w}_{adj}}$. Weight the regression analysis using this new final weight. The standard errors generated in the analysis will approximate the standard errors correctly adjusted for design effects.

It should be noted that direct computation of the standard errors is always recommended when the statistical significance of statements would be affected by small differences in the estimated standard errors. Although root design effects are sometimes used to approximate the standard errors from complex survey samples they are not appropriate for estimates of extremely rare events such as the occurrence of murder/suicide. This is due to the fact that the standard errors for such estimates are not meaningful even if they are computed using jackknife replication. For example, if no murders are reported in the sample, the jackknifed standard error would be 0, but this does not necessarily mean that the estimate is subject to no sampling error.

Table 8-8 shows how, for the SSOCS:2000, the average design effect varies for several school classification categories. In general, the average DEFT ranges between 1.0 and 1.4.

Table 8-8. Approximate sampling errors, selected average design effects: 2000

School characteristic	Average design effect
Totals	1.4
Instruction level	
Elementary	1.0
Middle	1.0
Secondary	1.1
Combined	1.1
Enrollment size	
Less than 300	1.1
300 – 499	1.3
500 – 999	1.4
1,000 or more	1.3
Type of locale	
City	1.4
Urban fringe	1.4
Town	1.4
Rural	1.3
Percentage minority	
Less than 5 percent/missing	1.3
5 to 19 percent	1.3
20 to 49 percent	1.4
50 percent or more	1.3

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety, 2000.

9. REINTERVIEW STUDY

Reinterviews with a subset of respondents from the original SSOCS:2000 survey were conducted in order to assess the reliability of the data provided in response to that survey as well as to provide insights into the questionnaire design. The primary objectives of the SSOCS:2000 reinterview study were to:

- Identify survey items that were not reliable (i.e., the two interviews did not elicit the same response);
- Quantify the magnitude of the response variability for items collected from the same respondent at two different times; and
- Provide feedback to improve the survey design for future surveys (especially since SSOCS is planned as a recurring survey).

The reinterview questionnaire was designed both to repeat questionnaire items taken from the original SSOCS:2000 questionnaire (so that the two sets of responses could be compared), and to ask some additional questions about the data that were provided. In order to limit the cost of the study and to limit respondent burden, only selected items from the SSOCS:2000 questionnaire were examined. This report discusses the results from that reinterview study conducted in coordination with the SSOCS:2000 survey.

Design and Procedures

A random sample of 185 schools was preselected from the original SSOCS:2000 sample to participate in the reinterview. The sample size of 185 was chosen in order to achieve a target of 150 respondents, assuming that both the SSOCS:2000 survey and the reinterview study achieved 90 percent response rates.[23]

Because the goal of SSOCS:2000 was to collect data for the original questionnaire, and because comparisons between the original and reinterview questionnaires were only possible if the original questionnaire was completed, the reinterview schools were only contacted for the reinterview after completing the original SSOCS:2000 questionnaire. Further, the schools must also have completed all data retrieval (i.e., telephone contacts to obtain answers for items that had been left blank, and/or to

[23] Information on the sample sizes needed to detect change is provided in table 2-3 elsewhere in this report. Cost was also a consideration in the selection of the sample size.

resolve apparently inconsistent answers given in different parts of the questionnaire), so that data collection would not be compromised by the reinterview study. Following the completion of data retrieval, a waiting period of 2 weeks was provided (so that respondents would have some time to forget their original responses) before the reinterview questionnaire was mailed.[24]

The items that were repeated on the reinterview questionnaire were chosen from among those items that had already been designated as key items, providing a mixture of both quantitative (i.e., numerical) and categorical responses. More specifically, the questions that were repeated on the reinterview survey were:

- Whether the school had a formal program to prevent or reduce violence (question 3, categorical) and, if so, whether each of eight components were included in that program (question 4, categorical);
- At what times the school regularly used paid law enforcement or security services personnel at school (question 8, categorical) and, if so, the average total number of hours that at least one such person was on duty (question 9a, quantitative), wore a uniform or other identifiable clothing (question 9b, quantitative), and/or carried a firearm (question 9c, quantitative);
- The extent to which various factors limited schools' efforts to prevent or reduce crime (question 12, categorical, with 5 items taken from the original list of 14 items);
- The number of incidents occurring at the school during the 1999-2000 school year (question 16, quantitative, with 3 rows taken from 15 in the original questionnaire);
- Following up on the question above (question 16), a set of four questions asked what was the primary source for the data that the respondent used, what the respondent did in order to provide separate counts for physical attacks or fights with weapons and without weapons, what the respondent did in order to limit the responses to thefts of $10 or more, and whether various alternatives would be easier to report than counting the number of incidents;
- The degree to which various disciplinary problems occurred at the school (question 19, categorical, using four of seven items in the original questionnaire);
- The disciplinary actions that were taken in response to physical attacks or fights by students (question 21g, quantitative, using 1 of 11 rows in the original questionnaire);

[24] If an original questionnaire was completed before the end of the school year, then sending the reinterview questionnaire was delayed until after the school year was completed, in addition to observing the 2 week waiting period. This was done so that the reinterview survey could measure any error associated with providing data for only a partial year. In fact, analysis of the data suggests that some respondents probably gave partial year data even after the school year was completed (as indicated by their later explaining a discrepancy by saying the later response included incidents that had not yet happened when the survey was first completed); this might happen if the school had not yet updated the records, or the school based its responses on summaries from the district and those were not yet updated to reflect the full year. In order to prevent the mode of the survey from being a source of discrepancies between the original and reinterview responses, and because some of the items might require respondents to refer to school records, the reinterview questionnaires were distributed by mail (similar to the original questionnaires).

- Following up on the question above (question 21), a set of two questions asked what was the primary source for the data that the respondent used, and whether the category "removal with no continuing school services for at least 1 year" was different from the school's definition of expulsion (and if so, what the school's definition of expulsion was);

- The disciplinary outcomes for special education students who committed an offense that normally would result in a suspension or expulsion of more than 10 school days for children without disabilities (question 22, quantitative);

- Following up on the question above (question 22), a question asked what was the primary source for the data that the respondent used; and

- The crime level of the area in which the school's students lived (question 27, categorical).

The complete reinterview questionnaire is provided in appendix K so that the precise question wording and structure can be reviewed.

If schools gave quantitative responses that were conflicting across the two surveys, follow-up telephone calls were made to determine the reason for the discrepancy. In the telephone calls, respondents were asked the reason for the discrepancy, which was recorded verbatim. In addition, respondents were provided with a set of yes/no items that gave possible reasons for the discrepancies (e.g., some incidents had occurred since the original survey was completed, the respondent provided estimates for one survey but checked school records for another, or different people were involved in completing each survey). Respondents also were asked which response should be considered the most accurate. Follow-up was limited to the quantitative responses because they were expected to have the greatest discrepancies (based on the potential burden or unavailability of obtaining the data from school records).

The reinterview questionnaire included some questions about how the data had been obtained that were not on the original questionnaire. For example, one concern was that respondents might collect data at a school in a way that conflicted with the definitions and distinctions used in the SSOCS:2000 questionnaire, possibly affecting the accuracy of the responses. For this reason, respondents were asked how they were able to provide separate estimates for the number of fights with weapons and the number without weapons (e.g., their records already made that distinction, or the number of incidents was small enough that they could make the distinction) (see question 16B in appendix K). Also, the reinterview questionnaire asked about the source of some of the quantitative data that respondents supplied, in order to better assess its reliability (e.g., electronic data file, or estimate) (see

question 16A in appendix K). Data that were based on school records were expected to be more reliable than those based on estimates, which are subject to problems of recall and potential bias. The data source also potentially affected survey burden since it may have been time consuming to manually search through records depending on the number of records and how they were organized.

Response Rates

Of the 185 schools selected for the reinterview sample, 143 (77 percent) responded to SSOCS:2000 and thus were eligible for the reinterview.[25] Of these, 114 respondents (80 percent) completed the reinterview questionnaire. Two of these were later rejected for having too much missing data on the original questionnaire, and were dropped from both surveys. Thus there were a total of 112 respondents available for reinterview analysis, comprising 79 percent of the 141 who were eligible for the reinterview survey.

In any survey, the existence of nonresponse creates the potential for bias in the survey results depending on the degree to which the survey respondents differ systematically from the nonrespondents and those differences are systematically related to the survey responses. For example, small schools were more likely to reply to the original survey than large schools; thus if small schools also tended to collect their data in a different way, then the results on survey reliability might differ from what would be obtained if all schools had responded. However, the survey weight included a correction for nonresponse to the original SSOCS:2000 survey, and an analysis of the nonresponse indicated that the adjusted weights appeared to be effective in eliminating those biases that could be identified. Thus, nonresponse to the original survey is probably not a critical issue.

To test the degree to which the 112 SSOCS:2000 reinterview respondents may be considered to represent the full sample, chi-square tests were performed with each of the four variables used for sampling (instructional level, type of locale, enrollment size, and minority status).[26] None of the tests revealed statistically significant biases in the sample. Nevertheless, when the reinterview data were reweighted to represent the full survey population, a nonresponse adjustment was also included to further lessen the likelihood that nonresponse bias might affect the results (as well as to allow national totals if

[25] Though some data on the reinterview questionnaire would be informative even if the original survey had not been completed (e.g., concerning the ways in which schools store the data), the primary focus of this analysis is on comparing principals' responses to the original questionnaire with their responses to the reinterview questionnaire. Nonrespondents to the original survey therefore were excluded since no data would be available for such comparisons.

[26] This analysis was based on weighted tables in which the 112 reinterview respondents were compared with the remainder of the 2,270 respondents to the original survey.

desired). The nonresponse adjustments were based on the locale of the school (chosen because, of the four sampling variables, it appeared to show some of the largest differences with regard to response rates) and the seriousness of the incidents reported by the schools (chosen to represent differences in crime levels among the schools).

Sampling and Nonsampling Errors

Every survey can have errors of two forms: sampling error (in which the sample differs in some way from the population it represents, because it is based only on a small sample of all schools) and nonsampling error (in which errors are introduced for other reasons). Sampling errors are generally well understood, and can be examined quantitatively by the use of standard errors: tests of statistical significance may then be performed to describe the likelihood that a particular result may have occurred by chance because a particular sample was used.

Nonsampling errors, by contrast, often are understood to be a greater source of error than sampling error, but are difficult to measure and quantify (because of the multiple potential sources of error and the lack of appropriate data). These sources of error include coverage problems in the sample, nonresponse, and measurement errors, and are described below. However, the SSOCS:2000 reinterview study was designed to examine only one – measurement error. Coverage error occurs when the data used to select the sample fail to correspond fully with the intended population: for example, to the extent that the Common Core of Data fails to include some schools (e.g., because they were created after the CCD data were collected, or because they are charter schools, which sometimes are listed as districts rather than as schools in CCD), a sample drawn from CCD may fail to reflect the full population of schools. Because the basic source of coverage error is a lack of appropriate data, it is difficult to quantify the degree to which the population is not covered unless an alternate source of data is available. The reinterview survey was not designed to measure coverage error.[27]

Nonresponse errors include both unit nonresponse (e.g., when a school fails to return a questionnaire) and item nonresponse (when some items on the questionnaire are not completed). These

[27] Even though coverage error occurs during the process of sampling, it is not sampling error. Sampling error has a very specific meaning, and refers to errors that may occur because an unrepresentative random sample could be selected by chance even if one has a complete list of all members of the population. For example, it is possible (though highly unlikely) that a purely random sample would result in selecting only schools with high levels of poverty. (However, the use of stratification helps to lessen the likelihood of unrepresentative samples, to the degree that important variables are either included among the stratification variables or correlated with them.) Because of the way that sampling error is defined, it is possible to set up mathematical models to describe the likelihood of a sample being unrepresentative, using standard errors and tests of statistical significance. Nonsampling error, by contrast, cannot be examined simply through a mathematical model, but is based on missing or incorrect data.

are relatively easy to quantify in terms of the amount of nonresponse (though in some surveys it is difficult to discriminate between nonresponse and ineligibility for the sample, such as when there is no answer to a telephone survey), but the impact of the nonresponse error depends on the degree to which the nonrespondents differed in some systematic ways from the respondents. Nonresponse errors also were not examined in the reinterview study. Unit nonresponse was examined in a separate report, and tables describing the item nonresponse are provided in both the user's manual and this detailed data documentation.

Measurements of Data Reliability

Measurement errors occur for a variety of reasons, such as the failure of a respondent to understand a question, an unwillingness to provide data which might be harmful or embarrassing, or providing inaccurate responses (e.g., through faulty recall, or the use of approximations rather than seeking to obtain and provide a precise response). Measurement errors can be especially difficult to quantify because an incorrect response may appear quite reasonable (especially on variables that show a wide amount of variation), and there may be no clear indication that a response is inaccurate.

This report provides a way of measuring the reliability of the data, even though it does not cover all sources of error. Reliability refers to the degree to which the same answer is given consistently under the same conditions; it is a necessary characteristic in order for data to be considered trustworthy, but not the only characteristic. Data could still be incorrect even though respondents give the same answer time after time. For example, there might be a problem in the sample's coverage (e.g., if new schools were systematically left out because of the lack of data on such schools), but coverage error is a separate issue from reliability and the responses when comparing the original and reinterview surveys would remain consistent across both surveys. The reinterview study would not provide any data on such errors. Similarly, respondents might consistently make the same errors on both the original questionnaire and the reinterview questionnaire (e.g., if they misinterpreted the question in the same way both times): these errors would not be ascertained by the reinterview study because the two sets of responses are consistent. Also, if respondents tried to lessen their workload by copying their responses from the original questionnaire onto the reinterview questionnaire, the responses would be consistent, but the reinterview study would provide no measure of errors in the original responses. (In order to minimize this possibility, the cover letter that accompanied the reinterview questionnaire explained the basis for the study, and asked respondents not to refer to their original questionnaires.)

