B. Collection of Information Employing Statistical Methods

This submission requests clearance for sampling and school recruitment activities for the High School Longitudinal Study of 2009 (HSLS:09) field test and full-scale study to be completed in 2008 and 2009, respectively. This section provides a description of the target universe for this study, followed by an overview of the sampling and statistical methodologies proposed for the field test and the full-scale study. We will also address suggested methods for maximizing response rates and for tests of procedures and methods, and we will introduce the statisticians and other technical staff responsible for design and administration of the study.

1. Target Universe and Sampling Frames

The target population for the HSLS:09 full-scale study consists of 9th grade students in public and private schools that include 9th and 11th grades; their parents; and corresponding math and science teachers, school administrators, and high school counselors. (School eligibility in the field test sample, however, is based on schools that have both a 9th and 12th grade, since fall 12th graders are a psychometric proxy for the spring 11th graders of the HSLS:09 first follow-up.) The needed respondent samples will be selected from all public and private schools with 9th and 11th grades in the 50 states and the District of Columbia.[1] Excluded from the target universe will be specialty schools such as Bureau of Indian Affairs schools, special education schools for people with disabilities, area vocational schools that do not enroll students directly, and schools for the dependents of U.S. personnel overseas.

The primary sampling units (PSU) of schools for this study will be selected from the two databases of the U.S. Department of Education. The Common Core of Data (CCD) will be used for selection of public schools, while private schools will be selected from the Private School Survey (PSS) universe files. To eliminate overlap between the field test and full-scale study samples, the full-scale study sample of schools will be selected prior to the field test sample. However, the early selected full-scale study sample will be "refreshed" by a small supplemental sample of schools that will become eligible in the time between the administration of the field test and of the full-scale study. The secondary sampling units (SSU) of students will be selected from student rosters that will be secured from the sample schools. The PSU and SSU sampling procedures for this study are detailed in the next section.

2. Statistical Procedures for Collecting Information

The following section describes sampling procedures for the field test and full-scale study for which clearance is requested. First discussed is the selection plan for the full-scale study sample of schools, followed by the selection plan for the field test sample, to reflect the sequence that will be observed for PSU selections. Next, selection procedures for the student samples will be presented for the field test and full-scale study that will be conducted in 2008 and 2009, respectively. This section also includes descriptions of the procedures that will be followed after data collection, including survey weight adjustments, to measure and reduce bias due to nonresponse.

[1] While the full-scale HSLS:09 sample will include only 9th-grade students, the field test sample will include both 9th- and 12th-grade students to prognosticate the progression that will be observed when reassessing the sample 9th-grade students in 2012.

a. School Frames and Samples

The latest CCD (2005–2006) will be used as the public school sampling frame and PSS (2005–2006) as the private school sampling frame. Given that these two sample sources provide comprehensive listings of schools, and that CCD and PSS data files have been used as school frames for a number of other school-based surveys, it is particularly advantageous to use these files in HSLS:09 for comparability and standardization across NCES surveys.

As mentioned earlier, the survey population for the full-scale study of HSLS:09 consists of all ninth-graders in the 50 states and District of Columbia enrolled in

- regular public schools, including state department of education schools, that include 9th and 11th grades; and
- Catholic and other private schools that have 9th and 11th grades.

Excluded for this study will be the following:

- schools with no 9th or 11th grade;
- ungraded schools;
- Bureau of Indian Affairs schools;
- special education schools;
- area vocational schools that do not enroll students directly;
- Department of Defense schools; and
- closed public schools.

The school samples will be selected using a stratified PPS methodology for which a composite size measure methodology developed by Folsom, Potter, and Williams (1987) will be used. This methodology will support the desired oversampling of students in key analytical domains (e.g., Asians and Pacific Islanders), maintain near equal sampling weights for students within each domain, and result in approximately equal total student sample sizes within sampled schools. Details of school sample selection for the full-scale study and field test are provided next.

Full-Scale Study School Samples

The public and private school samples for the full-scale study will be large enough to secure 800 participating schools, combined. The needed samples were selected from the CCD (2005–2006) and PSS (2005–2006) within sampling strata defined by

- school type: Public, Catholic, or Other private schools;
- Census region: Northeast, Midwest, South, or West; and
- locality: City, Suburban, Town, or Rural.

As illustrated in table 10, the starting sample of selected schools will be proportional to the number of ninth-grade students within each stratum, based on information from the CCD and PSS. When enrollment information was unavailable for certain schools, missing enrollment counts was imputed as the median value of the enrollment for ninth-graders within race/ethnicity categories in each school stratum. The full-scale and field test samples of schools have been

selected, with the full-scale sample selected first from the entire sampling frames unconditionally.

