

CHAPTER 7

SUMMARY

7.1 Introduction

This report reflects our current knowledge of the quality of SASS data. Direct, quantitative measures are available for some components of error, including sampling error, simple response variance (from reinterviews) and, for a few topics, reporting bias (from the Teacher Transcript Study). Only indirect measures are available for some other components of error. Unit and item nonresponse rates identify specific subpopulations and data items that are most likely to be affected by nonresponse bias. Comparisons among SASS and other surveys provide indications of the extent to which survey estimates may have failed to include some members of the target populations of school districts, schools, school administrators and teachers. As described later in this chapter, experimental and evaluation studies now underway will provide additional information.

It is not feasible to combine all of these indications of quality in some way to develop precise point estimates of total survey error for individual data items. Nevertheless, we believe that the information on quality that is now available or is being developed will help users decide how much confidence to place in the estimates that are of interest to them and determine how best to use these data in their analyses. The Quality Profile has been developed primarily as a convenient source of information about quality for users of SASS data.

We believe that the *Quality Profile*, with periodic updates, will also serve as a useful guide to the survey designers and managers at NCES and the Census Bureau in the effective allocation of additional resources to their continuing efforts to improve the quality of SASS data. There were many significant changes in the design, procedures and instrumentation used for the SASS surveys between the first and second rounds, and additional changes have been introduced in the third round, covering school year 1993-94. Many of these changes are designed to improve the quality of SASS data, as disclosed by the direct and indirect measures and indicators of quality presented in this report.

Section 7.2 of this chapter summarizes the information about different sources and components of error that has been presented in Chapters 2 through 5 for the four basic surveys and Chapter 6 for the Teacher Followup Survey. Section 7.3 describes additional research and evaluation activities that were underway at the time this report was being prepared. Section 7.4 presents some suggestions for data users on how to make effective use of SASS data products. Special attention is given to the possible effects of procedural and design changes on analyses of change between Rounds 1 and 2.

7.2 Principal sources of error

Coverage error The units of analysis for SASS are schools, school administrators, public school districts (LEAs) and teachers. Coverage errors can occur when units in the survey

target populations are omitted from sampling frames, when they appear more than once in the frames (if the duplication is not discovered prior to release of the estimates) or, in the case of schools and school districts, when a single unit in the target population is treated as more than one unit in the sampling frame, or vice versa. Coverage errors for schools can have a direct effect on the coverage of teachers and school administrators and, for public schools, may also affect the coverage of LEAs.

The target population for the School Survey was redefined between Rounds 1 and 2, in conjunction with the change in the frame used for selection of the public school sample. The Quality Education Data (QED) list used in Round 1 defined schools primarily in terms of physical location, whereas the Common Core of Data (CCD) list used in Round 2 defines them in terms of administrative units, so that it is possible to have more than one school at a single physical location or a single school with more than one location.

There are no direct estimates of gross or net coverage errors available for any of the SASS surveys. Comparisons of SASS estimates with data from other NCES surveys provide some indications of possible coverage error. As described in Section 2.5, SASS Round 2 estimates of the number of public schools by state were compared with the CCD counts. For the total U.S., the SASS estimate was 97.9 percent of the CCD count for the same school year. For most states the SASS estimates were within 5 percent of the CCD counts. Other things being equal, one would expect the CCD counts of public schools to be somewhat higher than the SASS estimate for the same year because the list frame for the School Survey was based on the CCD for the second year preceding the reference year, and therefore did not include all schools that started operation in the reference year or the preceding year. However, other factors, such as inclusion of out of scope units in the CCD may have operated to cause differences in the other direction. Similar considerations apply to the list frame used for private schools.

As described in the four chapters covering the base-year surveys, there are several indications, some of them quantitative, of potential coverage error. These include:

- The use, for both the public and private school surveys, of list frames constructed two years prior to the reference school year for the survey (Section 2.2).
- The need to use an area sample to supplement the list frame for private schools. The area sample accounted for about 22 percent of the estimated number of private schools in Round 1 and about 21 percent in Round 2, indicating no significant improvement of coverage by the list frame in Round 2 (Section 2.2).
- In Round 2, it was discovered that some multi-site special education programs of the State of California had been listed on the CCD as single schools. Adjustments were necessary to eliminate duplication for those sites located at existing schools and to select a sample of the other sites (Section 2.2).

- Discovery in both rounds, subsequent to sample selection, of some duplicate listings in the private school list frame. (Section 2.2).
- In Round 1, exclusion from the QED frame of 275 small Nebraska LEAs with about 2,800 students (Section 2.5).
- For the teacher surveys, use of teacher listing forms that ask only for teachers working at the sample schools at the time the forms were being completed. Teachers who begin working later in the reference year have no chance of inclusion (Section 5.2).
- In both rounds, counts of teachers on the teacher listing forms were, on the average, lower than the counts reported for the same schools on their School Survey questionnaires (Sections 5.2 and 5.4).

