

LINKING THE NAEP AND SASS SURVEYS: METHODOLOGICAL ISSUES AND POTENTIAL USES

Gary Skaggs, West Mesa Associates, Inc., Steve Kaufman, National Center for Educational Statistics
Gary Skaggs, 145 Dixon Drive, Hardy, VA 24101

KEY WORDS: NAEP, SASS, Survey Methods

Background

The National Center for Education Statistics (NCES) is currently planning to conduct a research and development study that would link survey data from State National Assessment of Educational Progress (NAEP) and the Schools and Staffing Survey (SASS) in the 1999-2000 school year. The purposes of the study are to determine if: 1) survey linking is feasible on a larger scale and 2) reliable and valid data can be produced from the linkage. The purpose of this paper is to describe the current plan for this project and discuss the methodological and substantive issues it will address.

NAEP began as a survey of what students at various ages know and can do in a variety of subjects. Beginning in 1989, the focus of NAEP shifted toward measuring what students *should* know and be able to do. A major aspect of this shift involved setting achievement levels and reporting results in terms of these levels. Another aspect of this shift concerned reporting results for individual states. NAEP surveys consist of subject area assessment items, student background variables, teacher background variables, and school characteristics. A number of NCES publications reports national NAEP results in terms of percent of students falling at or above achievement levels. In the 1999-2000 school year, state NAEP was administered to fourth and eighth grade students in mathematics and science.

SASS originated in the mid-1980's as a response to the school reform movement that began with *A Nation at Risk*. SASS was administered in 1987-88, 1990-91, and 1993-94. It focused on issues of teacher supply and demand, teacher and administrator working conditions, and basic conditions in schools. The 1999-2000 SASS shifts its emphasis toward measures of teacher and school quality and away from teacher quantity. SASS surveys consist of school, school district, principal, and teacher questionnaires.

From the initial SASS conceptual design, there has been consideration given to directly linking SASS and NAEP. Both surveys can potentially benefit from such a linkage. SASS would get a direct measurement of student outcomes in specific subject areas, while NAEP would get a large numbers of district, school,

principal and teacher characteristics that it currently does not have. The respondent burden consideration has always been the overriding issue that has stopped such a linkage. This was especially true in the early SASS and early state NAEP years, where reaction to the respondent burden for the individual surveys was unknown.

During the 1993-94 school year, NAEP and SASS were fielded together using independently selected samples. At that time, there were discussions about minimizing the overlap, as was done during the 1987-88 school year. In 1993, NAEP and SASS used different frames. The difficulty in matching the two frames, so an efficient minimization can be achieved, precluded the possibility of minimizing the overlap. However, this situation provided an opportunity to analyze the respondent burden, to determine whether a NAEP/SASS overlap sample is possible. During the 1999-2000 school year, it became possible to link both surveys, because, once again, they were being fielded together.

Sampling Plan

From the 1993-94 sample designs, it was expected that an overlap of approximately 800 4th grade schools and about 1,100 8th grade schools would occur naturally from the sampling methods used by both survey programs. However, the two programs employ vastly different sampling plans. For SASS, the U.S. Bureau of the Census derives a national probability sample. For State NAEP, Westat derives separate probability samples for each participating jurisdiction (States). Since both surveys are designed to produce state estimates, they both oversample small states and undersample large states. This means that most of the overlap in schools occurs in small states. As a result, national estimates will not be efficient.

For the 1999-2000 school year collection, using a sampling plan developed by Steve Kaufman of NCES, the natural overlap between the two survey programs has been minimized to about 200 4th grade and 500 8th grade schools. After the State NAEP sample was selected, Westat provided the Census Bureau with the NAEP selection probabilities for the entire NAEP sampling frame. With this information, Census selected a SASS sample that minimized the naturally occurring overlap between the NAEP and SASS school

samples. Once this was completed, Census selected nationally representative subsamples of 400 4th grade and 400 8th grade public State NAEP schools. These 800 schools then received the SASS public school questionnaires.

This sample size was chosen with stratification in mind for the four SASS geographic regions. In addition, each region can be further stratified by locale (central city, urban fringe/large town, and rural/small town). Regional and locale strata are major publication categories for SASS.

The 1999-2000 samples have an expected overlap of 600 4th grade schools and 900 8th grade schools. This is smaller than the 1993-94 expected overlap of 800 4th grade and 900 8th grade schools. So, in terms of this measurement of respondent burden, the 1999-2000 surveys provides less burden than the 1993-94 surveys, while including a nationally representative NAEP/SASS overlap sample. Additionally, the response rates in 1993-94 were not greatly affected by the large independent NAEP/SASS overlap, compared to previous rounds of the surveys. In fact, the 1994 NAEP response rates appeared to increase from the 1992 rates.

