Appendix A: Technical Notes

Technical Notes

General Information

The indicators in this report are based on information drawn from a variety of independent data sources, including national surveys of students, teachers, and principals, and data collections from federal departments and agencies, including the Bureau of Justice Statistics, the National Center for Education Statistics, the Federal Bureau of Investigation, and the Centers for Disease Control and Prevention. Each data source has an independent sample design, data collection method, and questionnaire design or is the result of a universe data collection. Universe data collections include a census of all known entities in a specific universe (e.g., all deaths occurring on school property). Readers should be cautious when comparing data from different sources. Differences in sampling procedures, populations, time periods, and question phrasing can all affect the comparability of results. For example, some questions from different surveys may appear the same, but were asked of different populations of students (e.g., students ages 12–18 or students in grades 9–12); in different years; about experiences that occurred within different periods of time (e.g., in the past 30 days or during the past 12 months); or at different locations (e.g., in school or anywhere).

All comparisons described in this report are statistically significant at the .05 level. Estimates displayed in the text, figures, and tables are rounded from original estimates, not from a series of rounding.

The following is a description of data sources, accuracy of estimates, and statistical procedures used in this report.

Sources of Data

This section briefly describes each of the datasets used in this report: the School-Associated Violent Deaths Surveillance Study, the Supplementary Homicide Reports, the Web-based Injury Statistics Query and Reporting System Fatal, the National Crime Victimization Survey, the School Crime Supplement to the National Crime Victimization Survey, the Youth Risk Behavior Survey, the Schools and Staffing Survey, and the School Survey on Crime and Safety. Directions for obtaining more information are provided at the end of each description.

School-Associated Violent Deaths Surveillance Study (SAVD)

The School-Associated Violent Deaths Surveillance Study (SAVD) is an epidemiological study developed by the Centers for Disease Control and Prevention in conjunction with the U.S. Department of Education and the U.S. Department of Justice. SAVD seeks to describe the epidemiology of school-associated violent deaths, identify common features of these deaths, estimate the rate of school-associated violent death in the United States, and identify potential risk factors for these deaths. The surveillance system includes descriptive data on all school-associated violent deaths in the United States, including all homicides, suicides, and unintentional firearm-related deaths where the fatal injury occurred on the campus of a functioning elementary or secondary school, while the victim was on the way to or from regular sessions at such a school, or while attending or on the way to or from an official school-sponsored event. Victims of such events include nonstudents as well as students and staff members. SAVD includes descriptive information about the school, event, victim(s), and offender(s). The SAVD Surveillance System has collected data from July 1, 1992, through the present.

SAVD uses a four-step process to identify and collect data on school-associated violent deaths. Cases are initially identified through a search of the Lexis/Nexis newspaper and media database. Then police officials are contacted to confirm the details of the case and to determine if the event meets the case definition. Once a case is confirmed, a police official and a school official are interviewed regarding details about the school, event, victim(s), and offender(s). A copy of the full police report is also sought for each case. The information obtained on schools includes school demographics, attendance/absentee rates, suspension/expulsions and mobility, school history of weapon-carrying incidents, security measures, violence prevention activities, school response to the event, and school policies about weapon carrying. Event information includes the location of injury, the context of injury (while classes were being held, during break, etc.), motives for injury, method of injury, and school and community events happening around the time period. Information obtained on victim(s) and offender(s) includes demographics, circumstances of the event (date/time, alcohol or drug use, number of persons involved), types and origins of weapons, criminal history, psychological risk factors, school-related problems, extracurricular

activities, and family history, including structure and stressors.

One hundred and five school-associated violent deaths were identified from July 1, 1992 to June 30, 1994 (Kachur et al. 1996). A more recent report from this data collection identified 253 school-associated violent deaths between July 1, 1994 and June 30, 1999 (Anderson et al. 2001). Other publications from this study have described how the number of events changes during the school year (Centers for Disease Control and Prevention 2001), the source of the firearms used in these events (Reza et al. 2003), and suicides that were associated with schools (Kauffman et al. 2004). The interviews conducted on cases between July 1, 1994, and June 30, 1999, achieved a response rate of 97 percent for police officials and 78 percent for school officials. The SAVD data are considered preliminary until interviews with school and law enforcement officials have been completed. The details learned during the interviews can occasionally change the classification of a case. For additional information about SAVD, contact:

Jeff Hall
Division of Violence Prevention
National Center for Injury Prevention and Control
Centers for Disease Control and Prevention
Mailstop K60
4770 Buford Highway NE
Atlanta, GA 30341
Telephone: (770) 488-4648
E-mail: jhall2@cdc.gov