Sample estimates of the number of schools were also affected in both rounds by school survey respondents who provided data for a unit other than the one intended on the basis of the sample selection. Some respondents reported combined data for two different schools at the same location, and some, especially in small LEAs, reported combined data for all schools in the LEA. Conversely, in the Teacher Demand and Shortage Survey, a few LEAs reported data for a single school rather than the entire LEA. Many of these erroneous reports were identified and corrected prior to data release, but some may have escaped detection (see Sections 2.5, "Prepublication checks" and 4.3, "Measurement errors associated with data collection").

Nonresponse error Unit and item response rates for each of the five surveys have been presented in Chapters 2 to 6. To permit comparisons among the surveys, Table 7.1 presents response rates for all five surveys by round, separately for the public and private school sectors. Response rates for public schools have consistently exceeded those for private schools. Response rates in Round 2 exceeded those for Round 1 for each of the four basic surveys and the Teacher Followup Survey, for both sectors. As mentioned in Chapter 2, these increases may have resulted in part from the more lenient criteria used in Round 2 in accepting questionnaires that had missing or incomplete responses for some items. Response rates for the Teacher and Teacher Followup Surveys are composite rates, reflecting losses from schools that did not supply teacher lists and nonresponding teachers from schools that did supply lists. Consequently these rates were, with one exception, lower than those for the other three surveys.

As shown in additional tables in Chapters 2 to 6, there was considerable variation in response rates within each sector. For the public school sector in Round 1, in each of the four basic surveys a few states had response rates of less than 80 percent. This was due in part to a small number of LEAs, some of them fairly large, that declined to participate in any of the surveys (Section 2.3). In Round 2, all states had response rates above 80 percent in the School Survey and the School Administrator Survey (see Tables 2.6, 3.2, 4.2 and 5.3). For

the private school sector, one or more association groups had response rates of less than 60 percent in each of the four basic surveys in Round 1 (see Tables 2.7a, 3.3a, 4.3 and 5.4). Comparable data for the Teacher Survey are not available for Round 2.

Adjustments for unit nonresponse are included in the estimation weights for all of the surveys; however, the success of such adjustments in reducing bias depends on the extent to which the characteristics of units that respond and do not respond are similar. Survey results for domains of analysis with low response rates should be interpreted with caution. A new analysis of the characteristics of nonrespondents and the possible effects of unit nonresponse will be released soon (Scheuren, Parke and Bureika, 1994).

Data on item response rates for the five surveys were presented in Tables 2.9, 3.5, 4.5, 5.6 and 6.2. These rates cannot be compared across surveys in a meaningful way because of differences in content. There appears to have been some reduction in item nonresponse between Rounds 1 and 2 for the School, School Administrator, Teacher Demand and Shortage (public school sector) and Teacher Followup Surveys. This improvement probably resulted in part from dropping items that proved especially difficult for respondents in Round 1. Changes in questionnaire format may also have contributed to the reduction in item nonresponse. For the Teacher Survey, in contrast, item nonresponse rates were somewhat higher in Round 2. With one exception, the items flagged as having low response rates in Round 2 were also used (some with slightly different formats) in Round 1 of the Teacher Survey, so there is no obvious explanation for the lower response rates observed for these items in Round 2.

Examination of the questionnaire items with the lowest response rates in each survey and round suggests that factors associated with item nonresponse include question format, respondent burden and sensitivity. "None" boxes were frequently overlooked, as were spaces for entering amounts associated with positive responses to a yes or no question, especially when the amount spaces were located well to the right of the yes and no boxes.

Responses were frequently incomplete for complex "matrix-style" items, an example being an item on staffing patterns in the Round 1 School Survey questionnaires. This item called for counts of teachers by subject and by their status in the reference and prior school years. Data for this item and a similar one in the Teacher Demand and Shortage Survey questionnaire were judged to be of such poor quality that they were not included on the public and restricted use data files for those surveys. Because of the high nonresponse and other reporting problems, questions on this topic were substantially simplified for Round 2.

Items for which low response may have been associated with item sensitivity included several from the Round 2 Teacher Survey relating to amounts of teacher income from sources other than teaching. Another one was an item from the Round 1 Teacher Followup Survey that asked teachers who had moved to a new school for that school's religious affiliation, if any.

In Round 1, most missing or inconsistent items were imputed for the School and the Teacher

Demand and Shortage Survey questionnaires, but there was only limited imputation for the other three surveys. In Round 2, missing items were imputed for all surveys. All items changed in the computer imputation phase of data processing (but not those changed in preceding operations) were flagged as being imputed on the public and restricted-use data tapes. Some missing items can be deduced with a fairly high degree of accuracy from other responses on the same questionnaire. Other items are imputed by the "hot-deck" method, which assumes the values of missing items to be similar to those reported by other units that have the same basic attributes. In Round 1, for the School and the Teacher Demand and Shortage Surveys, a single code was used to flag imputed items. In Round 2, separate codes were used to distinguish internal and donor-based (hot deck) imputations.

Measurement error Information about measurement (response) errors associated with data collection comes from several sources: reinterviews, a record-check study, in-depth interviews using cognitive research techniques, methodological experiments, reviews of completed questionnaires and analyses of errors and inconsistencies detected during data processing. Information for each survey from these sources is presented in Chapters 2 through 6. Here we summarize the main findings.