The SSOCS:2000 reinterview study differed from many reinterview studies by not simply repeating questions from the original questionnaire, but also including additional questions about the data. An advantage of this approach is that the study can provide more information about the quality of the data than would be provided only by comparing the original and reinterview responses, and it broadens the scope of the study beyond just the topic of consistency. For example, if a respondent indicated that the data on one survey were based on estimates while on the other survey they were based on school records, such a response might help to explain an inconsistency between the original and reinterview responses. In fact, the discrepancy interviews were used to verify whether this was a reason for the discrepancy. Even if no inconsistency appeared between the two surveys, knowing whether the responses were based on estimates or on school records would still provide some information about the accuracy of the data. One would generally expect that responses based on school records would be more accurate, because of the potential difficulty of recalling all incidents (especially if the number of incidents is large) plus a possible tendency to seek to provide approximations rather than precise numbers. Thus, knowing the percentage of schools that made estimates would provide information about the degree to which school reports might be subject to problems with recall. Aside from providing general information about survey reliability, this information could be valuable in determining the best timing for data collection in future surveys. The greater the number of schools providing the estimates, the more important it is that the survey be conducted relatively soon after the end of the school year, so that problems with recall will be minimized.

The classic component of a reinterview study remains the comparison of responses in the original questionnaire to those in the reinterview questionnaire. When a discrepancy appears, there are four possible explanations:

- Both responses may have been correct at the time they were given, but changes in circumstances may be responsible for the discrepancy. For example, the original responses of some schools may not reflect all of the crimes that occurred during the entire year, or may not include important changes in school programs that occurred after the survey was completed.

- The original data may have been incorrect. For example, a respondent may have given a rough estimate without looking up the actual data, a key data source may not have been available at the time the questionnaire was completed, or a respondent may have misinterpreted a question.

- The reinterview data may have been incorrect for similar reasons as above.

- Both answers may be incorrect. For example, approximations may have been used to provide the questionnaire responses.

Analysis Methods

Several statistics have been developed to examine various aspects of the reliability of reporting using original and reinterview responses. Two statistics in particular have been widely used by the National Center for Education Statistics in other reinterview studies, and are used here to measure the response variability (i.e., the degree to which the reinterview responses differed from the original responses): the gross difference rate and the index of inconsistency.[28] The gross difference rate is the average squared differences between the responses; for a binary variable this is equivalent to the percentage of cases with different responses in the two interviews. It is an absolute measure of the impact of response error on the estimates. The index of inconsistency is the ratio of the gross difference rate to the total variance of the estimate. Thus, it is a relative measure of this impact, used to measure the proportion of the total variability that arises due to random response error.

Of these two measures, in general the gross difference rate is the more appropriate measure of reliability for the SSOCS:2000 data because the index of inconsistency is much less stable: the index of inconsistency can vary greatly depending on the variance of the estimate (as will be discussed later), and can be high even when the number of inconsistencies is quite low. For this reason, the primary focus of this analysis is on the gross difference rate. However, the index of inconsistency also is provided for two reasons. First, it provides readers with an additional perspective on the data. Second, in the particular case of quantitative variables the categorization scheme that is often used to evaluate the response variability is not applicable, and the index of inconsistency has the possible advantage of imposing a relative scale on all items, which may facilitate comparisons across items. The two measures are discussed in greater detail in the sections that follow.

These statistics are typically computed based on the number of sample cases reported as having a particular characteristic in the original survey and in the reinterview. This approach is valid for simple random sampling or when the goal of the analysis is to evaluate and quantify response variability of the population of survey respondents. When the goal is to provide estimates of response variability of the national estimates, it is more appropriate to estimate these statistics using weights that adjust for the probability of selection. Since this was a main objective of the SSOCS:2000 reinterview, weighted data were used. More specifically, the weight developed for the original SSOCS:2000 survey was adjusted to weight the reinterview sample to the full sample and adjust for nonresponse.

[28] For example, see Salvucci, S., et. al. (1997). *Measurement Error Studies at the National Center for Education Statistics* (NCES 97-464). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

If a response was missing for a case in either the original survey or reinterview, then that case was excluded from the calculation.[29] If more than two responses were possible for a categorical variable, then that variable was collapsed to a binary variable in order to facilitate the calculation of the gross difference rate. More specifically, for question 12 the categories *limit in a major way* and *limit in a minor way* were combined to create a single category (e.g., *limit in some way*). For questions 19a and 19b the categories *happens daily* and *happens at least once a week* were combined into the category *happens at least once a week*, and the remaining categories were combined into the category *happens less than once a week.* For questions 19f and 19g all categories except for *never* were combined into the category *happens at all*, while *never* was retained as the alternate category. For question 27 the categories *high level of crime* and *moderate level of crime* were combined into the category *moderate or high level of crime*, and the remaining categories were combined into the category *low or mixed level of crime*. (Original and reinterview questionnaires can be found in appendices C and K.)

Table 9-1 shows the general format of the possible reporting outcomes from the original and reinterviews when the item has only two possible values. From tables formatted in this fashion, it is possible to estimate several characteristics relevant to the consistency of the reporting between the original survey and the reinterview. For example, the off-diagonal cells (b and c) estimate the proportion of responses that were reported differently in the original interview and the reinterview. Since most of the statistics computed in this report are based on weighted data, the values in the cells are actually weighted sums of the number of cases rather than the raw number of cases. The definitions of the statistics used in this report are given below, where the cell counts are the estimated totals.

Table 9-1. General format of interview-reinterview results

	Original interview		
Reinterview	Number of cases with characteristic	Number of cases without characteristic	Total
Number of cases with a characteristic	a	b	a+b
Number of cases without a characteristic	c	d	c+d
Total	a+c	b+d	n=a+b+c+d

[29] Imputed values were not used in this report because the purpose was to compare respondents' answers across the two surveys; the inclusion of imputed values could create false matches or discrepancies. Since the item response rate typically was very high for both the interview and reinterview, the exclusion of the missing values was not a significant problem. The questions that had response rates lower than 90 percent (counting a nonresponse to either survey) were: question 9B and question 9C (86 percent); question 21G4 (65 percent), question 21G5 (44 percent); and question 22 (ranging from 71 to 78 percent).

Gross Difference Rate

The gross difference rate is an estimate of the reliability or consistency of reporting. When applied to binary variables it often is expressed in percentage form as the weighted percentage of cases with different responses reported in the original interview and the reinterview. This interpretation does not apply to quantitative variables, however. The gross difference rate is the weighted ratio of the gross difference (i.e., the sum of the squared differences) divided by the estimated total number of cases. In its raw form, the gross difference rate for both categorical and quantitative variables can be written algebraically as

$$G = \sum w_i(x_{1i} - x_{2i})^2 / \sum w_I \tag{9-1}$$

where x_{1i} is the response to the original interview for case *i*, x_{2i} is the response to the reinterview for case *i*, and w_i is the original interview weight for case *i*, described above. Thus, the gross difference rate is the average squared difference between the responses.

As noted, for all binary variables the gross difference rate can be expressed in percentage form as the percentage of cases with different responses in the two interviews (i.e., those falling in the off-diagonal cells in table 9-1). The percentage form often is used because of its ease of interpretation, and because the use of percentages imposes a common scale that makes it easier to compare response variability across different questionnaire items. The percentage form can be calculated as

$$G\% = 100\frac{(b+c)}{n} \tag{9-2}$$

where b is the number of cases in which, for example, the original response was "Yes" and the reinterview response was "No," c is the number of cases where the original response was "No" and the reinterview response was "Yes," and n is the weighted number of cases. This is a special case of the summation formula above, in which the x_i terms only take on the values of 0 or 1.

To aid in the presentation of the gross difference rates, the following general rules may be used to estimate the impact of measurement error on the estimates. They are used to categorize the response variability of binary variables as measured by the gross difference rate:

- A gross difference rate in which less than 10 percent of the responses show disagreement across the two surveys is low response variability;

- A gross difference rate in which between 10 and 20 percent of the responses show disagreement across the two surveys is moderate response variability; and

- A gross difference rate in which above 20 percent of the responses show disagreement across the two surveys is high response variability.[30]

Since the gross difference rate is an absolute measure of the measurement error, this rule does not account for the fact that 5 percent disagreement is a much more serious response problem for a 1 percent statistic than a 50 percent statistic. Therefore, to account for the relative size of the estimate, the rule above best applies to estimates between 20 and 80 percent. Outside of this range, the gross difference rate and other measures of data quality should be considered with respect to the size of the estimate. Most of the categorical variables in this analysis had estimates that fell within 20 percent and 80 percent; where they did not, that fact is noted in the tables through the use of an asterisk. For the quantitative variables, these rules do not apply.

For example, the original survey estimate for question 3 (the percentage of schools with formal programs to prevent or reduce violence) was 73 percent, with a gross difference rate of 20.8 percent (see table 9-2). The estimate of 73 percent falls within the desired range of 20 to 80 percent, so that the categorization scheme for evaluating response variability is reasonable to apply. The gross difference rate exceeds 20 percent (indicating that more than 20 percent of the respondents gave different responses on the two surveys), and therefore is classified as high in response variability.

Index of Inconsistency

A second statistic used here is the index of inconsistency. It is the ratio of the gross difference rate to the total variance of the statistic. The general formula for both categorical and quantitative variables is:

$$I = G / (s_1^2 + s_2^2) \qquad (9\text{-}3)$$

where G is the gross difference rate defined above, s_1^2 is the sample variance for the original interview, and s_2^2 is the sample variance for the reinterview. Because SSOCS:2000 used a complex sample design,

[30] This rule is used in Brick, J.M., Rizzo, L., and Wernimont, J. (1997). *Reinterview Results for the School Safety and Discipline and School Readiness Components* (NCES 97-339). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

the standard formulas and statistical routines for computing the sample variance do not apply, and WesVarPC was used to perform the calculations.[31]

For binary data, the index can be expressed as a percentage as:

$$I\% = 100(b + c) / 2np(1 - p) \qquad (9\text{-}4)$$

where b + c is the weighted number of cases where the two responses disagree, n is the weighted number of cases, and p = (a + c)/n is the weighted percentage of cases in the original interview that gave the designated response (e.g., that responded with a "yes") (see table 9-1). The gross difference rate divided by 2 is an unbiased estimate of the simple response variance if the observations from the two interviews are independent and identically distributed.[32] Thus, the index of inconsistency is a ratio of the simple response variance and the total variance of the estimate.

The fact that the index of inconsistency is calculated by dividing the gross difference rate by the variance of the statistic helps to partially impose a scale on the measure of response variability, though it still is possible for the measure to take on values greater than 100. This can happen because the response variance may actually reduce the overall variability in the estimate. Hansen, Hurwitz, and Pritzker (1964) showed precisely this phenomenon for a binary random variable.[33] In fact, values greater than 100 did appear for some SSOCS:2000 variables, as is shown later in this analysis.

It is possible, and even likely, that the responses to the reinterview may be affected in some ways by the original interview experience. This conditioning of respondents means that the assumption of independent and identically distributed responses to the interviews may not be fully satisfied. Nevertheless, the index is a valuable measure of the relationship between response error (i.e., the numerator in equation 9-3) and sampling error (the denominator in equation 9-3), and it provides some basis for evaluating the level of response variability for those quantitative variables where the gross difference rate cannot be expressed as a percentage.

The index of inconsistency is a relative measure since the gross difference rate (an absolute measure) is divided by a term that depends on the variance of the estimate. Thus, in a sense the variance

---

[31] Additional information about the sampling, weighting, and standard errors is provided elsewhere in this detailed data documentation.

[32] Forsman, G., and Schreiner, I. (1991). "The Design and Analysis of Reinterview: An Overview," in *Measurement Error in Surveys,* eds. P. Biener, et. al. New York: John Wiley & Sons, 279-302.

[33] Hansen, M.H., Hurwitz, W.N., and Pritzker, L. (1964). "The Estimation and Interpretation of Gross Differences and Simple Response Variance." in *Contributions to Statistics*, ed. C.R. Rao. Research Triangle Park, NC: Research Triangle Institute, 111-136.

of the estimate provides a scale for evaluating whether the gross difference rate is high or low. Note that as the estimated percentage in the category (p) becomes extreme (close to 0 or 1), then the denominator of the index becomes very small (equation 9-4). As a result, even a small gross difference rate can result in a very large index. For example, assuming a constant value for the gross difference rate, the value of I differs by a factor of about 25 when the value of p varies from 1 percent to 50 percent. Thus, the index is most useful for binary variables for estimates between 20 and 80 percent because in this range the total variance is relatively constant, varying only between 16 and 25 percent.

The rules used here for interpreting the index for binary variables are that:

- An index of less than 20 percent is low relative response variability;
- An index between 20 and 50 percent is moderate relative response variability; and
- An index above 50 percent is high relative response variability.[34]

For example, returning to the example discussed above (question 3, appendix K — the presence of formal programs to prevent or reduce violence), the original survey estimate (73 percent) falls within the appropriate range for using the index of inconsistency (20 to 80 percent). As shown in table 9-2, the estimated value of the index is 54.5, which, like the gross difference rate for the same measure, indicates high relative variance (i.e., because it is above 50 percent). In this particular case, the gross difference rate and the index of inconsistency both lead to the same conclusion, but it is not necessary for them to do so in general.

No rules have been established for evaluating the index of inconsistency for quantitative variables.