Supplementary Homicide Reports (SHR)

The Supplementary Homicide Reports (SHR), which are a part of the Uniform Crime Reporting (UCR) program, provide incident-level information on criminal homicides, including situation (number of victims to number of offenders); the age, sex, and race of victims and offenders; types of weapons used; circumstances of the incident; and the relationship of the victim to the offender. The data are provided monthly to the Federal Bureau of Investigation (FBI) by local law enforcement agencies participating in the FBI's UCR program. The data include murders and nonnegligent manslaughters in the United States from January 1976 to December 2005; that is, negligent manslaughters and justifiable homicides have been eliminated from the data. Based on law enforcement agency reports, the FBI estimates that 594,277 murders (including non-negligent manslaughters) were committed from 1976 to 2005. Agencies provided detailed information on 538,210 victims and 597,359 offenders.

About 91 percent of homicides are included in the SHR. However, adjustments can be made to the weights to correct for missing reports. Estimates from the SHR used in this report were generated by the Bureau of Justice Statistics (BJS) using a weight developed by BJS that reconciles the counts of SHR homicide victims with those in the UCR for the 1992 through 2005 data years. The weight is the same for all cases for a given year. The weight represents the ratio of the number of homicides reported in the UCR to the number reported in the SHR. For additional information about SHR, contact:

Communications Unit
Criminal Justice Information Services Division
Federal Bureau of Investigation
Module D3
1000 Custer Hollow Road
Clarksburg, WV 26306
Telephone: (304) 625-4995
E-mail: cjis_comm@leo.gov

Web-based Injury Statistics Query and Reporting System Fatal (WISQARS™ Fatal)

WISQARS Fatal provides mortality data related to injury. The mortality data reported in WISQARS Fatal come from death certificate data reported to the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention. Data include causes of death reported by attending physicians, medical examiners, and coroners. It also includes demographic information about decedents reported by funeral directors, who obtain that information from family members and other informants. NCHS collects, compiles, verifies, and prepares these data for release to the public. The data provide information about what types of injuries are leading causes of deaths, how common they are, and who they affect. These data are intended for a broad audience—the public, the media, public health practitioners and researchers, and public health officials—to increase their knowledge of injury.

WISQARS Fatal mortality reports provide tables of the total numbers of injury-related deaths and the death rates per 100,000 U.S. population. The reports list deaths according to cause (mechanism) and intent (manner) of injury by state, race, Hispanic origin, sex, and age groupings. For more information on WISQARS Fatal, contact:

National Center for Injury Prevention and Control
Mailstop K59
4770 Buford Highway NE
Atlanta, GA 30341-3724

Telephone: (770) 488-1506
E-mail: ohcinfo@cdc.gov
Internet: http://www.cdc.gov/ncipc/wisqars

National Crime Victimization Survey (NCVS)

The National Crime Victimization Survey (NCVS), administered for the U.S. Bureau of Justice Statistics by the U.S. Census Bureau, is the nation's primary source of information on crime and the victims of crime. Initiated in 1972 and redesigned in 1992, the NCVS collects detailed information annually on the frequency and nature of the crimes of rape, sexual assault, robbery, aggravated and simple assault, theft, household burglary, and motor vehicle theft experienced by Americans and their households each year. The survey measures both crimes reported to police and crimes not reported to the police.

Readers should note that in 2003, in accordance with changes to the Office of Management and Budget's standards for the classification of federal data on race and ethnicity, the NCVS item on race/ethnicity was modified. A question on Hispanic origin is followed by a question on race. The new question about race allows the respondent to choose more than one race and delineates Asian as a separate category from Native Hawaiian or Other Pacific Islander. Analysis conducted by the Demographic Surveys Division at the U.S. Census Bureau showed that the new question had very little impact on the aggregate racial distribution of the NCVS respondents, with one exception. There was a 1.6 percentage point decrease in the percentage of respondents who reported themselves as White. Due to changes in race/ethnicity categories, comparisons of race/ethnicity across years should be made with caution.

Due to changes in survey methodology in 2006 that mainly affected rural areas, 2006 estimates are not comparable to estimates based on NCVS data from previous years. Continuity between urban and suburban areas in the sample for both years (2005 and 2006) enabled year-to-year comparisons for these areas. The U.S. Census Bureau, the Bureau of Justice Statistics (BJS), and a panel of outside experts extensively reviewed the 2006 NCVS data and determined that there was a break in series between 2006 and previous years that prevented annual comparison of criminal victimization at the national level. This was mainly the result of three major changes in the survey methodology: (1) introducing a new sample to account for shifts in population and location of households that occur over time, (2) incorporating responses from households that were in the survey for the first time, and (3) using computer-assisted

personal interviewing (CAPI). For more information on the 2006 NCVS data, see *Criminal Victimization, 2006* at http://www.ojp.usdoj.gov/bjs/pub/pdf/cv06 .pdf and the technical notes at http://www.ojp.usdoj .gov/bjs/pub/pdf/cv06tn.pdf.