- Reinterviews have shown that the items asking for the opinions, perceptions and future expectations of teachers and school administrators are, almost without exception, subject to high response variability. Moderate reductions in variability can be achieved by combining responses to 4-point scales into two categories (Sections 3.3, 5.3 and 6.3).
- Evidence from several sources suggests that the quality of information obtained by mail is superior to that obtained in telephone followups to nonrespondents. There are several possible reasons for this: questionnaires were not explicitly designed for use in telephone interviews; some of the questions can be answered more accurately by referring to records, which is harder to do in a telephone interview; persons who do not respond by mail are less likely to be motivated to provide accurate information; and there were indications that the training and supervision for the telephone followup interviews could have been improved. In the Round 2 School Survey, about one-third of the public school questionnaires and nearly one-half of the private school questionnaires were completed by telephone.
- An experiment, the State Data Project, was undertaken in connection with the Pretest for Round 2 of SASS to test the feasibility of obtaining data for the public sector Teacher Demand and Shortage Survey from state rather than local education agencies. A comparison of data collected from both sources for the same sample of LEAs showed a high frequency of substantial differences (more than 10 percent in either direction) for several variables. The experiment did not include any means of determining which of the two sources had provided more accurate data. Based on these findings, it was decided not to try to collect data from state agencies in Round 2 (see Section 4.3, "The State Data Project").

- Some of the concepts adopted for SASS data collection appear to be unfamiliar to respondents and to cause them considerable difficulty in formulating appropriate responses. One such concept is that of full-time equivalent (FTE) teachers used in the School and the Teacher Demand and Shortage Surveys. A school that has part-time teachers should report numbers of FTE teachers that are lower than their teacher counts. Nevertheless, many such schools reported the same numbers for teacher counts and FTE teachers (see Section 2.5, "Evaluation of published estimates: Round 1").
- A record-check study, the Teacher Transcript Study, compared teachers' self-reports of their educational backgrounds with data from their college transcripts. The main conclusion was that self-reports of types and years of degrees earned and major fields of study were reasonably accurate, but that self-reported information on courses and credit hours in specific fields was less accurate (see Section 5.3, "Measurement error: the Teacher Transcript Study").
- For all surveys and in both rounds of SASS, it was common for respondents to ignore skip instructions and consequently to try to answer questions that did not apply to them. Such errors have little or no direct effect on the quality of data, because most inapplicable responses can readily be deleted in clerical and computer edits. However, it is possible that frustration induced by trying to respond to irrelevant items may lead to a falling off, as respondents proceed through a questionnaire, in their level of commitment to providing complete and accurate information.

The foregoing and other findings relating to measurement error led to numerous changes in survey instruments and procedures between Rounds 1 and 2, and additional changes were made for Round 3. Some of the questions identified as being especially difficult have been eliminated, reduced in scope or modified. For a few items which were found to have especially severe nonresponse or other reporting problems, no estimates were published and individual responses were eliminated from the public-use data tapes. Included in this category were: item 28 on the Round 1 Teacher Questionnaire, which asked for information about hours spent on school-related activities; items 9 and 10 on the Round 1 Teacher Demand and Shortage Questionnaire, which asked for detailed information about full-time equivalent teaching positions, by specialty; and item 32 (35 for private schools) on the Round 1 School Survey Questionnaire, which asked for teachers by primary field of assignment in the current and preceding school year.

Data processing and estimation error In contrast to the sources of error discussed up to this point, there is relatively little documentation of processing and estimation errors. There are three clerical operations, each with the potential for both resolution of errors from earlier processing stages and introduction of new errors: clerical review of incoming questionnaires (in the Teacher Survey, this includes the coding of some industry entries), data entry, and resolution of rejects from the computer pre-edit operation. The first and third of these operations include some telephone contacts with respondents to obtain missing items or

resolve inconsistencies. Data entry is verified for 33 percent of the Teacher Survey questionnaires and for 100 percent of the questionnaires in the other three base-year surveys.

There have been no formal studies or experiments with alternative imputation procedures for the SASS surveys. Information relevant to the weighting procedures used in the School Survey was obtained in a study by Shen, Parmer and Tan (1992). After examining the correlates of nonresponse in that survey, they recommended some changes in the definition of the nonresponse adjustment cells for the private school sector and in the order of collapsing small cells for both public and private schools (see Section 2.4, Weighting). Most of the recommended changes were made in defining the nonresponse adjustment cells in Round 3.

As noted in Section 5.4, a new component was introduced into the weights used for estimates from Round 2 of the Teacher Survey. The purpose of this component, called the teacher adjustment factor, was to force agreement between teacher counts from the School and the Teacher Surveys. The differences prior to adjustment apparently are due primarily to the fact that schools do not include, on the teacher lists used for sampling, all of the teachers they include in their counts in the School Survey. If the teachers not included on the lists differ in some respects from those who are included, use of the adjustment factor will not necessarily eliminate all biases resulting from their not being represented in the sample.