Though there is a logic to comparing the simple response variance to the total variance of the estimate (i.e., it is sensible that if the total variance of the estimate is low, then one would also want a low gross difference rate), one may question whether this is the best way to create a scale. For example, if two variables both have the same gross difference rate, one might question whether one is less reliable than the other just because the total variance of the estimate (i.e., the denominator used to calculate the index of inconsistency; equation 9-3) is smaller. In a comparative sense, one might say that it is doing less well than might be expected based on the total variance of the estimate, but the actual number of

[34] This rule is used in S. Salvucci, et.al., "Measurement Error." Some other studies have defined moderate response variability using the range of 20 to 45 percent. The ranges are somewhat arbitrary and reflect the practices of other researchers, but there is no clearly established standard.

discrepancies between the original and reinterview responses might be exactly the same. Other weaknesses of the index of inconsistency are its potential for exceeding 100, the greater difficulty of interpreting the statistic (as compared with the gross difference rate, which can be directly related to the percentage of discrepancies appearing between the two surveys), and the greater instability of the statistic. Thus the gross difference rate appears to be a better statistic, though the index of inconsistency is also used here to provide an alternative measure and to compensate for the fact that the gross difference rate is less easy to evaluate for quantitative data than for categorical data.

Response Variability for the Categorical Questions

This section applies the statistics discussed above with regard to the categorical questions that were examined in the reinterview study. The categorical questions were a mixture of objective and subjective questions: question 3 and question 4 were yes/no questions about the nature of formal programs at school to prevent or reduce violence, question 8 asked at what times the school regularly used paid law enforcement or security services at school, question 12 asked about factors limiting the school's efforts, question 19 asked for perceptions of the frequency of different kinds of disciplinary problems, and question 27 asked for the crime level of the area in which the students lived.[35] Because these questions were relatively simple and did not require referring to records to provide an answer, they were expected to show relatively little error.

Table 9-2 shows, for each question, the sample size (i.e., the unweighted number of cases available for analysis after excluding missing data), the actual population estimate based on the original survey (e.g., the weighted percentage of schools indicating they had formal programs intended to prevent or reduce violence), the gross difference rate, and the index of inconsistency. For example, all 112 schools responding to the reinterview survey answered question 3 (appendix K). The population estimate was that 73 percent of all regular public schools have a formal program to prevent or reduce violence. The gross difference rate (using the percentage format) was 20.8 percent, meaning that 20.8 percent of the respondents gave answers that differed when comparing the original and reinterview responses. This is interpreted as high response variability. The index of inconsistency was 54.5 percent, and shows the relationship between the gross difference rate and the total variance of the estimate (i.e., the total variance of the population estimate of 73 percent and the comparable estimate for the reinterview study). Because the questions in table 9-2 were all categorical and were reported as percentages, the gross difference rate is the most appropriate statistic for examining their reliability (as discussed earlier).

[35] See appendix K for the actual questions.

Table 9-2. Estimates of gross difference rates and indexes of inconsistency for categorical questions in SSOCS:2000

Question	Description	Un-weighted sample size	Original survey estimate (percent-age)	Gross difference rate		Index of inconsistency	
				Estimate	Level	Estimate	Level
Q3	Formal program prevent/reduce violence	112	73	20.8	High	54.5	High
Q4a	Prevention training (e.g.,social skills)	110	65	24.1	High	58.7	High
Q4b	Behavioral modification for students	112	66	27.3	High	61.6	High
Q4c	Student counseling/social work	111	66	22.9	High	54.5	High
Q4d	Individual mentoring/tutoring students	111	63	22.9	High	52.4	High
Q4e	Recreation/enrichment student activities	110	53	23.5	High	47.5	High
Q4f	Student involvement resolving problems	112	45	17.4	Moderate	35.2	Moderate
Q4g	Promote sense of community/integration	111	57	32.2	High	68.1	High
Q4h	Hotline/tipline to report problems	111	22	14.7	Moderate	38.7	Moderate
Q8a	Security used during school hours	112	31	8.0	Low	18.4	Low
Q8b	Security while students arrive/leave	110	26	6.0	Low	15.9	Low
Q8c	Security at selected school activities	110	42	24.7	High	49.8	High
Q8d	Security when school not occurring	110	15	16.8	Moderate*	69.0	High*
Q8e	Other times security used	111	5	7.7	Low*	72.7	High*
Q12a	Efforts lmtd by lack of tchr training	111	50	27.7	High	57.6	High
Q12b	Efforts lmtd by lack of altrntive plcmnt	111	67	26.5	High	55.5	High
Q12e	Efforts lmtd by lack of parent support	111	42	37.0	High	79.3	High
Q12l	Efforts lmted by fed policies/disabled	110	60	28.1	High	56.7	High
Q12m	Efforts limited by other fed. policies	110	39	24.0	High	51.3	High
Q19a	How often student racial tensions	112	3	3.1	Low*	63.1	High*
Q19b	How often student bullying occurs	112	29	16.4	Moderate	42.3	Moderate
Q19f	How often undesirable gang activities	112	19	12.6	Moderate*	39.3	Moderate*
Q19g	How often undesirable cult activities	112	7	5.0	Low*	29.0	Moderate*
Q27	Crime where students live	111	24	16.1	Moderate	45.4	Moderate

*Survey estimate is outside of the range from 20 to 80 percent, making this categorization scheme less meaningful.

NOTE: Gross difference rates less than 10 are considered low, between 10 and 20 are moderate, and greater than 20 are high. Indexes of inconsistency less than 20 are considered low, between 20 and 45 are moderate, and greater than 45 are high. The indexes of inconsistency depend on the total variance of the estimate, and can be high when there is very little variation in the responses. All estimates in the table are weighted. The survey estimate column is based on the full SSOCS:2000 sample, not the reinterview sample. The complete question wording is shown in appendix K.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Table 9-2 shows that the categorical questions often were subject to high error rates. Based on the gross difference rate, 13 of the 24 categorical variables on the reinterview survey had high response variability (i.e., above 20 percent).[36] The index of inconsistency largely results in the same classification of response variability as the gross difference rate, except that for four variables it shows higher response variability. All four of these variables had survey estimates below 20 percent, suggesting

[36] Although some of the 13 variables had survey estimates that were outside of the range of 20 to 80 percent, the gross difference rate still may be classified as high. If one were to adjust the boundaries stated in the rules for classifying the response variability, one would adjust it by lowering the threshold below 20 percent rather than raising it. That is, if a gross difference rate is considered high relative to the actual estimate when the gross difference rate is above 20 percent, then it is even higher relative to the actual estimate when the estimate is quite small.

that both categorization schemes are less meaningful. For these four variables, the classification scheme for the gross difference rate probably understates the response variability, based on a comparison of the error rates to the survey estimates. On the other hand, the index of inconsistency probably overstates the response variability, because of its sensitivity to the variance of the estimate.

As seen in table 9-2, the gross difference rate is often high for questions 3 and 4 (seven of the nine variables had gross difference rates above 20 percent), which discuss formal programs to reduce or prevent violence. It is also high for question 12 (all five variables had gross difference rates above 20 percent), which asks for each principal's evaluations of the factors limiting his/her school's efforts to prevent or reduce crime. The response variability also is high (24.7) for question 8C (which asks about the use of security services at selected school activities).

Response Variability for the Quantitative Questions

One focus of the reinterview study was on the accuracy of the quantitative variables which measured the frequency of incidents and disciplinary actions at school. These quantitative questions were questions 9 (amount of time law enforcement or security services personnel were used regularly at the school), 16 (the number of various types of incidents at the school), 21 (the number of various disciplinary actions taken), and 22 (various outcomes for offenses by special education students). It was assumed that these questions would likely show relatively high levels of error, based on the possibility that respondents would find it difficult to obtain the requested quantitative data. For example, respondents might choose to make estimates rather than referring to records to obtain such data, the data might not be kept at all, the data might not be kept in a way that was consistent with the definitions used in the questionnaire, or additional incidents might have occurred since the original survey was completed. In order to better assess the accuracy of the quantitative data and the reasons for problems that appeared, the reinterview survey included additional questions about how the quantitative data were obtained. In addition, telephone calls were made to check on the reasons for discrepancies between the original responses and the reinterview responses.

Table 9-3 presents three measures of the response variability, rather than the two used in table 9-2. The gross difference rate has a slightly different interpretation than previously, and an additional measure (the percentage of responses that did not match) is added to display

Table 9-3. Estimates of gross difference rates and indexes of inconsistency for quantitative questions in SSOCS:2000

Question	Description	Un-weighted sample size	Percentage that did not match	Original survey estimate	Gross difference rate	Index of inconsistency
Q9a	Total hours security on duty per week	70	64.1	623,044	85.3	15.0
Q9b	Total hours security wore uniform	50	62.7	489,464	153.8	25.5
Q9c	Total hours security carried a firearm	51	61.1	354,133	151.4	25.0
Q16c1_1	# of attacks/with weapon - total	111	7.1	11,982	0.4	74.2
Q16c1_2	# of attack/weapon/reported police	111	7.1	5,339	0.4	75.5
Q16c1_3	# of attack with weapon/hate crimes	111	0.3	155	0.0	100.3
Q16c1_4	# of attack with weapon/gang-related	110	0.4	615	0.0	100.4
Q16c2_1	# of attacks/no weapon – total	111	60.8	806,784	218.9	25.8
Q16c2_2	# of attacks/no weapon/reported	111	27.4	137,637	7.2	12.5
Q16c2_3	# of attacks/no weapon/hate crimes	110	3.4	7,603	36.0	101.9
Q16c2_4	# of attacks/no weapon/gang-related	110	5.3	11,923	1.0	36.0
Q16f1	# of theft/larceny - total	108	51.0	217,875	11.4	11.2
Q16f2	# of incidents theft/larceny/reported	108	27.8	105,475	4.9	5.9
Q16f3	# of incident theft/larceny/hate crime	108	0.0	355	0.0	*
Q16f4	# of incidents theft/larceny/gang	108	0.3	1,155	0.0	100.3
Q21g1	# of removals for attacks/fights	111	18.3	29,927	87.2	101.5
Q21g2	# of transfers for attacks/fights	111	14.3	19,640	3.2	28.7
Q21g3	# of suspensions for attacks/fights	111	35.7	282,887	57.0	50.0
Q21g4	# of other actions for attacks/fights	69	56.8	427,974	344.3	39.9
Q21g5	# of no actions for attacks/fights	44	2.5	6,495	0.2	102.2
Q22a1_1	Placement changed after hearing/total	86	24.2	42,120	4.1	51.8
Q22a1_2	Placement chngd/hearing, drugs/weapons	71	8.0	7,458	0.1	61.6
Q22a2_1	Placement changed after injunction/total	81	6.0	3,078	0.2	103.3
Q22a2_2	Placement changed/injnction, drugs/wpns	78	0.4	484	0.0	100.4
Q22a3_1	Placement chnge w/o hearing, total	83	12.3	18,718	3.0	72.2
Q22a3_2	Placement chnge w/o hearing, drgs/wpns	75	1.4	1,908	0.0	70.2
Q22b1_1	No change,hearing/session not held,total	80	6.3	24,985	1.2	101.8
Q22b1_2	No change,hearing not held,drugs/wpns	73	0.6	1,784	0.0	100.6
Q22b2_1	Hearing did not approve change, total	80	4.8	8,775	0.5	96.7
Q22b2_2	Hearing did not approve chnge, drgs wpns	75	0.0	1,166	0.0	*
Q22b3_1	Court did not approve change, total	78	0.0	468	0.0	*
Q22b3_2	Court did not approve chnge, drgs/wpns	75	0.0	303	0.0	*

*The index of inconsistency cannot be computed for these variables because of the lack of variation in the reinterview sample (all responses were zero).

NOTE: Because these statistics are for quantitative variables, there is no uniform standard for evaluating the rates as low, moderate, or high. The indexes of inconsistency depends on the total variance of the estimate, and can be high when there is very little variation in the responses. All estimates in the table are weighted. The survey estimate column is based on the full SSOCS:2000 sample, not the reinterview sample. The complete question wording is shown in appendix K.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

information that previously was incorporated in the gross difference rate. For quantitative data (e.g., the number of hours or incidents), the gross difference rate cannot be interpreted as a percentage, and the standards used earlier for judging low, moderate, or high gross difference rates do not apply. Instead, the gross difference rate must be evaluated by comparing it with the original survey statistic (though no

guidelines have been established for classifying the response variability as high or low). Further, for binary variables (as in table 9-2) the gross difference rate is equivalent to the percentage of responses that did not match, but for quantitative variables these two measures are not equivalent. Thus, both the gross difference rate and the percentage of responses that did not match are presented here, along with the index of inconsistency. The gross difference rate is a better measure of response variability than the percentage of responses that did not match because it measures the distance between the two responses (i.e., not just whether the two responses were inconsistent, but also how different they were).[37] For example, it could happen that the original responses and reinterview responses never agreed but that they were always so close that for practical purposes there was little difference on the two surveys (e.g., the percentage that did not match was 100 percent, while the gross difference rate could be quite low relative to the original estimate). Still, the percentage of responses that did not match also can provide useful information. Table 9-3 shows that many variables matched exactly between the two surveys for more than 90 percent of all respondents, another indication of the consistency of the responses.

The categorization scheme for the index of inconsistency also does not apply to quantitative variables, though the way that it is computed (i.e., by dividing the gross difference rate by the total variance of the estimate) helps to impose a partial scale on the statistic (i.e., it typically results in a value lower than 100, though not necessarily so). Some of the variables examined in table 9-3 showed extremely skewed distributions (i.e., almost all of the responses were zero). Because of the low amount of variation, the index of inconsistency should not be considered very reliable in such cases. No standard has been established for identifying those situations in which the index of inconsistency is unreliable (or indeed for interpreting the index of inconsistency when it is applied to quantitative variables). However, table 9-3 shows situations both where the index of inconsistency cannot be calculated (because of the lack of variation in the reinterview responses) and where it appears high despite the presence of few inconsistencies (e.g., for three variables the index of inconsistency is above 100, and for an additional six variables the index of inconsistency is 60 or higher while the gross difference rate is 5 or lower. As in the previous section, the primary focus here is on the gross difference rate since it is a more reliable statistic.