The number of NCVS eligible households in the sample in 2006 was about 41,800. They were selected using a stratified, multistage cluster design. In the first stage, the primary sampling units (PSUs), consisting of counties or groups of counties, were selected. In the second stage, smaller areas, called Enumeration Districts (EDs), were selected from each sampled PSU. Finally, from selected EDs, clusters of four households, called segments, were selected for interview. At each stage, the selection was done proportionate to population size in order to create a self-weighting sample. The final sample was augmented to account for households constructed after the decennial Census. Within each sampled household, U.S. Census Bureau personnel interviewed all household members age 12 and older to determine whether they had been victimized by the measured crimes during the 6 months preceding the interview.

The first NCVS interview with a housing unit is conducted in person. Subsequent interviews are conducted by telephone, if possible. About 68,000 persons age 12 and older are interviewed each 6 months. Households remain in the sample for 3 years and are interviewed seven times at 6-month intervals. Since the survey's inception, the initial interview at each sample unit has been used only to bound future interviews to establish a time frame to avoid duplication of crimes uncovered in these subsequent interviews. Beginning in 2006, data from the initial interview have been adjusted to account for the effects of bounding and are included in the survey estimates. After their seventh interview, households are replaced by new sample households. The NCVS has consistently obtained a response rate of over 90 percent at the household level. The completion rates for persons within households were about 86 percent. Thus, final response rates were about 78 percent in 2006. Weights were developed to permit estimates for the total U.S. population 12 years and older. For more information about the NCVS, contact:

Katrina Baum
Victimization Statistics Branch
Bureau of Justice Statistics
U.S. Department of Justice
810 7th Street NW
Washington, DC 20531
Telephone: (202) 307-5889

E-mail: Katrina.Baum@usdoj.gov
Internet: http://www.ojp.usdoj.gov/bjs

School Crime Supplement (SCS)

Created as a supplement to the NCVS and codesigned by the National Center for Education Statistics and Bureau of Justice Statistics, the School Crime Supplement (SCS) survey was conducted in 1989, 1995, 1999, 2001, 2003, 2005, and 2007 to collect additional information about school-related victimizations on a national level. This report includes data from the 1995, 1999, 2001, 2003, 2005, and 2007 collections. The 1989 data are not included in this report as a result of methodological changes to the NCVS and SCS. The survey was designed to assist policymakers as well as academic researchers and practitioners at the federal, state, and local levels so that they can make informed decisions concerning crime in schools. The SCS asks students a number of key questions about their experiences with and perceptions of crime and violence that occurred inside their school, on school grounds, on a school bus, or on the way to or from school. Additional questions not included in the NCVS were also added to the SCS, such as those concerning preventive measures used by the school, students' participation in after school activities, students' perceptions of school rules, the presence of weapons and gangs in school, the presence of hate-related words and graffiti in school, student reports of bullying and reports of rejection at school, and the availability of drugs and alcohol in school, as well as attitudinal questions relating to fear of victimization and avoidance behavior at school.

In all SCS survey years through 2005, the SCS was conducted for a 6-month period from January to June in all households selected for the NCVS (see discussion above for information about the NCVS sampling design and changes to the race/ethnicity item made for 2003 onward). It should be noted that the initial NCVS interview has always been included in the SCS data collection. Within these households, the eligible respondents for the SCS were those household members who had attended school at any time during the 6 months preceding the interview, were enrolled in grades 6–12, and were not home schooled. In 2007, the questionnaire was changed and household members who attended school any time during the school year were included. The age range of students covered in this report is 12–18 years of age. Eligible respondents were asked the supplemental questions in the SCS only after completing their entire NCVS interview.

The prevalence of victimization for 1995, 1999, 2001, 2003, 2005, and 2007 was calculated by using NCVS incident variables appended to the 1995, 1999, 2001, 2003, 2005, and 2007 SCS data files. The NCVS type of crime variable was used to classify victimizations of students in the SCS as serious violent, violent, or theft. The NCVS variables asking where the incident happened and what the victim was doing when it happened were used to ascertain whether the incident happened at school. For prevalence of victimization, the NCVS definition of "at school" includes in the school building, on school property, or on the way to or from school. Only incidents that occurred inside the United States are included.