Sampling error At present, there are two ways for users of SASS data to determine the sampling errors of estimates that are of interest to them. Users of data from publications will find that standard errors are provided for many of the published estimates. Users of microdata files can compute standard errors for any estimate by employing readily available software for variance estimation by the balanced half-sample replication method. Half-sample replication weights for this purpose are included in the microdata files.

A recent study has confirmed the feasibility of including generalized variance functions in SASS publications (Salvucci and Holt, 1992; Salvucci, Galfond and Kaufman, 1993). These functions, which relate the sampling error of an estimate to its size, can be used by those who do not work with microdata files or lack the software for the replication method to produce approximations to the sampling errors associated with their estimates of interest. Specific parameter values for the four basic surveys in Round 1 have been computed and will be used for internal analyses. Parameter values for Round 2 are being developed and will be made available in a forthcoming NCES technical report.

The balanced half-sample replication method assumes that sampling units have been selected with replacement, but in fact sampling without replacement is used in all of the SASS surveys. Violation of the assumption leads to overestimates of the true variances, but the effects are small unless the sampling fractions are quite large. Large sampling fractions do occur in the selection of samples of LEAs for the public sector Teacher Demand and Shortage Survey in some of the smaller states, so it is likely that sampling errors in those states will be substantially overestimated (see Section 4.4, Variance estimation).

Comparisons with data from external sources Comparisons of school, teacher and enrollment counts with other NCES surveys have been discussed in connection with "Coverage error" at the beginning of this section. Here we summarize comparisons of SASS data with information available from sources other than NCES:

- The Census Bureau collects data on school enrollment annually in the October Supplement to the Current Population Survey (CPS), a monthly national sample survey of households. SASS estimates of private elementary and secondary school enrollment from Round 1 exceeded the CPS estimates for the same school year (1987-88) by 15 percent. NCES surveys of private schools prior to SASS had shown similar differences with CPS enrollment estimates during the 1980s (see Section 2.5, Evaluation of published estimates: Round 1).
- The National Catholic Education Association conducts an annual census of Catholic schools. SASS Round 1 estimates of the number of Catholic schools and their enrollment exceeded the Association's census counts by 6.1 and 7.8 percent, respectively (see Section 2.5, Evaluation of published estimates: Round 1).
- Public school administrators' salaries reported in the Round 1 School Administrator Survey were compared with data obtained directly from state education agencies in selected states. The two sets of data were not fully comparable; however, the patterns were similar and there were no obvious inconsistencies (Section 3.5).
- Round 1 estimates of teachers' salaries were compared with data from private organizations. The Teacher Survey estimate of average base salary, \$26,231, was 6.6 percent below a \$28,071 estimate of average salary for the same school year from a 1989 survey conducted by the American Federation of Teachers and 6.4 percent below an estimated average salary of \$28,029 reported by the National Education Association (see Section 5.5, Evaluation of estimates: Round 1).

7.3 Current research

This section describes several research, development and evaluation activities which are in various stages of completion. Some are just getting underway. For others, data have been collected or compiled and the results are being analyzed. Included in this section is a description of a new reinterview procedure that was introduced in Round 2 of the Teacher Followup Survey.

Expansion of coverage and content: library media centers and staff Two projects are related to plans to expand the coverage and content of SASS. As part of a pretest for Round 3 of SASS that was conducted during school year 1991-92, questionnaires for collecting data about public and private school library media centers and library media staff specialists were tested. The media center questionnaires included items on staffing, facilities, collections, equipment, expenditures, users and services provided. The questionnaires for library media specialists

asked for information about training and experience, current status and roles, perceptions and attitudes toward work, compensation, incentives and demographic characteristics.

The pretest sample included 682 public and private schools. Questionnaires were returned for 525 school media centers and 410 media specialists. The lower response for media specialists occurred in part because some schools do not have such persons on their staffs. Item nonresponse and other features of the pretest responses were analyzed and the questionnaires redesigned for use in Round 3 of SASS (Williams, 1992).

Expansion of coverage and content: students Collection of data about students is another area of expansion for SASS. Data from school records for a sample of students are being collected from a subsample of schools in Round 3 (1993-94). The subsample for this component includes an oversample of Indian schools, public schools with high Indian enrollment, and public schools located in Alaska. The student data collected from school records will be linked to data for their schools and teachers. Topics covered include: students' demographic characteristics; disabilities; course work, including advanced placement courses; and participation in special programs (Colaciello, 1993b).

Procedures for selecting samples of students and obtaining information about them from school records were tested in 1991. The initial sample for the study consisted of 200 public and private elementary and secondary schools. Of these, 192 were found to be eligible for participation and 174 (90.6 percent) of the eligible schools agreed to participate in the test.

For each participating school, selection of a sample of students proceeded in three stages: selection of a sample of up to five teachers, selection of a day and class period for each of the sample teachers, and selection of a sample of five students from the rosters for the selected class periods. For the third stage, two different procedures were used. Half of the schools were asked to submit a teacher's student roster for each of the selected class periods and the sample of students was selected by the Census Bureau. The schools were then asked to record information for the selected students on a student questionnaire. The other half of the schools were asked to select the student samples themselves, according to instructions that were included on the questionnaire to be used for recording the information about the sample students.