Structure of the Database

The SASS/NAEP overlap sample will consist of public schools that participate in State NAEP. The primary linkage between the two databases will occur at the school level in the form of a single school-level data file which will combine, for each participating school, variables from the SASS public school file and the school files from State NAEP.

In addition to this linked “master” school-level file, there will be several files unique to each program. SASS district, teacher, and principal files will be created and linked to the master school file. A NAEP student file will be created and linked to the master school file. To simplify linkage, a school district code will be used throughout all data files, and a school code will be used in all school, teacher, and student files. By having these codes common to all files, it should be easy for both student-level and school-level analyses to be carried out.

Descriptions of the data files are provided below.

Master School File. This file will contain responses to the items from the SASS School Questionnaire (public schools) and the additional variables created by the SASS program. It will also contain responses to the items from the NAEP School Characteristics and

Policies Questionnaire and the additional variables created by the NAEP program. The usual sampling weights for each survey program will be replaced by a single set of school weights, developed by Census, for this database.

In addition, the file will contain aggregates of student performance from NAEP. These will consist of five NAEP school means (one for each set of plausible values) in each subject area and a measure of within school variation (e.g. percents of students falling within each achievement level, calculated by summing posterior distributions for each student). Most of the 400 schools will receive both math and science, so there will be space for records on both subjects.

SASS School Principal File. This file will contain the usual data found in the SASS school principal file plus principal weights unique to the overlap sample. The file is linked to other files through the school district control number and the school control number.

SASS School District File. This file will contain data from the SASS School District Questionnaire. The file is linked to other files through the school district control number. There will be no district weights for two reasons. First, SASS samples schools, then administers the district questionnaire to the district associated with each school. Therefore, analyses at the district level would not be based on any sampling within districts. Second, the number of schools per district varies widely with one school per district as the mode. For this exploratory study, district variables can be linked to schools and analyzed as school level variables.

SASS Teacher File. This file will contain the usual data found in the SASS teacher file plus two teacher weights, the final teacher weight and within-school teacher weight. The file is linked to other files through the school district control number and the school control number. The SASS teacher questionnaire is administered to a random sample of teachers within each school. While it would be most desirable to link this questionnaire to the NAEP teacher questionnaire and control for grade level, this will not be possible in this study. By using final teacher weights, relationships between teacher variables and NAEP achievement can be carried out.

On the other hand, because of the random selection within a school, the SASS teacher variables can be aggregated and treated as school-level variables. For this purpose, the within school weight can be used. Certain key variables, such as years of experience, teacher perceptions of school climate, and numbers of

out-of-field teachers, from the teacher questionnaire will be included in the Master School file. A number of issues remain to be worked out, such as how to treat missing data.

NAEP Student File. This file will contain the usual data found in the NAEP student file, including NAEP proficiencies, item responses, and student and teacher background item responses, and a unique set of weights developed by Census or Westat. The file is linked to other files through the school district control number and the school control number.

A first step in looking at possible analyses is to examine the anticipated data structure of the overlap database. Table 1 below shows target and expected sample sizes for grade 4 in numbers of schools. Table 2 shows the same figures for grade 8. The target numbers are the first in each cell. They are based on Steve Kaufman's sample design and Census's implementation of it. Briefly, of the 400 overlap schools at each grade, 80 were chosen for each SASS region. The rest were selected to increase the representation of the larger states and to permit the use of NAEP regions.

The target numbers represent the numbers of schools that both NAEP and SASS will contact. However, while every effort will be made to persuade schools to

participate, it is well known that not all schools will do so. In NAEP, substitute schools are chosen. Because SASS data collection begins in September, it will not be possible to use the substitute schools in the overlap database. In the 1993-94 SASS, about 92 percent of schools responded to the public school questionnaire. In the 1996 State NAEP in Science in grade 8, about 87 percent of schools in participating states participated prior to substitution. If participation decisions in the two programs are not related, and if participation rates are equal across all strata, then we can expect about 80 percent of the 400 schools to provide both NAEP and SASS data at the school level. The second number in each cell represents this expected sample size.

Tables 3 and 4 translate the numbers of schools into projected numbers of students. The 2000 State NAEP sampling plan calls for 30 students to be assessed in each selected school in each subject. For various reasons, not all 30 student records are available for analysis. The average number of students per school varied considerably. In the 1996 State NAEP in Science, across 45 participating states, the average was slightly less than 28 students per school. The numbers of students in Tables 3 and 4 were generated by multiplying the numbers of schools from Tables 2 and 3 by 28.