In 2001, the SCS survey instrument was modified from previous collections. First, in 1995 and 1999, "at school" was defined for respondents as in the school building, on the school grounds, or on a school bus. In 2001, the definition for "at school" was changed to mean in the school building, on school property, on a school bus, or going to and from school. This change was made to the 2001 questionnaire in order to be consistent with the definition of "at school" as it is constructed in the NCVS and was also used as the definition in 2003, 2005, and 2007. Cognitive interviews conducted by the U.S. Census Bureau on the 1999 SCS suggested that modifications to the definition of "at school" would not have a substantial impact on the estimates.

A total of 9,728 students participated in the 1995 SCS, 8,398 in 1999, 8,374 in 2001, 7,152 in 2003, 6,297 in 2005, and 5,618 in 2007. In the 2007 SCS, the household completion rate was 90 percent. In the 1995, 1999, 2001, 2003, and 2005 SCS, the household completion rates were 95 percent, 94 percent, 93 percent, 92 percent, and 91 percent, respectively, and the student completion rates were 78 percent, 78 percent, 77 percent, 70 percent, and 62 percent respectively. For the 2007 SCS, the student completion rate was 58 percent.

Thus, the overall unweighted SCS response rate (calculated by multiplying the household completion rate by the student completion rate) was 74 percent in 1995, 73 percent in 1999, 72 percent in 2001, 64 percent in 2003, 56 percent in 2005, and 53 percent in 2007. Response rates for most survey items were high—typically over 95 percent of all eligible respondents. The weights were developed to compensate for differential probabilities of selection and nonresponse. The weighted data permit inferences about the eligible student population who were enrolled in schools in 1995, 1999, 2001, 2003, 2005, and 2007.

Due to the low unit response rate in 2005 and 2007, a unit nonresponse bias analysis was done. There are two types of nonresponse: unit and item nonresponse. Unit response rates indicate how many sampled units have completed interviews. Because interviews with students could only be completed after households had responded to the NCVS, the unit completion rate for the SCS reflects both the household interview completion rate and the student interview completion rate. Nonresponse can greatly affect the strength and application of survey data by leading to an increase in variance as a result of a reduction in the actual size of the sample and can produce bias if the nonrespondents have characteristics of interest that are different from the respondents. Furthermore, imputation, a common recourse to nonresponse, can lead to the risk of underestimating the sampling error if imputed data are treated as though they were observed data.

In order for response bias to occur, respondents must have different response rates and responses to particular survey variables. The magnitude of unit nonresponse bias is determined by the response rate and the differences between respondents and nonrespondents on key survey variables. Although the bias analysis cannot measure response bias since the SCS is a sample survey and it is not known how the population would have responded, the SCS sampling frame has four key student or school characteristic variables for which data is known for respondents and nonrespondents: sex, race/ethnicity, household income, and urbanicity, all of which are associated with student victimization. To the extent that there are differential responses by respondents in these groups, nonresponse bias is a concern.

In 2005, the analysis of unit nonresponse bias found evidence of bias for the race, household income, and urbanicity variables. White (non-Hispanic) and Other (non-Hispanic) respondents had higher response rates than Black (non-Hispanic) and Hispanic respondents. Respondents from households with an income of $35,000–$49,999 and $50,000 or more had higher response rates than those from households with incomes of less than $7,500, $7,500–$14,999, $15,000–$24,999 and $25,000–$34,999. Respondents who live in urban areas had lower response rates than those who live in rural or suburban areas. Although the extent of nonresponse bias cannot be determined, weighting adjustments, which corrected for differential response rates, should have reduced the problem.

In 2007, the analysis of unit nonresponse bias found evidence of bias by race, household income, and urbanicity variables. Hispanic respondents had lower response rates than other race/ethnicities. Respondents from households with an income of $25,000 or more had higher response rates than those from households with incomes of less than $7,500. Respondents who live in urban areas had lower response rates than those who live in rural areas. However, when responding students are compared to the eligible NCVS sample, there were no measurable differences between the responding students and the eligible students, suggesting the nonresponse bias has little impact on the overall estimates.

For more information about SCS, contact:

Kathryn A. Chandler
National Center for Education Statistics
1990 K Street NW
Washington, DC 20006
Telephone: (202) 502-7486
E-mail: kathryn.chandler@ed.gov
Internet: http://nces.ed.gov/programs/crime

Youth Risk Behavior Survey (YRBS)

The National School-Based Youth Risk Behavior Survey (YRBS) is one component of the Youth Risk Behavior Surveillance System (YRBSS), an epidemiological surveillance system developed by the Centers for Disease Control and Prevention (CDC) to monitor the prevalence of youth behaviors that most influence health.[1] The YRBS focuses on priority health-risk behaviors established during youth that result in the most significant mortality, morbidity, disability, and social problems during both youth and adulthood. This report uses 1993, 1995, 1997, 1999, 2001, 2003, 2005, and 2007 YRBS data.