For each student selected by either method, the schools were asked to provide information on demographic characteristics, current grade level, attendance, participation in special programs, disabilities and, for students in grades 9 to 12, grade-point average, type of program and current enrollment in mathematics and science courses. Information about the student's attendance at classes taught by each of the sample teachers in his or her school (sometimes called multiplicity information) was also requested to provide a basis for the determination of selection probabilities for the students included in the sample.

Analysis of the sampling operations suggested that either procedure could be used successfully. Selection of the student sample by the Census Bureau maintains tighter control

over the process, but requires more time and an additional mailing to the schools (Frazier, 1992).

A review of the data from the 1991 pilot test showed that the multiplicity items, which provide the data needed to weight the sample students correctly, were not producing high quality results. Research was undertaken, using cognitive interviews with school administrators and teachers, to develop and test improved versions of these items and to determine whether they could be more readily answered by school administrators or by teachers. Teachers were found to be better respondents for the multiplicity items.

A field test of the new questionnaires and procedures was conducted in the spring of 1993, using a sample of 282 public schools and 194 private schools to test the collection of student data by mail with telephone followups. For a separate sample of 28 schools in 5 states, the sampling and collection of data for students were carried out by personal visits. The majority of this latter group were Indian Schools and public schools with high Indian enrollment (Colaciello, 1993a).

Using the procedures developed in these tests, student data are being collected for a subsample of schools in Round 3. Information about estimation procedures will be included in a forthcoming paper by King and Kaufman (1994).

Periodicity Rounds 1 to 3 of SASS have been conducted at three-year intervals, but some thought has been given to the possibility of a different cycle. An initial exploration of the implications of cycles of varying length was undertaken to provide guidance to the decision on when the second round of SASS should be conducted. Models were developed to explore the tradeoffs between the cost of a survey cycle and the errors of key estimates, with the cost depending on periodicity and sample sizes and the error being expressed as a composite of sampling error and the error of prediction based on prior year estimates. Several other factors, including response burden and the need for time to evaluate the Round 1 content and methodology, influenced the decision in favor of a three-year interval between Rounds 1 and 2.

The question of periodicity is now being reexamined. The assumptions underlying the models that were used previously are being reviewed and the results of various optimization calculations based on estimates of key variables from Rounds 1 and 2 will be evaluated (Ghosh, Kaufman and Smith, 1994; Smith and Ghosh, 1994).

Alternative reinterview procedure A new reinterview procedure has been tested in Round 2 of the Teacher Followup Survey. For all previous reinterviews in Rounds 1 and 2, responses were obtained for selected questionnaire items, but no attempt was made to reconcile differences between responses given in the initial interviews and the reinterviews. The data from the original interviews and the reinterviews were used to estimate simple response variances and other measures of response consistency for the items included in the reinterviews.

The Round 2 reinterviews for the Teacher Followup Survey included a procedure for reconciling differences. Responses from the initial interviews were transcribed to the questionnaires used by the Census Bureau field representatives who conducted the reinterviews. After completing all of the selected items in the normal way, the field representatives were instructed to compare the interview and reinterview responses. For every item that had a different response, they were to try to determine, using a specified set of questions, which of the two responses was correct and why the difference occurred (Harris, 1992a,b). This information will provide the basis for estimating both response variance and response bias, and is expected to be useful in improving the wording and format of the questionnaire items included in the reinterviews. Some initial results are given by Jenkins and Wetzel (1994a,b).

Alternative modes of data collection Possible changes in the modes of data collection for SASS are being evaluated. As mentioned in Chapters 1 and 2, development and testing of computer-assisted methods of response for schools and LEAs has begun. Interactive diskettes with the survey questions will be mailed to respondents, who will complete them using their own computers (see Section 1.2, Evolution of the SASS design). A prototype will be tested in a small sample of schools during Round 3 of SASS. This method of data collection has already been used successfully by the NCES for the completion, by state offices, of questionnaires relating to public libraries and the completion of questionnaires for academic libraries (Kindel, 1992).

A first attempt to evaluate the feasibility of collecting data for LEAs from state education agencies was inconclusive. There were substantial differences between items reported directly by LEAs and the corresponding values reported by the state agencies. Further research would be needed to understand the reasons for the differences and to identify specific items which might be adequately reported at the state level (see Section 4.3, "The State Data Project").

Improvement of response rates When telephone followups are necessary for teachers who do not mail in their questionnaires, it has proved difficult to reach them at their schools and complete the interviews by telephone while they are there. In the pretest for Round 3 of SASS, conducted during school year 1991-92, postcards were sent to teachers during the mail followup phase asking them to supply their home telephone numbers if they were willing to be contacted at home (Section 5.3, Test of new followup procedure). The proportion responding was low; hence this procedure is not being used in Round 3.