**Table 1
Target/Expected Sample Sizes (Number of Schools) for Grade 4**

SASS Region/ CCD Locale	1=Central city	2=Urban fringe /large town	3=Rural/ small town	Total
1=Northeast	28/22	42/37	22/18	92/77
2=Midwest	30/24	28/22	37/30	95/76
3=South	38/30	34/27	48/38	120/95
4=West	33/26	37/30	23/18	93/74
Total	129/102	141/116	130/104	400/322

**Table 2
Target/Expected Sample Sizes (Number of Schools) for Grade 8**

SASS Region/ CCD Locale	1=Central city	2=Urban fringe /large town	3=Rural/ small town	Total
1=Northeast	25/20	39/31	27/22	91/73
2=Midwest	27/22	26/21	46/37	99/80
3=South	30/24	29/23	59/47	118/94
4=West	27/22	32/26	33/26	92/74
Total	109/88	126/101	165/132	400/321

Table 3
Expected Sample Sizes (Number of Students) for Grade 4

SASS Region/ CCD Locale	1=Central city	2=Urban fringe /large town	3=Rural/ small town	Total
1=Northeast	616	1036	504	2156
2=Midwest	672	616	840	2128
3=South	840	756	1064	2660
4=West	728	840	504	2072
Total	2856	3248	2912	9016

Table 4
Expected Sample Sizes (Number of Students) for Grade 8

SASS Region/ CCD Locale	1=Central city	2=Urban fringe /large town	3=Rural/ small town	Total
1=Northeast	560	868	616	2044
2=Midwest	616	588	1036	2240
3=South	672	644	1316	2632
4=West	616	728	728	2072
Total	2464	2828	3696	8988

The estimates in the tables above are only approximations based on past experience, but they do give an idea of the limitations for analysis work. Both SASS and NAEP have statistical guidelines for minimum sample sizes based on NCES guidelines. For SASS, that minimum is 30 schools per cell. Some but not all of the cells in the above tables meet this requirement. For NAEP, there is a minimum subgroup size of 62 if those students are adequately distributed across schools. Therefore, the database will be able to produce results for some variables, such as region and locale, but not others, such as by state. Further, NAEP has had in place for some time policies and considerations regarding treatment of missing data, hypothesis testing, and Type I error rates. These will be followed in publications based on this database. Given these considerations, two types of descriptive analyses are proposed.

Methodological Issues

The primary purposes of the study are to answer questions about future feasibility and utility and about the validity of further analyses with this database. However, a number of methodological issues need to

be addressed in order to meet these purposes. The analyses include the following:

Participation Rates. As was the case in the 1993-94 SASS, there is still a major concern that the additional burden of responding to both NAEP and SASS will have an adverse impact on the response rates. The additional burden is on principals (or other administrator who completes the school surveys) and teachers. It may that some teachers respond to both the NAEP and SASS teacher questionnaires, but we will have no way of identifying them. Therefore, the response burden analysis should be conducted for each questionnaire as well as overall school participation. If feasible, some qualitative information would be useful, such as debriefing interviews with survey managers to determine if there were complaints from respondents.

Comparison of SASS questionnaire responses to full SASS. These comparisons will help determine how representative the overlap sample is of the full SASS. However, at the present time, there are 41 jurisdictions participating in State NAEP. Therefore, this analysis reduces to the comparison of estimates from the overlap sample to those from the 41 State NAEP jurisdictions. These estimates will be produced from SASS questionnaire items or variables created from SASS questionnaire items.

Comparison of NAEP proficiencies and background questions to NAEP. These comparisons will help determine how representative the overlap sample is of the main NAEP. The overlap sample will consist of states participating in the 2000 State NAEP. Yet, the interest in this database is in national and regional estimates. To accomplish this analysis, the State NAEP scale needs to be transformed to the Main NAEP reporting scale. Once this is completed, estimates from NAEP questionnaire items or variables created from NAEP questionnaire items will be derived for the aggregate of the 41 participating jurisdictions. These estimates will then be compared to those from the overlap sample.

Comparison of NAEP and SASS Items that Measure Similar Content. There is some overlap between the two sets of school questionnaires. An analysis of the content of the two questionnaires indicates that most items in the NAEP school questionnaire are addressed in some way in the SASS school questionnaire, but the SASS questionnaire includes many topics not addressed by NAEP. For the most part, questions measuring similar content are asked in different ways on the two questionnaires. It will be important to both survey programs to determine the degree to which results are influenced by question format. For example, NAEP obtains race/ethnicity from student background questionnaires while SASS collects it from school surveys. Also, both NAEP and SASS school questionnaires ask questions about parental involvement but do so in different formats. Do the two sources produce the same estimates of minority membership and parent involvement?