The YRBS uses a three-stage cluster sampling design to produce a nationally representative sample of students in grades 9–12 in the United States. The target population consisted of all public and private school students in grades 9–12 in the 50 states and the District of Columbia. The first-stage sampling frame included selecting primary sampling units (PSUs) from strata formed on the basis of urbanization and the relative percentage of Black and Hispanic students in the PSU. These PSUs are either counties, subareas of large counties, or groups of smaller, adjacent counties. At the second stage, schools were selected with probability proportional to school enrollment size.

---

[1] For more information on the YRBSS methodology, see Brener et al. (2004).

The final stage of sampling consisted of randomly selecting, in each chosen school and in each of grades 9–12, one or two classrooms from either a required subject, such as English or social studies, or a required period, such as homeroom or second period. All students in selected classes were eligible to participate. Three strategies were used to oversample Black and Hispanic students: (1) larger sampling rates were used to select PSUs that are in high-Black and high-Hispanic strata; (2) a modified measure of size was used that increased the probability of selecting schools with a disproportionately high minority enrollment; and (3) two classes per grade, rather than one, were selected in schools with a high minority enrollment. Approximately 16,300, 10,900, 16,300, 15,300, 13,600, 15,200, 13,900, and 14,000 students participated in the 1993, 1995, 1997, 1999, 2001, 2003, 2005, and 2007 surveys, respectively.

The overall response rate was 70 percent for the 1993 survey, 60 percent for the 1995 survey, 69 percent for the 1997 survey, 66 percent for the 1999 survey, 63 percent for the 2001 survey, 67 percent for the 2003 survey, 67 percent for the 2005 survey, and 68 percent for the 2007 survey. NCES standards call for response rates of 85 percent or better for cross-sectional surveys, and bias analyses are required by NCES when that percentage is not achieved. For YRBS data, a full nonresponse bias analysis has not been done because the data necessary to do the analysis are not available. The weights were developed to adjust for nonresponse and the oversampling of Black and Hispanic students in the sample. The final weights were constructed so that only weighted proportions of students (not weighted counts of students) in each grade matched national population projections.

State level data were downloaded from the Youth Online: Comprehensive Results web page (http://apps.nccd.cdc.gov/yrbss/). Each state and local school-based YRBS employs a two-stage, cluster sample design to produce representative samples of students in grades 9–12 in their jurisdiction. All except a few state and local samples include only public schools, and each local sample includes only schools in the funded school district (e.g., San Diego Unified School District) rather than in the entire city (e.g., greater San Diego area).

In the first sampling stage in all except a few states and districts, schools are selected with probability proportional to school enrollment size. In the second sampling stage, intact classes of a required subject or intact classes during a required period (e.g., second period) are selected randomly. All students in sampled classes are eligible to participate. Certain states and districts modify these procedures to meet their individual needs. For example, in a given state or district, all schools, rather than a sample of schools, might be selected to participate. State and local surveys that have a scientifically selected sample, appropriate documentation, and an overall response rate greater than or equal to 60 percent are weighted. The overall response rate reflects the school response rate multiplied by the student response rate. These three criteria are used to ensure that the data from those surveys can be considered representative of students in grades 9–12 in that jurisdiction. A weight is applied to each record to adjust for student nonresponse and the distribution of students by grade, sex, and race/ethnicity in each jurisdiction. Therefore, weighted estimates are representative of all students in grades 9–12 attending schools in each jurisdiction. Surveys that do not have an overall response rate of greater than or equal to 60 percent and do not have appropriate documentation are not weighted and are not included in this report.

In 2007, a total of 39 states and 22 districts had weighted data. In sites with weighted data, the student sample sizes for the state and local YRBS ranged from 1,118 to 13,439. School response rates ranged from 69 to 100 percent, student response rates ranged from 60 to 92 percent, and overall response rates ranged from 60 to 90 percent.

Readers should note that reports of these data published by the CDC do not include percentages where the denominator includes less than 100 unweighted cases. However, NCES publications do not include percentages where the denominator includes less than 30 unweighted cases. Therefore, estimates presented here may not appear in CDC publications of YRBS estimates and are considered unstable by CDC standards.