As an example, there were 70 cases that had non-missing data for question 9a on both the original and reinterview questionnaires. Among those 70 cases, 64 percent gave different responses on the two questionnaires. If this were a binary variable, then 64 percent would also be the gross difference rate; however, for quantitative variables, the gross difference rate is presented as an absolute number (not

[37] Because the categorical variables were all recoded to be binary (if they were not binary already), no meaning can be attached to "distance" for the previous analyses (e.g., when multiple responses are considered, then the distance between, say, "limit in a major way" and "limit in a minor way" may be different than the distance between "limit in a minor way" and "does not limit;" when the responses are recoded to be binary, then only one measure of distance exists — e.g., "limits in a major or minor way" versus "does not limit" — and all differences between responses therefore have the same distance). For quantitative variables, however, the size of the difference is relevant.

a percentage), and should be compared in size to the actual survey estimate. For this variable, the original survey estimate (i.e., the sum of all nonmissing values using the original survey) is 623,044, while the gross difference rate is 85.3. (This particular estimate happens to be somewhat artificial because one normally would look at the mean for this variable rather than the sum. However, the gross difference rate applies to sums, not to means.) Though as a percentage a gross difference rate of 85.3 would be high, when compared as an absolute number to the original survey estimate, it appears small. In essence, this means that while the original and reinterview responses often disagreed, the amount of disagreement was small when compared to the actual values that were reported.[38] The index of inconsistency (15.0) also suggests a low response variability, though the categorization scheme used in table 9-2 does not strictly apply here since question 9a is a quantitative variable.

The quantitative questions showed less response variability than expected. Often the gross difference rate was extremely low (e.g., 23 of the 32 statistics were below 10, despite the fact that the original survey estimate ranged from 155 to 137,637 for these 23 variables). Even the larger gross difference rates were low in comparison with the survey estimates (e.g., the remaining nine variables had gross difference rates that were 1 percent or lower of the survey estimates, as also was true of the 23 variables with low gross difference rates). In fact, none of the gross difference rates appear to be high; those eight variables that did have gross difference rates higher than 20 were based on estimates of 7,600 or higher, and six of the eight variables had estimates higher than 275,000.

The index of inconsistency tended to show higher response variability than the gross difference rate, but still somewhat better results than for the categorical variables. To evaluate the index of inconsistency, this analysis uses the categorization scheme used for the categorical variables as a rough guide, though this scheme is not intended to be applied to quantitative variables. Still, using the total variance of the estimate as the denominator helps to create a common scale (equation 9-3), which both provides some basis for using the categorization scheme and precludes comparing the index of inconsistency to the actual estimates (which do not have a common scale). By this measure, the quantitative variables showed lower response variability than the categorical variables. While only 2 of 24 categorical questions showed an index of inconsistency lower than 20 on table 9-2, 4 of 32 did for the quantitative variables on table 9-3, along with another 4 items that showed no discrepancies between the original and reinterview surveys (the index of inconsistency cannot be computed for this second group of

[38] Because these are quantitative variables, the gross difference rate cannot be expressed as percentages in the same way that they can for categorical variables. It is possible to calculate them as a percentage of the survey estimate (to create a context for judging whether they are high or low), but this should be understood as a different way of examining the statistics, not a way of creating an equivalency with the approach used earlier. (In the categorical variables, the gross difference rates were not based on the survey estimate but as a percentage of the maximum level of differences that were possible for binary variables.) Alternatively, a closer way of matching the analysis used for categorical variables is that, if the quantitative variable were redefined to be a binary variable that indicated whether or not the reinterview response matched the original response, then the gross difference rate would be equivalent to the percentage that did not match, as shown in table 9-4.

4 items because of the lack of variation: all answers were 0 for both surveys). One reason that the index of inconsistency is often low is that no incidents or offenses were reported by many schools in a particular area, making it relatively easy for a school to give the same response in both surveys. On the other hand, the large number of zeroes often resulted in low variances of the total estimates (i.e., the denominator for the index of inconsistency, equation 9-3), which sometimes made the index of inconsistency appear artificially high. For example, the number of attacks with weapons that were hate crimes (question 16c1_3, appendix K) had no non-zero responses in the original survey among the reinterview respondents and an unweighted total of one in the reinterview. This variable showed low response variability on table 9-3 by two measures: there was a 99.7 percent agreement rate between the original responses and the reinterview responses, and the gross difference rate was 0.0; at the same time, the index of inconsistency was high (100.3). More generally, all ten of the variables for which the index of inconsistency was above 100 had relatively few non-zero responses (only question 21g1 had more than ten non-zero responses in either survey, with an unweighted total of 13 non-zero responses in the original survey among the reinterview sample, and 12 in the reinterview), suggesting an unstable index of inconsistency.

In short, the quantitative variables generally showed low response variability based on comparing the gross difference rate to the actual estimates, while the index of inconsistency also often either showed low response variability or might be invalidated based on the low variance of the total estimates. Thus, contrary to the original expectation, the quantitative variables showed greater reliability than many of the categorical variables. Much of the low response variability can be explained by the lack of incidents at many schools, making it relatively easy for the schools to consistently report the same statistic (i.e., zero) for both the original and reinterview surveys. Another possible explanation is that the quantitative data often were subject to greater telephone verification than the categorical data, because they allowed reviewers to check the ranges to see if they were reasonable, and to check for consistency among the various responses. Thus, the higher consistency might reflect the greater time and effort devoted to verifying the accuracy of the responses.

Sources of the Quantitative Data

For all of the quantitative questions on table 9-3 except question 9, additional questions were asked in the reinterview survey to provide more information about school administrative practices regarding these data. These additional questions asked about the source of the data, the use of distinctions and definitions specified in SSOCS:2000, the method of counting incidents, and the frequency of electronic record updates. Understanding how schools keep records and what they keep records of, is important information for implementing future surveys of this nature.

Some of the reinterview questions asked respondents to describe the primary source of data for counting the number of incidents, the number of disciplinary actions, and the number of offenses for special education students (see questions 16A, 21A, and 22A in appendix K). Response options describing the data source included: an electronic data file or tabulations, records were counted manually, the respondent made an estimate, the respondent knew the answer because of the small number involved, and other reasons. The distinction between electronic and manual records was made in the reinterview questionnaire because electronic records are likely to require less respondent effort unless the number of incidents is extremely small. However, it is not necessarily important from an accuracy viewpoint if the records were electronic or counted manually since both practices are subject to error. Manual counts allow the possibility of counting errors and electronic records allow for the possibility of programming errors, and may additionally allow less flexibility to check whether the definitions in the questionnaire were followed precisely. Either way, manual and electronic record keeping are a formalized system for collecting and maintaining information about certain aspects of student behavior. The benefit of using records is that it does not rely on the cognitive recall abilities of an administrator that may introduce significant bias.

As shown in table 9-4, between 39 and 58 percent of respondents said that they knew the answer because of the small number involved, depending on the particular type of data provided. In general, if no incidents occurred, or if there were only a very small number of incidents, a respondent would likely be able to remember the exact answer without referring to records, though respondents' memories could be faulty. Further, even if errors were made, the errors were likely to be small in size in such cases. However, one should use caution and consider the potential impact that small numbers can have on certain subgroup estimates. Depending on the statistic that is used, errors in reporting counts can have a large impact on smaller estimates. For example, over or under reporting five violent incidents when the estimate is ten is proportionally much larger compared to when the estimate is 100. One estimate is off by 50 percent and the other by 5 percent. The importance of such errors depends on several factors. One is whether there is a bias in the estimates (e.g., if underreporting is more common than overreporting). If there is no bias, then the actual estimates should not be greatly affected, but the reliability of the estimates is affected (i.e., the estimates would have larger standard errors). Another factor is the type of statistic being examined. The total number of incidents in the country would not be greatly affected by such errors, as also would be true for the percentage of incidents that occur at a particular type of school. On the other hand, the average number of incidents at a particular kind of school might be affected substantially in proportional terms.

Another group of schools based its statistics on records in some way, either by using an electronic data file or computer tabulations (11-21 percent of schools) or counting the records manually (10-20 percent of schools). Between 8 and 24 percent said their responses were estimates.

Table 9-4. Percentage of schools using various primary sources of data when reporting incidents, offenses, and disciplinary actions: 2000

Questionnaire item	Electronic data file or computer tabulations	Counted records manually	Made estimate	Knew because of small number	Other
Number of incidents of fights and of theft/larceny (Q16)					
Total number	19	16	24	39	3
Number reported to police	17	18	10	52	4
Hate crimes	18	10	9	58	4
Gang-related	21	11	8	56	4
Number of disciplinary actions (Q21)					
Removals for at least 1 year	20	17	14	49	—
Transfers to specialized schools	14	20	17	49	—
Out-of-school suspensions	15	20	21	45	—
Other	18	18	21	42	0
No disciplinary action	14	17	16	50	3
Data on offenses by Special Education students (Q22)	11	13	14	56	7

—No responses appeared in this category.
NOTE: Detail may not add to totals because of rounding. All estimates are weighted. The complete question wording is shown in appendix K.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Survey Distinctions and Definitions

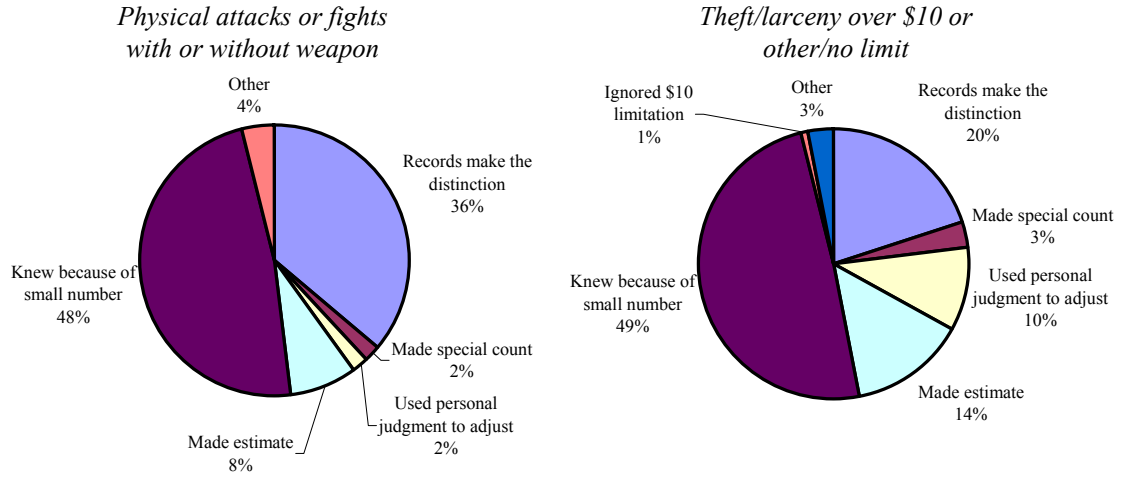
Another concern was that respondents might not answer questions making the distinctions or using the definitions specified in the SSOCS:2000 questionnaire. Three specific types of survey questions in which these distinctions and definitions were thought to be especially important were examined. First, respondents were asked to provide separate responses for fights with weapons and fights without weapons, while school records or tabulations may not have made that distinction. Second, respondents were asked to report only on thefts of $10 or more, while school records may not have indicated the amount or may have used a different cutoff. Third, schools were asked for the number of removals with no continuing services for at least 1 year, instead of using the word *expulsion*. (The word *expulsion* was purposely not used because schools might define expulsions differently; instead, the word *removal* was used, and the time period of the removal was clearly stated.) Thus, all three of these areas created the potential for inaccurate responses because of schools' inability or difficulty in providing the specific types of data that were requested.

The reinterview questionnaire asked respondents how they were able to distinguish between fights with and without weapons (see question 16B in appendix K). Response options included: their records already made that distinction; they made a special count; they determined the total number of fights, and used personal judgment to divide them; they made their best estimate without reference to data files, tables, or records; they knew the answer because of the small number involved; or other reasons. They were also asked what they did to limit their response to thefts of $10 or more (see question 16C in appendix K). The same response options were given, plus an additional one indicating that the limitation was ignored and the number that was available was provided.

Again, from a data accuracy viewpoint, the ideal would be for institutions to have records that made the distinctions requested, so that the data collection would eliminate the need to impose a distinction on information not collected in the original record. Further, if the information was organized in this fashion it would presumably be easier to supply. Other options included making a special count (e.g., by manually searching through the records for the requested information) or saying the respondents knew the answer because of the small number of incidents involved. Both of these options offer the possibility of accurate data, but allow more room for human error, and in the case of making special counts, may increase respondent burden. In terms of data accuracy, respondent burden is important for several reasons: respondents may be less likely to complete the entire questionnaire or individual items if it is burdensome, and burden increases the likelihood that respondents will take shortcuts in completing the questionnaire that may affect data accuracy. Other possible options, such as somehow adjusting the totals taken from records to allow for the distinction, making an estimate, or simply ignoring the distinction (in the case of thefts over $10), create additional room for respondent error and increase the likelihood of supplying inaccurate data.

Figure 9-1 shows that 48 percent of the reinterview respondents said they were able to distinguish between fights with weapons and fights without weapons because of the small number of incidents involved, while 36 percent said their records already made that distinction. Eight percent made their best estimate without referring to records, four percent stated other reasons, two percent said they made a special count, and another two percent used personal judgment to divide the total number of fights. With regard to limiting responses concerning the number of thefts to $10 or more, 49 percent said they knew because of the small number involved. Twenty percent said their records made the distinction, 14 percent said they made an estimate, 10 percent used personal judgment, 3 percent made a special count, another 3 percent gave other reasons, and 1 percent said they ignored the $10 limitation.