A study is underway, using data from all of the SASS surveys in Round 2, to compare the characteristics of nonrespondents and respondents, based on the sampling frame information that is available for both groups. It is hoped that the results of the study will suggest methods of improving response rates for problem groups and also possible improvements in the nonresponse adjustments used in developing estimates from the data for responding units (Moonesinghe, Smith and Gruber, 1993; Scheuren, Parke and Bureika, 1994).

Coverage improvement The quality of SASS data is affected in many ways by the quality of the sampling frames for schools, LEAs and teachers. Frame imperfections -- omissions, duplications and incomplete or incorrect information about the characteristics of units included in the frames -- can cause both biases and increases in the sampling errors of the survey estimates. The target populations change over time and the frames must be updated to reflect these changes.

Several current evaluation and research projects are aimed at the improvement of the sampling frames and other features of the SASS surveys that relate to coverage. For public schools and LEAs, the CCD has been adopted, starting in Round 2, as the frame of choice. As discussed in Section 2.2 (Evaluation of the sampling frames), some problems of omission, duplication and incomplete information on school and LEA characteristics have been encountered in using the CCD. A plan has been developed for a detailed assessment of the quality of data collected in the CCD surveys, including the data that are used to create and maintain the LEA and public school sampling frames (Nisselson, Parke, Streett, Salvucci and Fink, 1993; Peng, Gruber, Smith and Jabine, 1993).

For private schools, NCES requested the Census Bureau to undertake a detailed analysis of private school list and area frames and the procedures for updating them (Bynum, 1992; Dillen and Jackson, 1992). Results of 1991 updating operations for both frames were evaluated to determine which sources and strategies were most effective for frame updates. A preliminary analysis of the additions to the list frame is available (Jackson, 1993) and additional results are presented by Jackson and Frazier (1994).

In Rounds 1 and 2, there have been several instances of LEAs and public schools completing questionnaires for units other than those intended. For example, a school questionnaire may be completed for two different schools at the same physical location or for all of the schools in a small school district. An LEA may complete a questionnaire for a single school, rather than all of the schools under its jurisdiction. (For further detail, see Section 2.3, "Measurement error: findings from in-depth interviews", Section 2.5, and Section 4.3, "Measurement errors associated with data collection".) Work is continuing on efforts to redesign the instructions and initial items on the school and LEA questionnaires to make it easier for respondents to identify the units for which they are being asked to report.

Finally, as mentioned in Chapter 5, the numbers of teachers listed by the schools for sampling purposes are, on average, smaller than the teacher counts reported on the School Survey questionnaires (Section 5.2, Frame evaluation). During the 1992-93 school year an extensive test, the Teacher Listing Validation Study, was undertaken to seek answers to 3 questions:

- (1) What kinds of problems do schools have in completing the teacher listing forms?
- (2) For public schools, would LEAs be able to provide teacher listings that are more accurate than those prepared by the schools?

- (3) What is the relative accuracy of teacher counts from the school questionnaires and the teacher listing forms?

The first two questions were investigated for a sample of 300 private schools, 290 public schools and 254 LEAs (some LEAs had more than one sample school). The third question was investigated for a separate sample of 300 public schools and 290 private schools. All of the schools in these two samples were asked to complete teacher listing forms and the LEAs were asked to complete teacher listing forms for the sample school(s) in their districts. Various techniques involving personal and telephone interviews for a subsample of schools were then used to investigate the study questions. The field phase of the Teacher Listing Validation Study has been completed and some results are now available (Royce, 1993, 1994).

Evaluation of estimates For several variables, SASS obtains information from more than one survey. Estimates of the number of teachers, for example, can be obtained from the School, Teacher, and Teacher Demand and Shortage Surveys. When aggregate estimates for school districts, states and other domains are compared, the differences are sometimes larger than could be accounted for by sampling variability. A Cross-Questionnaire Estimates Comparison Study is being undertaken to systematically document comparable estimates that can be produced from more than one SASS survey, compare them at several levels of aggregation, and identify possible reasons for differences (Kasprzyk and Scheuren, 1994; Fink, 1994).

Additional information Results of these ongoing research, development and evaluation activities will be documented in internal memoranda, contractor reports and, where appropriate, in NCES working papers, technical reports and papers presented at professional association meetings or in journals. Based on this *SASS Quality Profile*, NCES is reviewing past and ongoing research on the quality of SASS data, with a view toward identifying gaps in our knowledge and establishing priorities for future research activities. For a forthcoming document based on this review and for further information about the status of specific projects, write to:

SASS Quality Profile
555 New Jersey Avenue, N.W.
Washington, D.C. 20208-5651

7.4 Suggestions for users

User options The three basic means of user access to SASS data are publications, public-use data tapes and restricted-use data tapes. Data tapes contain individual records for schools, school administrators, public school districts or teachers, so that users may tabulate or analyze the records as required to meet their specific needs and, with some restrictions, link data across surveys. The public-use data tapes can be obtained through the Inter-university Consortium for Political and Social Research. Their content is limited in order to protect the confidentiality of individual respondents. Data for Round 1 are also available on CD-ROM

and data for Round 2 will be issued in this format also. For information, write to:

Superintendent of Documents
U.S. Government Printing Office
P.O. Box 371954
Pittsburgh, PA 15250-7954