In 1999, in accordance with changes to the Office of Management and Budget's standards for the classification of federal data on race and ethnicity, the YRBS item on race/ethnicity was modified. The version of the race and ethnicity question used in 1993, 1995, and 1997 was:

How do you describe yourself?
A. White—not Hispanic
B. Black—not Hispanic
C. Hispanic or Latino
D. Asian or Pacific Islander
E. American Indian or Alaskan Native
F. Other

The version used in 1999, 2001, 2003, 2005, and in the 2007 state and local surveys was:

How do you describe yourself? (Select one or more responses.)
A. American Indian or Alaska Native
B. Asian
C. Black or African American
D. Hispanic or Latino
E. Native Hawaiian or Other Pacific Islander
F. White

In the 2005 national survey and all 2007 surveys, race/ethnicity was computed from two questions: (1) "Are you Hispanic or Latino?" (response options were "yes" and "no"), and (2) "What is your race?" (response options were "American Indian or Alaska Native," "Asian," "Black or African American," "Native Hawaiian or Other Pacific Islander," or "White"). For the second question, students could select more than one response option. For this report, students were classified as "Hispanic" if they answered "yes" to the first question, regardless of how they answered the second question. Students who answered "no" to the first question and selected more than one race/ethnicity in the second category were classified as "More than one race." Students who answered "no" to the first question and selected only one race/ethnicity were classified as that race/ethnicity. Race/ethnicity was classified as missing for students who did not answer the first question and for students who answered "no" to the first question but did not answer the second question.

CDC has conducted two studies to understand the effect of changing the race/ethnicity item on the YRBS. Brener, Kann, and McManus (2003) found that allowing students to select more than one response to a single race/ethnicity question on the YRBS had only a minimal effect on reported race/ethnicity among high school students. Eaton, Brener, Kann, and Pittman (2007) found that self-reported race/ethnicity was similar regardless of whether the single-question or a two-question format was used.

For additional information about the YRBS, contact:

Laura Kann
Division of Adolescent and School Health
National Center for Chronic Disease Prevention and Health Promotion

Centers for Disease Control and Prevention
Mailstop K-33
4770 Buford Highway NE

Atlanta, GA 30341-3717
Telephone: (770) 488-6181
E-mail: lkk1@cdc.gov
Internet: http://www.cdc.gov/yrbs

Schools and Staffing Survey (SASS)

This report draws upon data on teacher victimization from the Schools and Staffing Survey (SASS), which provides national- and state-level data on public schools and national- and affiliation-level data on private schools. The 1993–94, 1999–2000, and 2003–04 SASS were collected by the U.S. Census Bureau and sponsored by the National Center for Education Statistics (NCES). SASS consists of four sets of linked surveys, including surveys of schools, the principals of each selected school, a subsample of teachers within each school, and public school districts. In 1993–94, there were two sets of teacher surveys, public and private school teachers. In 1999–2000, there were four sets of teacher surveys, public, private, public charter, and Bureau of Indian Affairs (BIA) school teachers. In 2003–04, there were three sets of teacher surveys, public (including public charter), private, and BIA. For this report, BIA and public charter schools are included with public schools.

The public school sampling frames for the 1993–94, 1999–2000, and 2003–04 SASS were created using the 1991–92, 1997–98, and 2001–02 NCES Common Core of Data (CCD) Public School Universe Files, respectively. In SASS, a school was defined as an institution or part of an institution that provides classroom instruction to students; has one or more teachers to provide instruction; serves students in one or more of grades 1–12 or the ungraded equivalent and is located in one or more buildings. It was possible for two or more schools to share the same building; in this case they were treated as different schools if they had different administrations (i.e., principals). Since CCD and SASS differ in scope and their definition of a school, some records were deleted, added, or modified in order to provide better coverage and a more efficient sample design for SASS. Data were collected by multistage sampling, which began with the selection of schools.

This report uses 1993–94, 1999–2000, and 2003–04 SASS data. Approximately 10,000 public schools and 3,300 private schools were selected to participate in the 1993–94 SASS, 11,100 public schools (9,900 public schools, 100 BIA-funded schools, and 1,100 charter schools) and 3,600 private schools were selected to participate in the 1999–2000 SASS, and 10,400 public schools (10,200 public schools and

200 BIA-funded schools) and 3,600 private schools were selected to participate in the 2003–04 SASS. Within each school, teachers selected were further stratified into one of five teacher types in the following hierarchy: (1) Asian or Pacific Islander; (2) American Indian, Aleut, or Eskimo; (3) teachers who teach classes designed for students with limited English proficiency; (4) teachers in their first, second, or third year of teaching; and (5) teachers not classified in any of the other groups. Within each teacher stratum, teachers were selected systematically with equal probability. In 1993–94, approximately 57,000 public school teachers and 11,500 private school teachers were sampled. In 1999–2000, 56,300 public school teachers, 500 BIA teachers, 4,400 public charter school teachers, and 10,800 private school teachers were sampled. In 2003–04, 52,500 public school teachers, 700 BIA teachers, and 10,000 private school teachers were sampled.