Figure 9-1. Percentage of schools reporting various ways in which distinctions were made when providing data on the number of incidents: 2000



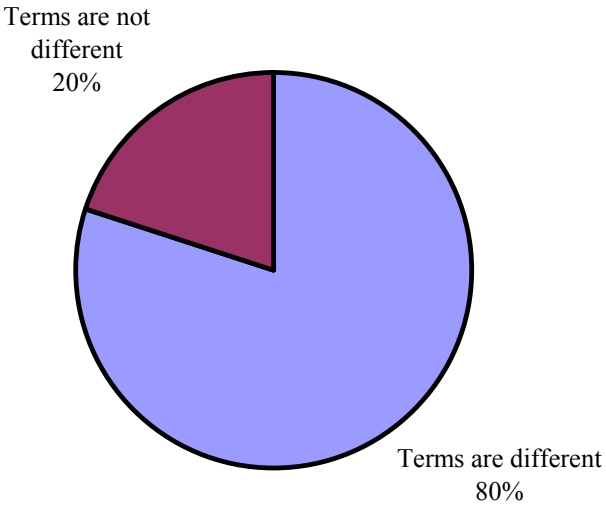
NOTE: All estimates are weighted. The complete question wording is shown in appendix K.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

These results suggest that a large number of schools can make the requested distinctions, and it is encouraging that only 1 percent said that they ignored the $10 limitation when reporting the number of thefts. The two categories that probably allowed the most room for error were the use of estimates and personal judgment: between 8 and 14 percent made estimates and between 2 and 10 percent used personal judgment.

Finally, the reinterview questionnaire asked respondents whether the term "removal with no continuing school services for at least 1 year" differed from the school's definition of expulsion, and if so to provide their definition of expulsion (see question 21B in appendix K). This question was asked both because the data may be regarded as more accurate if they are requested in a format readily available to the schools, and because the construction of the questionnaire might be simplified in future collections if the word *expulsion* could be used. Eighty percent of the reinterview respondents said that the two terms were different (figure 9-2). For example, the ways that the schools' definitions differed were that expulsion could be for less than 1 year (e.g., until the end of the term or, at some schools, any removal for more than 10 days), or expulsion came with some kind of services (e.g., at-home schooling or education at an alternative school). The large amount of variation found in the definitions that schools use suggests that *expulsion* is not an appropriate term for usage in the questionnaire. By using the word *removals*, one at least draws attention to the specific definition that is being used, and thus increases the chances that the question will be answered consistently across schools. Note, however, that drawing schools' attention to the definition does not guarantee that they use it. If schools do not maintain data on removals but only on expulsions, as defined by that school, some might still respond with their data on expulsions, thinking it

provides the closest approximation to what was being requested. Thus, using *removals* instead of *expulsions* does not guarantee complete uniformity in how schools respond to the question, but does provide some way of standardizing the data provided across schools.

Figure 9-2. Percentage of schools indicating that *removal with no continuing school services for at least 1 year* was different from their definition of *expulsion*: 2000



NOTE: All estimates are weighted. The complete question wording is shown in appendix K.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

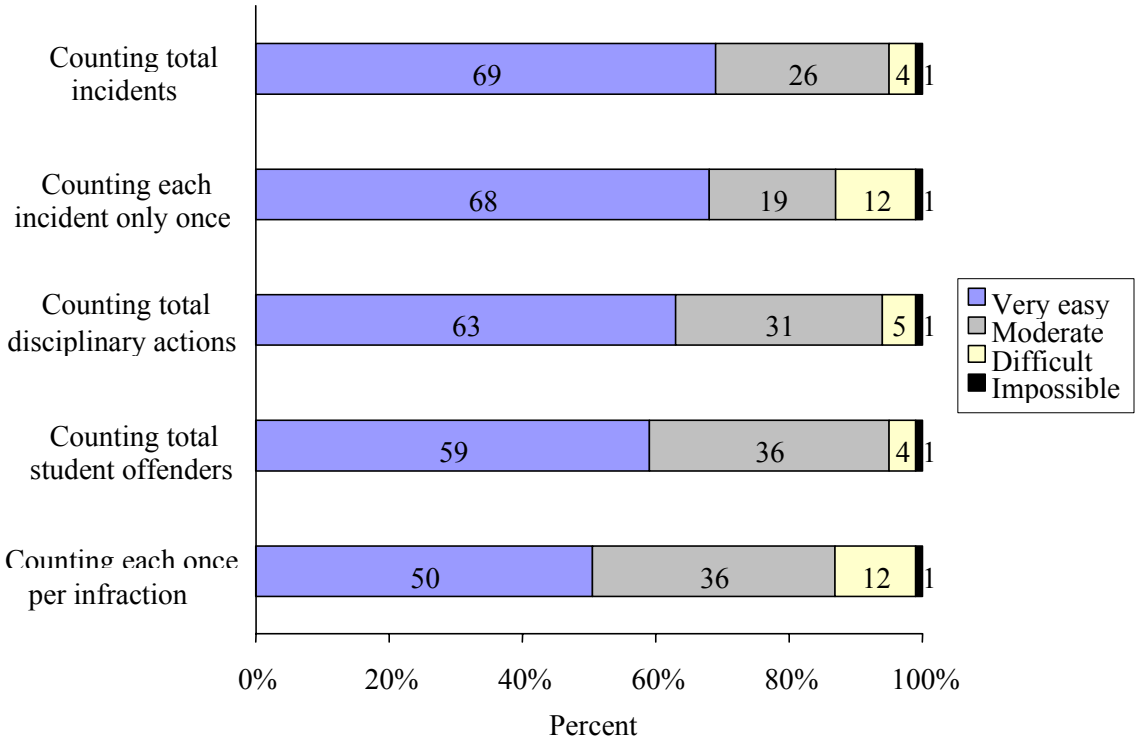
Method of Counting Incidents

Question 16 of the original SSOCS:2000 questionnaire (see appendix C) asked schools to provide the number of incidents, rather than the number of victims or offenders, in each of 15 categories. It also asked that each incident be counted only once, counting only the most serious offense when an incident involved multiple offenses (e.g., if an incident included both rape and robbery, to count only the rape). Thus, one possible source of data inaccuracies could occur if respondents had difficulty counting incidents in the way requested.

To examine this issue, respondents were asked how easy (using the categories *very easy*, *moderate*, *difficult*, and *impossible*) it would be to provide counts in several different ways when reporting incidents (see question 16D in appendix K). These alternatives included: counting each incident only once (as on the current questionnaire as described above), counting each incident once per infraction (i.e., allowing double counting of incidents when there were multiple infractions), counting the total number of incidents, counting the total number of student offenders, and counting the total number of disciplinary actions.

Figure 9-3 shows that 50 percent or more of the respondents thought each of the listed counting formats would be very easy, and 85 percent or more thought each would be either very easy or moderately easy. Among the five approaches listed, 69 percent of the schools said it would be very easy to provide the total number of incidents, 68 percent said it would be very easy to count each incident only once (the current format of the question), 63 percent said it would be very easy to provide the total number of disciplinary actions taken in response, 59 percent said it would be very easy to provide the total number of student offenders, and 50 percent said it would be very easy to provide each incident once per infraction. All of the approaches appear feasible, so if one is clearly the most useful from an analytic perspective, then that approach could be used in subsequent SSOCS collections. The current approach used in the SSOCS:2000 survey, counting each incident only once, is as easy to provide by the respondent as the alternatives (when comparing "very easy" categories).

Figure 9-3. Percentage of schools indicating ease of providing various counts when reporting incidents: 2000



NOTE: Detail may not add to totals because of rounding. All estimates are weighted. The complete question wording is shown in appendix K.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Frequency of Electronic Record Updates

Another way in which the accuracy of the SSOCS:2000 data can be assessed is by determining how current (and thus complete) they were when reported. If records are updated frequently, then the data supplied by the schools are more likely to be current as of the time the survey was completed, and thus as complete as possible. On the other hand, the greater the delay in updating the records, the more likely it is that the data supplied on the questionnaire might be incomplete. The frequency of updating records may be a more important issue when electronic (rather than manual) records are kept: there may be a separate process to get data entered into an electronic record system, or the data might first have to be transferred somewhere else (e.g., to the district office) before it is updated. Thus, since electronic records probably present the most potential for delayed updates, the reinterview study asked about the frequency of updates for only electronic records. However, one should also consider the fact that some schools may not update records daily, weekly or even monthly if no incident has occurred. An incident may only happen one or two times per school year for some incident types or for some schools.[39]

Specifically, the reinterview questionnaire asked how often schools updated their electronic records of crimes (see question 16E in appendix K). The response categories offered were: daily, weekly, monthly, less frequently than monthly, and the school does not have electronic records. Sixty-four percent of the schools reported that they kept electronic records.

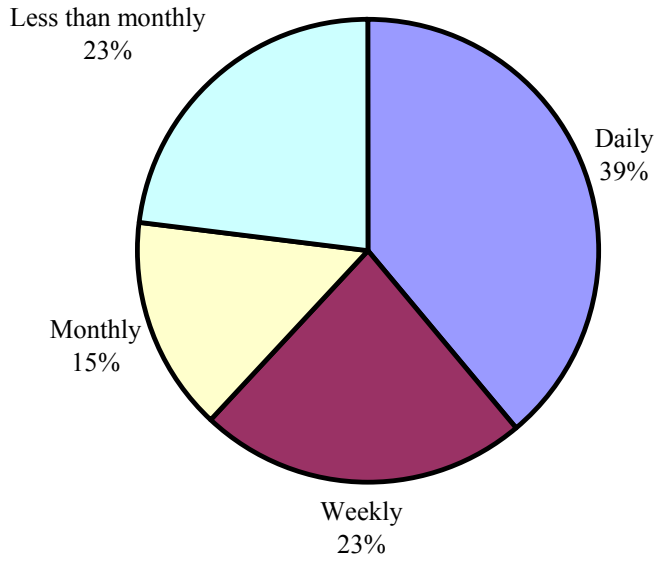
Focusing on just those schools with electronic records, figure 9-4 shows that 39 percent of these schools updated the records on a daily basis and another 23 percent updated them on a weekly basis.[40] The remainder updated the data monthly (15 percent) or less often (23 percent). The 62 percent of schools that updated their electronic records either weekly or more often is a substantial proportion, but that still leaves 38 percent that updated their data only monthly or less often. This is a potentially large proportion of schools that may be supplying outdated data or may be updating on an as-needed basis if these events occur infrequently at their schools.

---

[39] It is not clear how a school would interpret this question if updates were rare because incidents were rare. A school might respond in terms of the time period that would elapse between the incident and the records update; e.g., it might respond daily if records were updated within the next day (even though changes were not made on a daily basis), or weekly if records were updated within the next week. On the other hand, a school might report in terms of the time elapsed between updates, which could be much longer. Of those schools that reported five or fewer physical attacks or fights without weapons (question 16c1_2, which was the variable that showed the greatest range, and thus the most likely to prompt updates), 31 percent said they made updates on a daily basis, 24 percent made updates on a weekly basis, 13 percent made updates on a monthly basis, and 32 percent made updates less frequently than monthly.

[40] Note that these statistics are not directly comparable to those in table 9-5, in which respondents reported on the *primary* source of data that they used to answer specific questions rather than on whether computer tabulations were available.

Figure 9-4. Percentage of schools with electronic records of school crimes making record updates over selected time periods: 2000



NOTE: All estimates are weighted. The complete question wording is shown in appendix K.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Discrepancy Analysis

If respondents gave different answers to the quantitative questions in the reinterview questionnaire than they did on the original SSOCS:2000 survey, they were called back to verify the reasons for the discrepancy. An illustrative dialogue is presented in appendix K to show how these telephone interviews were conducted. First, the discrepancy was described, and the respondent was asked (through an open-ended question) to explain the reason for the discrepancy. Second, the respondent was provided with a list of five possible explanations plus a sixth open-ended item for providing other explanations, and asked to indicate which of them were reasons for the discrepancy. Thus all respondents considered a standard set of categorized options to facilitate comparison of their responses. The listed explanations were:

- My most recent responses included some incidents that hadn't happened when I first completed the survey.
- One answer was an estimate, while the other was based on checking our records.
- I tried to remember our original response, but didn't remember it exactly.
- A different person completed the question each time.

- I/we consulted with someone else when answering it one time, but did not talk to that person the other time.
- Other (with space to write in a response)

Third, if the respondent provided more than one reason for the discrepancy from this list, he/she was asked which one best explained the reason for the difference. Fourth, the respondent was asked which answer (i.e., the response in the original survey or the response in the reinterview survey) could be considered the most accurate. If there were multiple discrepancies, then this dialogue was followed for each discrepancy individually.

A total of 96 of the 112 reinterview respondents were eligible for the discrepancy interviews. Of those, 92 completed the discrepancy interviews, while 4 refused to participate further in the study. A total of 418 discrepancies were included in the analysis (weighting to 238,800 discrepancies in the entire population). The number of discrepancies among the 92 respondents ranged from 1 to 11.

Reasons for the Discrepancies

Table 9-5 shows the number of discrepancies that appeared for each question set, and the percentage of discrepancies that were primarily explained by each reason. To simplify the analysis, and because there often were only a few discrepancies per questionnaire item, the responses are tabulated in terms of the overall question (i.e., questions 9, 16, 21, and 22) rather than the individual items within each question (see table 9-4 for descriptions of the individual items). For example, since the portion of question 16 that was reproduced on the reinterview questionnaire had three rows and four columns, there was a potential for up to 12 discrepancies to appear within that question. The statistics in table 9-5 are based on the number of discrepancies, not the number of respondents. If a school did not have any discrepancies for a particular question (e.g., question 9), then that school is not included in the statistics for that question; on the other hand, if a school had several discrepancies within a particular question (e.g., 3 discrepancies within the 12 items included for question 16), then each discrepancy is counted separately (e.g., as 3 responses in the example provided).