Restricted-use data tapes contain additional information, allowing users to analyze data for more detailed geographic areas and for the complete set of private school association groups. Researchers desiring access to restricted-use data tapes must apply to NCES for a site licensing agreement to use the tapes. Write to:

Associate Commissioner
Statistical Standards and Methodology Division
NCES/OERI, U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208

For detailed current information on SASS publications and how to obtain them, call 1 (800) 424-1616. For information about the purchase of data tapes write to:

U.S. Department of Education
OERI/EIRD/Data Systems Branch
555 New Jersey Avenue, NW
Washington, DC 20208-5725

(202) 219-1522 or 219-1847

Learning about SASS Even the most casual user of SASS data can benefit by learning something about the objectives, content and design of the SASS surveys, and how they relate to each other. A good starting point is the current version of *An Overview of SASS and the TFS*.

Purchasers of data tapes for a survey receive the data file user's manual for that survey. The Round 1 manual for each of the four base-year surveys (NCES, 1991a,b,c,d) contains information on the design and procedures for all four surveys, plus the data base documentation and copies of the questionnaires for the particular survey. For Round 2, there is a single manual, in three volumes, covering all four of the base year surveys (Gruber, Rohr and Fondelier, 1993). Volume I provides general information about survey content, design and methodology; Volumes II and III contain the detailed specifications for the Restricted-Use Version and the Public-Use Version of the datafile, respectively. For each round there is a separate manual for the Teacher Followup Survey (Faupel, Bobbitt and Friedrichs, 1992; Whitener, Rohr, Bynum, Kaufman and King, 1994 -- public-use and restricted-use versions).

Users of data from SASS publications are urged to read the technical notes and appendices included in those publications. Detailed technical notes for Round 1 of SASS are included in NCES Publication 92-120, *Schools and Staffing in the United States: A Statistical Profile* and in the corresponding publication for Round 2, *Schools and Staffing in the United States: A Statistical Profile, 1990-91*, NCES Publication 93-146. The publications in the *E.D. Tabs* series contain technical notes and some include copies of questionnaires for the surveys on which they are based. Detailed accounts of the sample design and estimation procedures for Rounds 1 and 2 of SASS are given in technical reports by Kaufman, *1988 Schools and Staffing Survey Design and Estimation*, NCES 91-127, and by Kaufman and Huang, *1990-91 Schools and Staffing Survey: Sample Design and Estimation*, NCES 93-449.

A SASS User Group, consisting of SASS data users in the Washington DC area, meets about once a year. NCES representatives announce plans for the next round of the survey and solicit user opinions about the availability of data and ease of use. Once every year NCES representatives meet with members of private school associations to share relevant information with them and solicit their views on various aspects of the survey. A SASS Review Board, consisting of distinguished researchers, meets with the NCES staff periodically to provide advice on technical questions.

Using cross-sectional data Conscientious analysts and researchers will want to have a thorough understanding of the nature and limitations of the data they are using. We recommend, of course, that they read the parts of this report that are relevant to the data they are working with. Some additional suggestions are:

- Review the questionnaires (available in the data file user's manuals) and examine the wording and format of the specific data items that are of interest to them.
- Take sampling errors into account. What is the confidence interval for an estimate of interest? Could observed differences have occurred by chance or are they statistically significant? As noted in Section 7.2, "Sampling error", SASS publications include sampling errors for key items, and users of data tapes may, if they wish, estimate sampling errors for their variables of interest. Generalized variance functions, which provide approximations of sampling errors for all estimates, based on their size, will be available soon for Round 2 estimates.
- Consider the possible effects of nonresponse error on the estimates of interest. For example, in making comparisons of public school data by state or private school data by association group, take note of the substantial variations in unit response rates by state and by association group. For specific items, note the item response rates. The data tapes include flags which identify all items that were changed in the computer imputation phase of data processing. For Round 1, a single flag was used to identify all items imputed at this stage in the School and Teacher Demand and Shortage Surveys; most missing items were not imputed and no flags were provided for those that were imputed in the other surveys. In Round 2 missing items were imputed for

all surveys. Items that were imputed following the computer edit stage of processing were assigned flags that distinguish imputations based on other information for the same unit (internal imputation) from those based on information for other units with similar characteristics (donor-based imputation). If desired, some or all imputed values may be omitted from analyses or reimputed by alternate methods.

Analyzing changes over time In working with data from the Round 2 surveys, many users will want to look at the changes that have occurred in the three-year interval between Rounds 1 and 2. We encourage SASS data users to do this, but at the same time we urge them to be aware of the many changes in the content, design and procedures for the two rounds of surveys and to consider how these changes may affect estimates of change. SASS is a complex, evolving system of surveys: comparability over time is highly desirable for periodic surveys, but changes are necessary at this early stage in order improve the quality of the data in ways suggested by the early indicators of quality that have been presented and discussed in this *Quality Profile*.