This report focuses on responses from teachers. The overall weighted response rate for public school teachers in 1993–94 was 88 percent. In 1999–2000, the overall weighted response rates were 77 percent for public school teachers, and 86 and 72 percent for BIA and public charter school teachers, respectively (which are included with public school teachers for this report). In 2003–04, the overall weighted response rates were 76 percent for public school teachers and 86 percent for BIA-funded school teachers (who are included with public school teachers). For private school teachers, the overall weighted response rates were 80 percent, 67 percent, and 70 percent in 1993–94, 1999–2000, and 2003–04, respectively. Values were imputed for questionnaire items that should have been answered but were not. For additional information about SASS, contact:

Kerry Gruber
National Center for Education Statistics
1990 K Street NW
Washington, DC 20006
Telephone: (202) 502-7349
E-mail: kerry.gruber@ed.gov
Internet: http://nces.ed.gov/surveys/sass

School Survey on Crime and Safety (SSOCS)

The School Survey on Crime and Safety (SSOCS) was conducted by NCES in the spring of the 2005–06 school year. SSOCS focuses on incidents of specific crimes and offenses and a variety of specific discipline issues in public schools. It also covers characteristics of school policies, school violence prevention programs and policies, and school characteristics that have been associated with school crime. The survey was conducted with a nationally representative sample of regular public primary, middle, high, and combined schools in the 50 states and the District of Columbia.

The sampling frame for the 2005–06 SSOCS was constructed from the 2003–04 NCES Common Core of Data (CCD) Public Elementary/Secondary School Universe File. The CCD is an annual national database of all public K–12 schools and school districts. Certain types of schools were excluded from the CCD Public School Universe File in order to meet the sampling needs of SSOCS, including those in the outlying U.S. areas[2] and Puerto Rico; overseas Department of Defense schools; newly closed schools, home schools, and schools with high grades of kindergarten or lower; special education, vocational, or alternative schools; schools sponsored by the Bureau of Indian Affairs; ungraded schools; and "intermediate units"[3] in California and Pennsylvania. The sample was stratified by instructional level, type of locale (e.g., city, urban fringe, etc.), and enrollment size. Within the primary strata, schools were also sorted by geographic region and by percentage of minority enrollment. In order to obtain a reasonable sample size of lower enrollment schools while giving a higher probability of selection to higher enrollment schools, the sample was allocated to each subgroup in proportion to the sum of the square roots of the total student enrollment in each school in that stratum. A total of 3,565 schools were selected for the study. In March 2006, questionnaires were mailed to school principals, who were asked to complete the survey or to have it completed by the person most knowledgeable about discipline issues at the school. A total of 2,724 schools completed the survey. The weighted overall response rate was 80.6 percent, and weighted item nonresponse rates ranged from 0.0 to 27.7 percent. A nonresponse bias analysis was conducted on the 13 items with weighted item nonresponse rates above 15 percent, and the detected bias was not deemed problematic enough to suppress any items from the data file. Weights were developed to adjust for the variable probabilities of selection and differential nonresponse and can be used to produce national estimates for regular public schools in the 2005–06 school year. For information on the 1999–2000 and 2003–04 iterations, see *Indicators of School Crime and Safety: 2006*. For more information about the School Survey on Crime and

---

[2] "U.S. outlying areas" include the following: America Samoa, Guam, Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands.
[3] These are generally schools specializing in special education, alternative education, or juvenile halls.

Safety, contact:

Kathryn A. Chandler
National Center for Education Statistics
1990 K Street NW
Washington, DC 20006
Telephone: (202) 502-7486
E-mail: kathryn.chandler@ed.gov
Internet: http://nces.ed.gov/surveys/ssocs

Accuracy of Estimates

The accuracy of any statistic is determined by the joint effects of nonsampling and sampling errors. Both types of error affect the estimates presented in this report. Several sources can contribute to nonsampling errors. For example, members of the population of interest are inadvertently excluded from the sampling frame; sampled members refuse to answer some of the survey questions (item nonresponse) or all of the survey questions (questionnaire nonresponse); mistakes are made during data editing, coding, or entry; the responses that respondents provide differ from the "true" responses; or measurement instruments such as tests or questionnaires fail to measure the characteristics they are intended to measure. Although nonsampling errors due to questionnaire and item nonresponse can be reduced somewhat by the adjustment of sample weights and imputation procedures, correcting nonsampling errors or gauging the effects of these errors is usually difficult.