Table 9-5. Number and percentage of discrepancies between the original and reinterview questionnaires that were explained by various reasons, by question: 2000

Number of and reasons for discrepancies	Primary reason for discrepancy: Question number: 9	16	21	22	All 4 questions
Total number of discrepancies					
Weighted	45,900	108,600	59,000	25,300	238,800
Unweighted	76	192	96	54	418
Reasons: Total percentage	100	100	100	100	100
Some incidents hadn't happened when survey was first completed	4	10	19	8	11
One answer was an estimate, and the other was based on records	9	30	24	18	23
Didn't remember original response exactly	3	11	2	6	7
Different person completed each questionnaire	2	8	7	10	7
Consulted someone else for only one of the surveys	0	3	2	6	2
Other	83	39	46	52	51

NOTE: Detail may not add to totals because of rounding. Percentages are weighted. The complete question wording is shown in appendix K. The counts in this table cannot be directly compared to those in table 9-4 because they do not represent all discrepancies that appeared, but rather all for which there are data in the discrepancy interviews.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Eighty-three percent of the discrepancies for question 9 had a primary reason other than the five main options examined in the reinterview questionnaire, along with 39 percent for question 16, 46 percent for question 21, and 52 percent for question 22. When respondents were asked to specify these idiosyncratic "other" reasons, some examples of their responses were human error (e.g., failing to include sports events as times when law enforcement personnel were present, thinking of a full-time officer as serving for 40 hours when a school week actually was 30 hours, or placing a number in the wrong location), failing to follow the directions closely (e.g., reporting an incident that occurred outside of the school year, confusion over the location of an event, and using different dollar limits for theft/larceny), misinterpretation of the question, and being unable to explain how a number was obtained.

Between 9 and 30 percent of the discrepancies were a result of one report being based on records, while the other was based on an estimate. Nineteen percent of the discrepancies for question 21 (and between 4 and 10 percent for the remaining questions) occurred because additional incidents happened between the time of the original survey and the reinterview survey.[41] Between 2 and 11 percent of the discrepancies occurred when respondents tried to remember their original response but did not remember it exactly. Between 2 and 10 percent of the discrepancies were explained by a different person completing each questionnaire. (The reinterview questionnaire was mailed to the same person who completed the original SSOCS:2000 questionnaire, with instructions for that person to complete the

[41] Since question 9 is not concerned with the number of incidents but rather the number of hours security personnel were on duty, one would not expect any principals to give this reason when explaining a discrepancy. However, they may be referring to changes in schools' security practices between the time they completed the original questionnaire and the time they completed the reinterview questionnaire.

reinterview survey. However, sometimes a different person did complete the reinterview survey, including occasions when the original person was not available.) Finally, the statistics that were provided sometimes depended on whether other people were available when one of the two questionnaires was completed. For example, a key person might not have been available because of other school demands, his/her vacation schedule, or for other reasons. Between 0 and 6 percent of the discrepancies were explained by such a person being consulted for only one of the two surveys.

Conclusions

The categorical questions in SSOCS:2000 showed less reliability than anticipated. This finding cannot necessarily be generalized to all of the categorical questions, however. Some categorical questions did perform well, while others did not. The greatest problems appeared to be associated either with questions 3 and 4, question 8c, and question 12. All of these questions should be examined further to determine what changes might make the responses more reliable. For example, changes in question wording and/or the use of more extensive examples might further clarify the question items and remind the respondent of programs that otherwise might be forgotten. Also, the removal of the skip pattern in questions 3 and 4 might keep respondents from prematurely reporting their schools had no formal programs before they had considered all relevant programs.

The quantitative questions, despite the higher level of burden associated with them, tended to outperform the categorical questions, and showed a high level of reliability. The amount of error that did appear was low when compared to the actual survey estimates, so that its impact was relatively small. Further, a range of 39 percent to 56 percent of respondents were able to answer the questions relatively easily because of the small number of incidents or offenses involved, and 24 percent to 37 percent did base their estimates on school records so their answers were not dependent on their cognitive recall (which may introduce significant bias). No changes appear necessary to the quantitative questions based on the reinterview results. Two of the reasons that the quantitative questions often outperformed the categorical questions may be that: (1) the data often were zeroes, which may have been easy to report; and (2) in general, greater telephone verification efforts probably were given to the quantitative questions because of the possibility of performing range changes and sometimes consistency checks (with other variables) to verify the accuracy of the reports.

From a data collection viewpoint, the timing of the survey administration (March 27 through August 15) was poor. Even the beginning of data collection was relatively close to the end of the school year, when administrators were busy with end-of-the-year activities, and delays in responding led to

increased conflicts with end-of-the-year activities. The summer was also difficult because many principals (and other administrators who might assist them) were not at their schools during the summer, or were present only for a limited number of hours. These problems in data collection were anticipated, and collecting data in the fall of 2000 was considered as an alternative; however, part of the reason for choosing the earlier data collection period was to improve data reliability for those schools that were making estimates, based on the assumption that if principals were contacted well after the end of the school year, their recall might be less accurate. The results from the reinterview study suggest that data reliability does not appear to be an important consideration with regard to survey timing, at least with regard to the time period that was observed. Only 11 percent of the discrepancies were explained in a way that was directly related to survey timing (i.e., that some incidents had not yet happened when the first response was given), while, more fundamentally, the number and size of the discrepancies were relatively small. Since the reinterview study was completed prior to the fall of 2000, the study cannot provide definitive estimates of how the reliability of a fall data collection period would compare with the period that was actually used.

Additional questions were asked in the SSOCS:2000 reinterview study that were designed to explore whether changes in question wording or in the administration of the survey would be helpful in future collections. These questions covered topics such as the use of distinctions and definitions in the survey, ways to count incidents, the understanding of the term *expulsion*, and the frequency of electronic record updates.

Several questions were asked in the reinterview study about the wording of questionnaire items. One such question concerned distinctions that respondents were asked to make in question 16 of the SSOCS:2000 questionnaire, where they were asked to provide separate data on physical attacks with weapons and those without weapons, and thefts over $10 in value. Most respondents were able to make the distinctions requested on the questionnaire, either because their records made that distinction or the number of cases was sufficiently small that they knew the correct answer. Thus, the continued use of such distinctions in future collections appears reasonable.

No clear answer emerged concerning the best way to ask schools to quantify the number of incidents in question 16. Respondents generally indicated that the various counting formats proposed would be very easy to provide. Unless future collections want to consider alternatives that may provide substantive differences based on analytical reasons (e.g., number of student offenders), the current format, asking for each incident only once, appears acceptable.

Most schools reported that they use a different definition of *expulsion* than the terminology used in the SSOCS:2000 questionnaire, in question 21: *removal with no continuing school services for at least 1 year*. But the meaning of *expulsion* varies widely from one school to another, so schools' answers would not be comparable even if that term were used. Further, even if a definition of *expulsion* was provided, schools might still use their own definition. Therefore, using the terminology about removals in its current form on the questionnaire helps to promote consistency across schools in their reporting.

Other questions were asked that might help in the administration of future SSOCS collections. One such question was the frequency of updating electronic records. The data show that 39 percent of the schools with electronic records updated their electronic records monthly or less often, suggesting that it may be desirable to allow at least one month after the end of the school year for records to be updated.[42] This might result both in more up-to-date reports, and potentially also increase the number of schools using electronic data files or computer tabulations as their primary data source. However, collecting SSOCS data after the school year ends may mean that some key school personnel could be unavailable to complete the survey. There are thus multiple tradeoffs in selecting the best survey administration dates. Collecting data late in the school year risks using data that have not been updated, conflicts with end-of-the-year activities (i.e., increasing the difficulty of obtaining responses), and risks collecting data that are incomplete because some incidents have not yet occurred. Collecting data in the summer may allow time for the records to be complete and updated, but is difficult because many people are not at the school during the summer. Collecting data in the fall of the following school year avoids the end-of-the-year and summer conflicts, and allows the data to be complete and updated, but may make recall more subject to error for those respondents giving estimates.

Some problems and errors discussed in this chapter might also be explained by the mailed self-response mode of data collection used in SSOCS:2000 and by other factors often associated with reinterview studies. Mail surveys have both advantages and disadvantages when compared to the use of in-person or telephone interviews. Mailed self-report questionnaires by their very nature reduce the opportunities of respondents to query the interviewer about problems they have with such things as definitions, question formatting, and question applicability. SSOCS:2000 did provide a toll-free number so that respondents could ask questions if they desired, but this might be less convenient than asking an interviewer who was already present (in person, or on the telephone). On the other hand, the use of telephone interviews for data collection might lead respondents to provide estimates (as a way of providing an instant response) rather than checking their records when reporting on the frequency of

[42] It may be that some schools updated their records promptly but indicated their updates were infrequent because incidents were rare. For such schools, a delay might not be necessary in order to get accurate reports.

incidents, which could reduce data reliability. Mail surveys also might be less threatening than personal interviews when supplying data on sensitive topics, because of their more impersonal nature. Mail surveys have mixed advantages and disadvantages with regard to skip patterns on questionnaires. The use of an interviewer helps to ensure that a skip pattern is observed, but this can be either a strength or a weakness. Sometimes a respondent's mistaken failure to follow the skip pattern, or the respondent's skimming of the skipped question to verify what alternatives were available, leads to a discovery that the respondent's initial response was incorrect. Such mistakes might not be caught when an interviewer is used. Given the tradeoffs associated with using a mailed self-report survey versus an in-person or interviewer type data collection mode, alternative solutions such as simple adjustments to the questionnaire (e.g., a revised question format or revised directions) could help alleviate some problems.

The administrative problems associated with implementing a reinterview study may also have contributed to problems and errors, although it was not always possible to disentangle whether errors were associated with administration of the original survey or whether they were generated in part by the reinterview survey's inability to replicate the original survey conditions. For example, the reinterview study shows that some schools used estimates to provide responses for one survey and records for another. But it did not ascertain which source was used when. Thus, the study cannot show the specific effect of the reinterview methodology on these choices. In addition, those idiosyncratic "other" factors discussed above, such as human error, failing to follow directions, and misinterpreting questions, may have also contributed to problems and errors.

Standard Errors for Tables and Figures in Chapter 9

Table 9-2a. Standard errors of estimates of gross difference rates and indexes of inconsistency for categorical questions in SSOCS:2000

Question	Description	Original survey estimate (percentage)	Gross difference rate	Index of inconsistency
Q3	Formal program prevent/reduce violence	1.8	6.6	15.1
Q4a	Prevention training (e.g.,social skills)	1.4	6.7	14.6
Q4b	Behavioral modification for students	1.3	5.2	10.1
Q4c	Student counseling/social work	1.4	6.6	14.6
Q4d	Individual mentoring/tutoring students	1.4	6.2	13.4
Q4e	Recreation/enrichment student activities	1.4	6.8	13.6
Q4f	Student involvement resolving problems	1.4	5.7	11.8
Q4g	Promote sense of community/integration	1.4	6.3	11.9
Q4h	Hotline/tipline to report problems	0.9	4.9	12.6
Q8a	Security used during school hours	1.2	3.3	6.9
Q8b	Security while students arrive/leave	0.9	3.1	7.8
Q8c	Security at selected school activities	1.3	5.3	10.5
Q8d	Security when school not occurring	1.0	4.9	11.7
Q8e	Other times security used	0.7	3.1	34.7
Q12a	Efforts lmtd by lack of tchr training	1.7	5.0	11.5
Q12b	Efforts lmtd by lack of altrntive plcmnt	1.4	5.5	13.4
Q12e	Efforts lmtd by lack of parent support	1.2	6.7	17.6
Q12l	Efforts lmted by fed policies/disabled	1.6	6.1	12.3
Q12m	Efforts limited by other fed. policies	1.5	5.7	10.0
Q19a	How often student racial tensions	0.4	1.5	26.6
Q19b	How often student bullying occurs	1.2	4.1	9.8
Q19f	How often undesirable gang activities	0.9	3.7	11.1
Q19g	How often undesirable cult activities	0.5	1.4	8.2
Q27	Crime where students live	1.3	5.0	14.4

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Table 9-3a. Standard errors of estimates of gross difference rates and indexes of inconsistency for quantitative questions in SSOCS:2000