In the chapter for each of the 4 base surveys we have included information on changes between Rounds 1 and 2 in the content, design and procedures for each survey. Following are some key points:

- As described in more detail in the introductory sections of Chapters 2 through 5, there have been numerous changes in content, with deletion of some topics, addition of others, changes in the kinds of information collected for each topic and changes in the wording and format of individual items and sub-items. Users interested in estimates of change are advised to obtain copies of the questionnaires for both rounds and review the questions and response categories for the items that are of interest to them.
- As described in Chapter 2, a new, more rigorous procedure was used in Round 2 to develop a locale or "urbanicity" code to describe the type of community in which each sample school was located. This locale classification will differ in many instances from the code for the self-report of community type obtained in both Rounds 1 and 2 on the School and School Administrator questionnaires (Gruber, Rohr and Fondelier, 1993, p.137; Johnson, 1993).
- In Round 1 a separate Teacher Demand and Shortage Survey questionnaire was sent to private schools in the sample. In Round 2, a single questionnaire containing the questions for both the School and the Teacher Demand and Shortage Surveys was sent to the sample of private schools. As shown in Table 7.1, the response rate to the combined questionnaire in Round 2 was substantially higher than the response rate to the separate Teacher Demand and Shortage Survey questionnaire used in Round 1.
- The shift from the use of the QED list as a sampling frame for schools in Round 1 to the CCD list in Round 2 was accompanied by a change in the definition of a school. The definition of a school for Round 2 was the same as the CCD definition, and the

public-use and analysis tapes use this definition. However, to permit comparisons, some tabulations will be prepared using both the QED and CCD definitions (Holt and Scanlon, 1994; Saba, Zhang and Chang, 1994).

- In Round 1, virtually all missing or inconsistent data items were imputed for schools and school districts, but there was only limited imputation of missing data for school administrators and teachers. In Round 2, missing or inconsistent data items were imputed for all surveys.
- Items changed in the computer imputation phase of data processing are identified on the data tapes by imputation flags for schools and school districts in Round 1, but not for school administrators and teachers. Imputation flags are provided for all of the surveys in Round 2. The flags used in Round 2 distinguish imputations based on other data for the same units (internal imputation) from those based on data for other similar units (donor-based imputation).
- In Round 2, a new weighting factor was applied to data from the Teacher Survey to force agreement between estimates of teacher counts based on the School and Teacher Survey questionnaires. This weighting factor was not used in Round 1, with the result that estimates of teacher counts from the Teacher Survey were, in general, below those from the School Survey. As a result, estimates of change in numbers of teachers between Rounds 1 and 2 based solely on the Teacher Survey will, in general, be overestimates.
- It is not necessarily safe to assume that measurement bias affecting specific items will remain constant over time and therefore will have little or no effect on estimates of change. There are indications that the effects of some kinds of measurement bias may have been smaller in Round 2. Table 7.1 shows that unit response rates increased for all of the base-year surveys in Round 2: as a result, biases associated with unit nonresponse are likely to have decreased. The systematic reinterview program has provided indications that improved wording and format of specific items has led to smaller response variance and bias for some items, one example being provided by the reports of degrees earned by school administrators and teachers, as discussed in subsections of Sections 3.3 and 5.3, both covering "Measurement error: findings from reinterviews".

The estimation of sampling errors for estimates of change is somewhat more complicated than for point estimates. As a first approximation, the variance (square of the standard error) for an estimate of change can be taken as the sum of the variances of the Round 1 and Round 2 estimates from which it was derived. In many instances this formulation will overestimate the true variance. It assumes no correlation between the two estimates, whereas there will in fact be some correlation because of the deliberate introduction of overlap in the samples of schools and school districts for the two rounds. More precise values and procedures for calculating sampling errors of estimated changes will be provided as they are developed.

User feedback If you are a user of SASS data, we invite you to let us know about your experiences in using the data, any problems you may have encountered, and your suggestions for improving the quality of data from any or all of the surveys. We also invite your comments on this *Quality Profile*. Has it been useful to you, and what additional information should we include in future versions to make it more useful? Please write to:

SASS Quality Profile
555 New Jersey Avenue, N.W.
Washington, D.C. 20208-5651

Table 7.1 Response Rates by Survey and Sector: Rounds 1 and 2

Survey and Sector	Round 1	Round 2	Type of Estimate
School			
Public	91.9	95.3	Weighted
Private	78.6	83.9 ^{1/}	Weighted
School Administrator			
Public	94.4	96.7	Weighted
Private	79.3	90.1	Weighted
Teacher Demand and Shortage			
Public	90.4	93.5	Weighted
Private	66.0	83.9 ^{1/}	Weighted
Teacher			
Public	83	86	Composite ^{2/}
Private	70	75	Composite ^{2/}
Teacher Followup: Current Teachers			
Public	81	84	Composite ^{2/}
Private	67	72	Composite ^{2/}
Teacher Followup: Former Teachers			
Public	78	79	Composite ^{2/}
Private	65	71	Composite ^{2/}

Notes:

1. For the private sector, the School and Teacher Demand and Shortage Survey questionnaires were combined in Round 2.
2. Combination of weighted and unweighted rates reflecting losses at all stages.

Sources:

See Tables 2.5, 3.1, 4.1, 5.1 and 6.1.