Sampling errors occur because observations are made on samples rather than on entire populations. Surveys of population universes are not subject to sampling errors. Estimates based on a sample will differ somewhat from those that would have been obtained by a complete census of the relevant population using the same survey instruments, instructions, and procedures. The standard error of a statistic is a measure of the variation due to sampling; it indicates the precision of the statistic obtained in a particular sample. In addition, the standard errors for two sample statistics can be used to estimate the precision of the difference between the two statistics and to help determine whether the difference based on the sample is large enough so that it represents the population difference.

Most of the data used in this report were obtained from complex sampling designs rather than a simple random design. The features of complex sampling require different techniques to calculate standard errors than are used for data collected using a simple random sampling. Therefore, calculation of standard errors requires procedures that are markedly different from the ones used when the data are from a simple random sample. The Taylor series approximation technique or the balanced repeated replication (BRR) method was used to estimate most of the statistics and their standard errors in this report.

Standard error calculation for data from the School Crime Supplement was based on the Taylor series approximation method using PSU and strata variables available from each dataset. For statistics based on all years of NCVS data, standard errors were derived from a formula developed by the U.S. Census Bureau, which consists of three generalized variance function (gvf) constant parameters that represent the curve fitted to the individual standard errors calculated using the Jackknife Repeated Replication technique.

The coefficient of variation (C_V) represents the ratio of the standard error to the mean. As an attribute of a distribution, the C_V is an important measure of the reliability and accuracy of an estimate. In this report, the C_V was calculated for all estimates, and in cases where the C_V was at least 30 percent the estimates were noted with a ! symbol (interpret data with caution). In cases where the C_V was greater than 50 percent, the estimate was determined not to meet reporting standards and was suppressed.

Statistical Procedures

The comparisons in the text have been tested for statistical significance to ensure that the differences are larger than might be expected due to sampling variation. Unless otherwise noted, all statements cited in the report are statistically significant at the .05 level. Several test procedures were used, depending upon the type of data being analyzed and the nature of the statement being tested. The primary test procedure used in this report was Student's t statistic, which tests the difference between two sample estimates, for example, between males and females. The formula used to compute the t statistic is as follows:

$$t = \frac{E_1 - E_2}{\sqrt{se_1^2 + se_2^2}} \qquad (1)$$

where E_1 and E_2 are the estimates to be compared and se_1 and se_2 are their corresponding standard errors. Note that this formula is valid only for independent estimates. When the estimates are not independent (for example, when comparing a total percentage with that for a subgroup included in the total), a covariance term (i.e., $2 * r * se_1 * se_2$) must be subtracted from the denominator of the formula:

$$t = \frac{E_1 - E_2}{\sqrt{se_1^2 + se_2^2 - (2 * r * se_1 * se_2)}} \qquad (2)$$

where r is the correlation coefficient. Once the t value was computed, it was compared with the published tables of values at certain critical levels, called alpha levels. For this report, an alpha value of .05 was used, which has a t value of 1.96. If the t value was larger than 1.96, then the difference between the two estimates is statistically significant at the 95 percent level.

A linear trend test was used when differences among percentages were examined relative to ordered categories of a variable, rather than the differences between two discrete categories. This test allows one to examine whether, for example, the percentage of students using drugs increased (or decreased) over time or whether the percentage of students who reported being physically attacked in school increased (or decreased) with their age. Based on a regression with, for example, student's age as the independent variable and whether a student was physically attacked as the dependent variable, the test involves computing the regression coefficient (b) and its corresponding standard error (se). The ratio of these two (b/se) is the test statistic t. If t is greater than 1.96, the critical value for one comparison at the .05 alpha level, the hypothesis that there is a linear relationship between student's age and being physically attacked is not rejected.

Some comparisons among categories of an ordered variable with three or more levels involved a test for a linear trend across all categories, rather than a series of tests between pairs of categories. In this report, when differences among percentages were examined relative to a variable with ordered categories, analysis of variance (ANOVA) was used to test for a linear relationship between the two variables. To do this, ANOVA models included orthogonal linear contrasts corresponding to successive levels of the independent variable. The squares of the Taylorized standard errors (that is, standard errors that were calculated by the Taylor series method), the variance between the means, and the unweighted sample sizes were used to partition the total sum of squares into within- and between-group sums of squares. These were used to create mean squares for the within- and between-group variance components and their corresponding F statistics, which were then compared with published values of F for a significance level of .05. Significant values of both the overall F and the F associated with the linear contrast term were required as evidence of a linear relationship between the two variables.

This page intentionally left blank.