Question	Description	Percentage that did not match	Original survey estimate	Gross difference rate	Index of inconsistency
Q9a	Total hours security on duty per week	5.6	25,728.9	32.5	5.6
Q9b	Total hours security wore uniform	8.3	21,032.7	59.1	10.5
Q9c	Total hours security carried a firearm	8.6	15,295.6	63.3	11.3
Q16c1_1	# of attacks/with weapon - total	2.1	2,481.2	0.1	17.1
Q16c1_2	# of attack/weapon/reported police	2.1	683.9	0.1	17.8
Q16c1_3	# of attack with weapon/hate crimes	0.3	71.5	0.0	*
Q16c1_4	# of attack with weapon/gang-related	0.4	178.1	0.0	*
Q16c2_1	# of attacks/no weapon – total	5.9	59,618.4	75.1	17.6
Q16c2_2	# of attacks/no weapon/reported	4.5	5,704.0	2.3	2.9
Q16c2_3	# of attacks/no weapon/hate crimes	1.9	3,936.9	35.5	2.8
Q16c2_4	# of attacks/no weapon/gang-related	1.9	1,431.3	0.5	39.7
Q16f1	# of theft/larceny - total	6.9	9,209.9	2.9	15.5
Q16f2	# of incidents theft/larceny/reported	7.8	5,562.7	2.2	15.6
Q16f3	# of incident theft/larceny/hate crime	0.0	181.4	0.0	*
Q16f4	# of incidents theft/larceny/gang	0.3	538.8	0.0	*
Q21g1	# of removals for attacks/fights	5.5	5,332.4	48.6	4.3
Q21g2	# of transfers for attacks/fights	3.8	2,141.7	2.2	26.5
Q21g3	# of suspensions for attacks/fights	5.9	25,861.7	34.1	22.8
Q21g4	# of other actions for attacks/fights	7.3	47,249.4	220.5	34.0
Q21g5	# of no actions for attacks/fights	1.8	2,332.4	0.1	1.3
Q22a1_1	Placement changed after hearing/total	4.8	2,869.1	2.3	14.6
Q22a1_2	Placement chngd/hearing, drugs/weapons	2.9	758.5	0.0	21.0
Q22a2_1	Placement changed after injunction/total	3.7	654.4	0.1	1.2
Q22a2_2	Placement changed/injnction, drugs/wpns	0.4	142.8	0.0	*
Q22a3_1	Placement chnge w/o hearing, total	4.1	2,596.6	1.9	28.3
Q22a3_2	Placement chnge w/o hearing, drgs/wpns	1.0	427.0	0.0	38.3
Q22b1_1	No change,hearing/session not held,total	2.4	2,641.5	0.7	0.7
Q22b1_2	No change,hearing not held,drugs/wpns	0.6	480.9	0.0	*
Q22b2_1	Hearing did not approve change, total	2.3	1,359.8	0.4	28.1
Q22b2_2	Hearing did not approve chnge, drgs wpns	0.0	414.3	0.0	*
Q22b3_1	Court did not approve change, total	0.0	225.5	0.0	*
Q22b3_2	Court did not approve chnge, drgs/wpns	0.0	184.0	0.0	*

*The index of inconsistency cannot be computed for these variables because of the small amount of variation in the reinterview sample.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Table 9-4a. Standard errors of percentage of schools using various primary sources of data when reporting incidents, offenses, and disciplinary actions: 2000

Questionnaire item	Electronic data file or computer tabulations	Counted records manually	Made estimate	Knew because of small number	Other
Number of incidents of fights and of theft/larceny (Q16)					
Total number	4.7	4.7	5.9	5.7	1.9
Number reported to police	5.4	4.7	2.5	6.1	2.1
Hate crimes	5.2	3.3	2.3	4.6	2.1
Gang-related	6.8	3.3	2.4	5.9	2.1
Number of disciplinary actions (Q21)					
Removals for at least 1 year	3.4	4.1	4.9	4.2	—
Transfers to specialized schools	3.3	4.3	6.4	5.1	—
Out-of-school suspensions	3.2	4.3	5.3	5.1	—
Other	3.5	4.9	6.1	6.8	0.3
No disciplinary action	3.2	4.3	4.7	3.9	3.1
Data on offenses by Special Education students (Q22)	2.8	3.6	3.2	6.3	4.2

—No responses appeared in this category.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Table 9-5a. Standard errors of number and percentage of discrepancies between the original and reinterview questionnaires that were explained by various reasons, by question: 2000

Number of and reasons for discrepancies	Primary reason for discrepancy: Question number: 9	16	21	22	All 4 questions
Total number of discrepancies (weighted)	11,967	14,531	10,457	6,063	31,343
Some incidents hadn't happened when survey was first completed	3.3	2.7	5.5	4.2	2.5
One answer was an estimate, and the other was based on records	4.4	5.9	7.1	8.3	4.3
Didn't remember original response exactly	2.3	6.6	1.1	3.6	3.1
Different person completed each questionnaire	1.5	2.7	3.3	4.6	2.3
Consulted someone else for only one of the surveys	0.0	1.6	1.2.	5.5	1.5
Other	5.7	8.5	5.9	11.9	5.1

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Figure table 9-1a. Standard errors of percentage of schools reporting various ways in which distinctions were made when providing data on the number of incidents: 2000

Method of making the distinction	*Physical attacks or fights with or without weapon*	*Theft/larceny over $10 or other/no limit*
Records made the distinction	5.6	5.0
Made special count	1.3	1.4
Used personal judgment to adjust	0.9	3.9
Made estimate	2.9	5.3
Knew because of small number	5.5	5.9
Ignored $10 limitation	NA	0.8
Other	1.9	1.9

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Figure table 9-2a. Standard errors of percentage of schools indicating that *removal with no continuing school services for at least 1 year* was different from their definition of *expulsion*: 2000

Definition of expulsion and removals	Percentage
Terms are different	4.6
Terms are not different	4.6

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Figure table 9-3a. Standard errors of percentage of schools indicating ease of providing various counts when reporting incidents: 2000

Measure of incidents	*Very easy*	*Moderate*	*Difficult*	*Impossible*
Counting total incidents	5.6	4.9	1.6	0.8
Counting each incident only once	6.3	4.6	3.1	0.8
Counting total disciplinary actions	5.6	5.2	1.6	0.8
Counting total student offenders	6.8	6.8	1.6	0.8
Counting each once per infraction	7.4	6.3	2.8	0.9

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

Figure table 9-4a. Standard errors of percentage of schools with electronic records of school crimes making record updates over selected time periods: 2000

Frequency of updating records	Percentage
Daily	7.5
Weekly	4.9
Monthly	4.9
Less than monthly	5.7

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS) Reinterview Study, 2000.

10. COMPARISON OF SSOCS:2000 ESTIMATES WITH STATISTICS FROM OTHER SOURCES

The primary purpose of SSOCS:2000 was to provide school-based statistics, while most other sources of school crime data are based on either household surveys or sources such as police records. For this reason, among others discussed below, SSOCS:2000 cannot be expected to provide estimates that are comparable to those other sources. Still, it is helpful to report estimates from other sources in order to gain a fuller perspective on the SSOCS:2000 data.

Populations of Interest and Data Sources

For SSOCS:2000, the population is defined as all regular public schools, excluding schools in the outlying U.S. territories, ungraded schools, and those with a high grade of kindergarten or lower. The sample was based on the 1997–98 NCES Common Core of Data (CCD) Public School Universe File, but the weights were poststratified to sum to the totals in the 1998–99 CCD Public School Universe File (the most recent available CCD file at the time of weighting) so that the totals would more closely correspond with the year of the SSOCS:2000 Survey (1999–2000).

Some other studies conducted by the federal government that report data on school crime and safety discussed in this chapter are the School Associated Violent Deaths Study (SAVDS), the National Crime Victimization Survey (NCVS), and the Fast Response Survey System (FRSS) Principal/School Disciplinarian Survey on School Violence. All data discussed from sampled studies are weighted estimates.

The School Associated Violent Deaths Study is an epidemiological study that provides descriptive data on all school associated violent deaths in the United States. School associated deaths include deaths at school, on the way to or from school, or at school-sponsored events. Deaths are first identified through database searches, and more detailed information is collected through interviews of police or examination of police records, and school officials.

The National Crime Victimization Survey (NCVS) is a nationally representative sample of U.S. households. Residents 12 years and older from about 55,000 households annually are interviewed every 6 months over the course of 3 years. Respondents are asked to recall the details of all victimizations that may have occurred to them in the previous six months reference period. The School Crime Supplement

(SCS) is an additional set of questions added to the NCVS interview. Residents ages 12 to 18 who have attended public or private school in the previous 6 months are asked about their experiences regarding school violence and safety. The final response rate for the 1999 NCVS and SCS are about 85 and 73 percent, respectively.

The Principal/School Disciplinarian Survey on School Violence was a nationally representative survey of 1,234 regular public schools conducted during the spring and summer of 1997. The survey was mailed to school principals, who were asked to have it completed by the person most knowledgeable about discipline issues at the school. The weighted response rate was 89 percent.

Methodological Considerations in Data Comparisons

While some findings from these studies are discussed here, readers should use caution in attempting to compare them to the findings from SSOCS:2000. These surveys use different levels of analysis, are not from the same populations, do not cover the same age groups, and are from different years.

General Comments on the SSOCS:2000 Estimates

The estimates presented here from SSOCS:2000 represent just a small portion of the data collected for the survey, and focus on overall totals rather than on detailed breakdowns by school characteristics. One reason is that only a relatively small number of items appear across the various surveys; for example, SSOCS:2000 collected a great deal of information about school practices that were not examined in SAVDS or NCVS. Another reason is that the purpose of this analysis is to provide a general sense of findings from the various surveys, not to provide a highly detailed analysis.

Unless noted otherwise, all SSOCS:2000 estimates presented here are weighted estimates in order to provide nationally representative results.

Data Findings

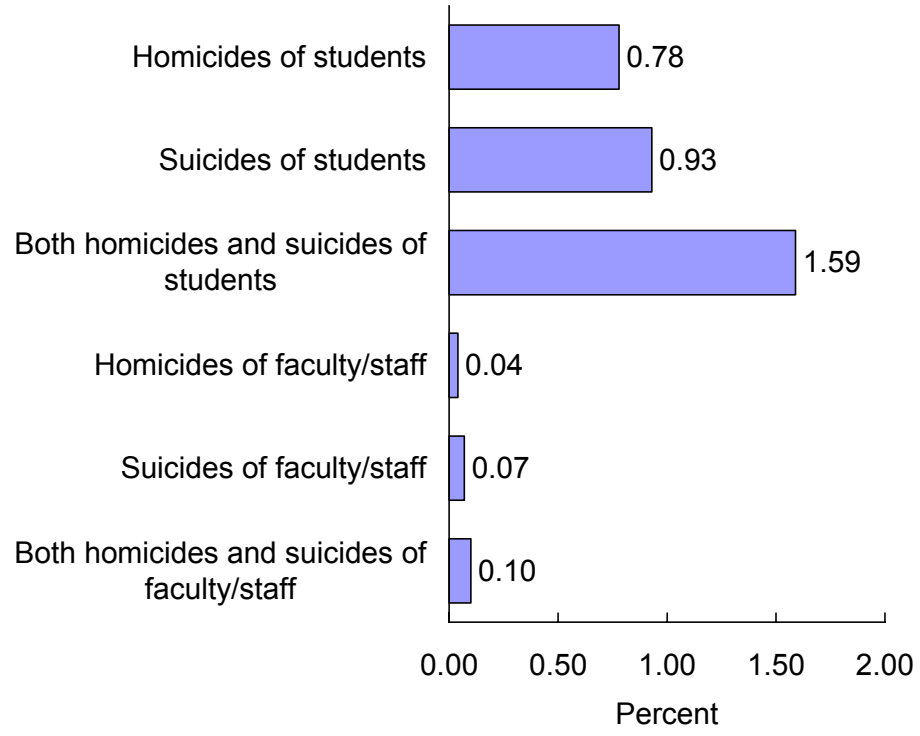
Deaths at School. Data from SSOCS:2000 as well as those collected elsewhere show that, while incidents of death at schools draw great attention and are a source of great concern for schools, they actually are relatively rare. In fact, during the 1999-2000 school year, not a single school in the SSOCS:2000 sample of 2,270 schools had a homicide on campus, and only one school reported a suicide (of a student) on campus (the statistics in this paragraph are unweighted). This survey result does not reflect the actual absence of deaths at schools, but rather the rarity of deaths, so that even a nationally representative sample of over 2,000 schools did not happen to include any.

Over the period July 1, 1997 through June 30, 1998, there were a total of 35 school-associated homicides of school-age children, 7 school-associated suicides of school-age children, 12 other school-associated homicides, and 5 other school-associated suicides (Indicators of School Crime and Safety, 2000, p. 2).[43] These statistics were collected from the School Associated Violent Deaths Study (SAVDS), which was a universe study based on mortality records. As such, the statistics are not subject to sampling error or reporting error by schools. Although they do not provide statistics on a school-level basis (e.g., the number of schools with deaths), the SAVDS provides the most accurate data on the number of school associated deaths.

Deaths Outside of School. SSOCS:2000 estimates show that some students and teachers died from violence outside of school in 1999-2000, though this also was relatively rare (figure 10-1). Only 1.6 percent of schools had any students who died from homicide or suicide outside of schools, and only 0.1 percent had any faculty or staff who died from homicide or suicide outside of schools. These findings suggest that deaths outside of schools are rare but more frequent than deaths at school. Deaths away from school were far more common than deaths that were school associated. Among school-age children (ages 5 though 19), there were 2,752 homicides according to the Federal Bureau of Investigation's Supplementary Homicide Reports, 1997–1998. The School Associated Violent Deaths Study indicates that 35 of those homicides were school related. Similarly, among this age group, statistics from the Centers for Disease Control and Prevention Vital Statistics of the U.S., 1997–1998 show there were 2,061 total suicides (with 7 being school associated) (Indicators of School Crime and Safety, 2000, p. 2). Methodological differences exist between SSOCS:2000 and SAVDS, including the presence of sampling error in SSOCS:2000 and the inability to present SAVDS data on a school-level basis.

[43] Kaufman, P. et al. (2000). *Indicators of School Crime and Safety, 2000* (NCES 2001-017). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Figure 10-1. Percentage of schools with homicides and suicides of students and faculty/staff outside of school: 2000



SOURCE: U.S. Department of Education, National Center for Education Statistics, School Survey on Crime and Safety (SSOCS), 2000.

Violent Incidents. SSOCS:2000 found a total of roughly 1.5 million incidents of violence in the 1999-2000 school year. (Violence is defined here as including rape or attempted rape, sexual battery, physical attack or fight, threats of physical attack, or robbery.) The National Crime Victimization Survey (which defines violent crime as rape, sexual assault, robbery, or assault) found 1.2 million violent crimes in 1997–98 (Indicators of School Crime and Safety, 2000, p. 4). Besides the difference in the definitions that were used, there were other important methodological differences that make comparisons of the two studies problematic. Some of these differences could result in NCVS producing lower estimates than SSOCS:2000, while some could result in higher estimates. The NCVS is a survey of students, and therefore can be expected to identify some crimes that were never reported to schools, and would not be included in school-maintained statistics. Other important differences between the two surveys are that the National Crime Victimization Survey includes crimes on the way to or from school (while SSOCS:2000 would include them only if they were on school-provided transportation). The NCVS results are for students of ages 12 through 18; therefore elementary school students are excluded

from the survey. But roughly half of the violent incidents reported in SSOCS:2000 were at elementary schools[44].

Incidents Reported to Police. The 1996–97 FRSS Principal/School Disciplinarian Survey on School Violence provides estimates of certain incidents that were reported to police (Heaviside, Rowand, Williams, and Farris, 1998).[45] The FRSS study was also a school-based study. The SSOCS:2000 and FRSS both provide estimates of rape or sexual battery, physical attack or fight (with or without a weapon), robbery, theft/larceny, and vandalism that occurred at school and were reported to the police.

Since the SSOCS:2000 and FRSS were conducted at different times, a comparison of the two studies may appear to provide measures of change over time. However, although SSOCS:2000 was partially modeled after the FRSS survey, there were many changes to the questionnaire. For this reason, the two cannot be compared.

In general, the SSOCS:2000 is a larger questionnaire with items located in different sections than on the FRSS. For example, in the questionnaire item that collects information on incidents reported to police, SSOCS:2000 uses different question wording and instructions to respondents. Another difference in the item is that SSOCS:2000 allows the respondent to report the total number of incidents and the total number of incidents reported to the police. For the FRSS, the respondent only reported the total number of incidents reported to police. The effect of the differences in wording, placement in the questionnaire, and additional information collected is unknown.

In addition to the general differences between the collection of incidents reported to the police for the SSOCS:2000 and FRSS, specific differences exist. In two cases, SSOCS:2000 items must be combined in order to be comparable to the FRSS survey. SSOCS:2000 distinguished between rape and sexual battery, while FRSS combined both together, and SSOCS:2000 distinguished between robbery with a weapon and robbery without a weapon, while FRSS collected a single statistic. Also, the statistics collected on theft/larceny may differ from those in FRSS because of a SSOCS:2000 limitation to items over $10 in value; however, small thefts are probably less likely to be reported to police, so this limitation may not be a big one.

---

[44] The Youth Risk Behavior Study (YRBS) also collects data on the number of students involved in fights (Kaufman, P., et al. Ibid., 148). However, the YRBS data are not directly comparable to the SSOCS:2000 data because they are counts of the number of students involved (some students might be involved in multiple fights), not the number of incidents.

[45] Heaviside, S., Rowand, C., Williams, C., and Farris, E. (1998). *Violence and Discipline Problems in U.S. Public Schools: 1996-1997* (NCES 98-030). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Disciplinary Actions. As with incidents reported to police, it may appear that the FRSS and SSOCS:2000 estimates of disciplinary actions for various incidents are comparable. However, there are many differences between the FRSS and SSOCS:2000 questionnaires with regard to the number of disciplinary actions taken by schools. The question wording was different on each survey. For example, SSOCS:2000 used removals with no continuing school services for at least 1 year while FRSS used expulsions, and SSOCS:2000 used transfers to specialized schools for disciplinary reasons for at least 1 year while FRSS used transfers to alternative schools or programs. The SSOCS:2000 reinterview study found that 79 percent of the sampled schools defined expulsion differently than it was defined in SSOCS:2000 (generally principals were more inclusive in their definitions of expulsion—including periods of less than a year or including students who were receiving services while expelled), so the FRSS estimates might tend to be higher based on these differences in definition.

Similar to the wording differences for disciplinary actions, SSOCS:2000 and FRSS differ in the wording of the incidents for which students were disciplined. SSOCS:2000 collected separate statistics for possession and use of firearms and other weapons, while FRSS combined possession and use. SSOCS:2000 did not ask about actions for the distribution of alcohol or the use of tobacco, while FRSS included both of these within a single larger category (possession, distribution, or use of alcohol or drugs, including tobacco).

APPENDIX A:

PROCEDURES FOR MINIMIZING OVERLAP BETWEEN NAEP/ECLS-K, SASS, FRSS, AND SSOCS:2000

PROCEDURES FOR MINIMIZING OVERLAP BETWEEN NAEP/ECLS-K, SASS, FRSS, AND SSOCS:2000

This technical appendix describes the general method used to minimize the overlap among the concurrent surveys: NAEP/ECLS-K, SASS, FRSS on teacher quality issues,[46] and SSOCS:2000. The method is an extension of the procedures that were used to minimize the overlap between the SASS school sample and the samples selected for NAEP and ECLS-K. Since the NAEP and ECLS-K samples are selected independently, they can be combined and treated as a single sample for the purpose of the sample minimization process.

The sample minimization algorithm involved the derivation of a set of "conditional" selection probabilities that were used to select the SSOCS:2000 sample. The term "conditional" probability refers to the probability of selecting a school after it has been given a chance of selection for one of the other studies, and depends on the outcomes of the previous sampling processes. The derivation of the conditional probabilities given below reflects the fact that the SASS sample was selected using a similar minimization strategy. As described in detail below, the application of the scheme required the relevant probabilities and conditional probabilities of selection of a school for each of the previous surveys.

Notation

Table A-1 sets the general notation, and table A-2 gives the notation for joint selection probabilities between two surveys.

Table A-1. General notation

Survey*	Sample	Probability of school *i* being	
		Selected	Not selected
(1) NAEP	s_1	$P_i(s_1)$	$P_i(\bar{s}_1) = 1 - P_i(s_1)$
(2) SASS	s_2	$P_i(s_2)$	$P_i(\bar{s}_2) = 1 - P_i(s_2)$
(3) FRSS	s_3	$P_i(s_3)$	$P_i(\bar{s}_3) = 1 - P_i(s_3)$
(4) SSOCS:2000	s_4	$P_i(s_4)$	$P_i(\bar{s}_4) = 1 - P_i(s_4)$

*For the purpose of minimizing overlap, the NAEP/ECLS-K samples were treated as a single sample.

[46] The Fast Response Survey System (FRSS) survey on teacher quality issues was conducted in early 2000.

Table A-2. Notation for joint selection probability

Probability of selecting *school i* for two surveys		Survey 2		Total
		Selected	Not selected	
Survey 1	Selected	$P_i(s_1 s_2)$	$P_i(s_1 \bar{s}_2)$	$P_i(s_1)$
	Not Selected	$P_i(\bar{s}_1 s_2)$	$P_i(\bar{s}_1 \bar{s}_2)$	$P_i(\bar{s}_1)$
	Total	$P_i(s_2)$	$P_i(\bar{s}_2)$	1

Similarly, in the case of three surveys, $P_i(s_1 s_2 s_3)$ is the probability that school i is selected for surveys 1, 2, and 3, whereas $P_i(s_1 s_2 \bar{s}_3)$ is the probability that school i is selected for survey 1 and survey 2, but not selected for survey 3. Similar notations will be used for the joint probabilities of selection for other combinations of surveys.

Objective

The objective was to derive the conditional probability of selecting school *i* for SSOCS:2000 (survey 4) in a way that (a) minimizes the overlap with NAEP/ECLS-K (survey 1), SASS (survey 2), and FRSS (survey 3) while (b) achieving the desired unconditional selection probability for SSOCS:2000. In other words, the required conditional probabilities (which were used to select the SSOCS:2000 sample) depended on whether or not the school was selected for one of the other studies as well as on the original selection probabilities. The "unconditional" selection probability, on the other hand, is simply the desired overall probability of selecting a school for SSOCS:2000 regardless of its selection status for the other studies. As shown later, the unconditional probabilities of selection for SSOCS:2000 equal the desired probabilities under the proposed sample design.

Derivation of Conditional Probabilities to Select SSOCS:2000 Sample

The conditional probability used to select a school for SSOCS:2000 depended on (a) the joint probabilities of selection for the previous surveys, (b) the desired (unconditional) probability of selection for SSOCS:2000, and (c) the selection status of the school (i.e., whether

or not it was selected for one or more of the other studies). The steps involved in deriving conditional probability for each school are described below.

Step 1: Compute prior and residual probabilities for each school

After selecting the samples for the first three surveys, four mutually exclusive and exhaustive groups can be formed as follows.

Group 1 (G_1): schools not selected for any of the previous surveys,
Group 2 (G_2): schools selected for only one of the three previous surveys,
Group 3 (G_3): schools selected for two of the three previous surveys,
Group 4 (G_4): schools selected for all three of the previous surveys.

To minimize overlap, the highest priority in selection will be given to the schools in group 1, then to the schools in group 2, then to the schools in group 3, and finally to the schools in group 4. The second column in table A-3 presents the inclusion probability of a school in each of these groups. These probabilities will be referred to as "prior" probabilities. The probabilities defined in the third column of table A-3 will be called "residual" probabilities. Ratios of these residual and prior probabilities in rows of table A-3 were used to derive conditional probabilities of selection for SSOCS:2000. Therefore, the first step was to calculate these four prior and four residual probabilities for each school in the frame.

Table A-3. Prior and residual probabilities in different priority groups

Group	Prior probability (i.e., the probability that a school would be included in the group)	Residual probability (i.e., desired probability – cumulative sum of prior probabilities)
G_1 (Group 1)	$P_i(G_1) = P_i(\bar{s}_1\bar{s}_2\bar{s}_3)$	$P_i(R_1) = P_i(s_4)$
G_2 (Group 2)	$P_i(G_2) = P_i(\bar{s}_1\bar{s}_2 s_3) + P_i(\bar{s}_1 s_2\bar{s}_3) + P_i(s_1\bar{s}_2\bar{s}_3)$	$P_i(R_2) = P_i(s_4) - P_i(G_1)$
G_3 (Group 3)	$P_i(G_3) = P_i(\bar{s}_1 s_2 s_3) + P_i(s_1\bar{s}_2 s_3) + P_i(s_1 s_2\bar{s}_3)$	$P_i(R_3) = P_i(s_4) - P_i(G_1) - P_i(G_2)$
G_4 (Group 4)	$P_i(G_4) = P_i(s_1 s_2 s_3)$	$P_i(R_4) = P_i(s_4) - P_i(G_1) - P_i(G_2) - P_i(G_3)$

Step 2: Determine inequality type for each school

Next, for every school in the frame, each of the four residual probabilities listed in table A-3 was compared with the corresponding prior probability. For a given pair of prior and residual probabilities, either the residual probability is less than the corresponding prior probability (YES) or it is not (NO). Thus, for example, if for a given school all four residual probabilities listed in table A-3 are less than the corresponding prior probability, then this can be denoted by the sequence YES, YES, YES, and YES. On the other hand, if the first residual probability is not less than the corresponding prior probability, but all of the remaining three residual probabilities are less than the corresponding prior probability, then this would be denoted by the sequence NO, YES, YES, and YES. It turns out that the four sequences listed in table A-4 (referred to as "inequality" types) are the only possible sequences due to the fact that once the residual probability is less than the prior probability for a given row of table A-3, the residual probability for all subsequent rows will also be less than the corresponding prior probability. Hence, the next step was to determine which of the four inequality types listed in table A-4 applied to each school in the sampling frame.

Table A-4. Possible combinations of the inequalities between prior and residual probabilities in different priority groups

Inequality type	Is residual probability less than or equal to prior probability in different groups in table A-3?			
	Group 1	Group 2	Group 3	Group 4
A	YES	YES	YES	YES
B	NO	YES	YES	YES
C	NO	NO	YES	YES
D	NO	NO	NO	YES

Step 3: Derive conditional probabilities of selection

Table A-5 presents the conditional probability of selection assigned to a school depending on the type of inequality it satisfied and the number of previous surveys it was included in. For example, if a school satisfied inequality type A and it was not included in any of the previous surveys then the conditional probability would be the ratio of the residual and prior probabilities given in table A-5. However, if the school was included in one or more of the previous surveys then the conditional probability would be zero. Similarly, if a school satisfied inequality type B then its conditional probability of selection would be 1 if it was not included in any of the previous surveys. On the other hand, the conditional probability of that school would be the ratio of the residual and prior probabilities given in table A-5 if it was included in only one of the previous surveys, and the conditional probability would be zero if it was included in 2 or 3 of the previous surveys.

To summarize, the steps involved in implementing the procedure for overlap minimization in SSOCS:2000 were as follows. First, the prior and residual probabilities in each priority group were computed using table A-3 for every school in the frame. Second, using table A-4 it was determined which inequality type the school satisfied. Third, the conditional probability of selection for SSOCS:2000 was computed from table A-5 depending on the inequality type and the number of previous surveys the school was included in. Finally, the derived conditional probabilities were used to select the sample.

Table A-5. Conditional probabilities of selection assigned to the *i* th school to minimize overlap of SSOCS with the previous three surveys: 2000

Inequality type	Number of surveys *i* th school was included in			
	None	1 survey	2 surveys	3 surveys
A	$\frac{P_i(R_1)}{P_i(G_1)}$	0	0	0
B	1	$\frac{P_i(R_2)}{P_i(G_2)}$	0	0
C	1	1	$\frac{P_i(R_3)}{P_i(G_3)}$	0
D	1	1	1	$\frac{P_i(R_4)}{P_i(G_4)}$

A proof of the unbiasedness of the overlap minimization procedure is beyond the scope of the detailed data documentation. However, this and further theoretical details on the derivation of the procedure can be found in Chowdhury, Chu, and Kaufman (2000).[47]

[47] Chowdhury, S., Chu, A., and Kaufman, S. (2000). Minimizing Overlap in NCES Surveys. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 174-179.