Table 10. Illustrative school sample allocation and expected yields (full-scale study HSLS:09)

	Total		Northeast		Midwest		South		West	
School stratum	Sampled	Participating	Sampled	Participating	Sampled	Participating	Sampled	Participating	Sampled	Participating
Total	1,349	800	242	144	338	201	504	298	265	157
Public, total	1,012	600	167	100	241	142	395	234	209	124
Public, city	280	167	42	25	59	35	106	63	73	44
Public, suburban	387	229	74	44	91	54	135	80	87	51
Public, town	118	70	23	14	28	16	41	24	26	16
Public, rural	227	134	28	17	63	37	113	67	23	13
Catholic, total	168	100	46	28	58	35	41	24	23	13
Catholic, city	96	58	21	13	33	20	30	18	12	7
Catholic, suburban	54	31	19	10	19	11	8	5	8	5
Catholic, town	16	10	4	4	6	4	3	1	3	1
Catholic, rural	2	1	2	1	0	0	0	0	0	0
Other private, total	169	100	29	16	39	24	68	40	33	20
Other private, city	74	44	11	6	15	9	28	17	20	12
Other private, suburban	56	32	8	5	16	8	25	15	7	4
Other private, town	17	10	3	1	4	4	8	4	2	1
Other private, rural	22	14	7	4	4	3	7	4	4	3

As mentioned earlier, however, a refresher sample of schools will be added to the full-scale sample to account for new schools or those that become eligible after the sampling frames are constructed. For this purpose, frame comparison will be conducted between the 2005–2006 CCD and the 2006–2007 CCD to determine the frequency of new public high schools. Moreover, districts associated with the refresher subsample of schools will be contacted to identify eligible schools recently opened in their jurisdiction. The districts will be provided with a list of all public schools on the sampling frame in their district to help them identify the appropriate schools. Analogous activities will be carried out for private schools using available information from relevant sources such as Quality Education Data (QED), since the 2006–2007 version of the PSS will not be available in time for refreshing the sample of private schools. However, there is a possibility that NCES will be able to secure an early release copy of the next PSS to include in this investigation. Should such a copy be available, it will be used for sample refreshing and related quality control activities.

Obviously, a sample size larger than 800 schools is necessary to compensate for the anticipated nonresponse and ineligibility. As per NCES standards, a weighted response rate of at least 70 percent at the school level will be targeted. In unweighted terms, this means that a sample of size 1,143 schools will be required to secure 800 (or, $1{,}143 \times 0.7$) participating schools. Based on experience with the Education Longitudinal Study of 2002 (ELS:2002), about 4 percent of sampled schools will emerge as ineligible for this study. Consequently, the projected

size for the starting sample will be 1,190 (or, 1,143 × 1.04) schools. Moreover, based on ELS:2002 response rates, the expectation is that an additional sample of 159 schools will be needed to secure 800 participating schools, for a grand total of 1,349 (or, 1,190 + 159) schools.

School recruitment activities will be closely monitored and additional schools will be released as needed to ensure that the goal of 800 participating schools is reached. To this end, in addition to the above sample of 1,349 schools, a reserve pool of 251 schools will be selected should observed yield rates fall below expectations. Operationally, the entire sample of 1,600 (or, 1,349 + 251) schools will be randomly partitioned within each stratum into two release pools and a reserve pool. The two release pools will compose the basic sample of 1,349 schools, and schools in the second pool will be released in waves as needed to achieve the sample size goal. The reserve pool will be released selectively in waves by simple random sampling within stratum for strata with low yield rates, when necessary.

Once the school sample has been selected, data from QED will be used to obtain principal and district superintendent names along with related information that will be needed for contacting schools. Contacted schools will be asked to provide student rosters for those expected to participate in the field test and the full-scale study, accordingly. For refusing schools, an abbreviated questionnaire will be used to obtain important school-characteristic data to complement frame information. The resulting information will enable a more effective analysis of nonresponse bias.

Field Test School Sample

Using probability-based selection of the full-scale study sample of 1,600 schools from the complete CCD and PSS sampling frames, sample schools will be removed from the frames so that a purposive sample can be selected from among the remaining schools to yield 55 participating schools for the field test study. This sample will be divided into 44 public and 11 private schools and will be selected from schools that have both 9th and 12th grades in the states of New York, California, Florida, Illinois, and Texas.

To the extent possible, the stratification plan to be used for selection of this sample will be similar to the one used for the full-scale study sample. Given the small sample size for the field test, however, a somewhat coarser stratification might become necessary to avoid empty strata. As illustrated in table 11, a slightly larger sample of 84 schools will be selected to ensure that at least 50 schools will provide student lists for the field test. Moreover, an additional sample of 20 schools will be selected and kept in a reserve pool should yield rates fall below expectations.

Table 11. Illustrative school sample allocation and expected yields (field test HSLS:09)

	Total		New York		California		Florida		Illinois		Texas	
School stratum	Sampled	Participating	Sampled	Participating	Sampled	Participating	Sampled	Participating	Sampled	Participating	Sampled	Participating
Total	92	55	19	11	19	11	18	11	18	11	18	11
Public, total	72	44	15	9	14	9	14	9	15	9	14	8
Public, city	22	14	4	3	4	3	5	3	5	3	4	2
Public, suburban	27	16	6	4	5	3	5	3	5	3	6	3
Public, town	8	5	2	1	2	1	1	1	2	1	1	1
Public, rural	15	9	3	1	3	2	3	2	3	2	3	2
Catholic, total	10	6	2	1	3	1	2	1	1	1	2	2
Catholic, city	5	3	2	1	2	1	1	1	0	0	0	0
Catholic, suburban	3	2	0	0	0	0	1	0	0	0	2	2
Catholic, town	1	0	0	0	1	0	0	0	0	0	0	0
Catholic, rural	1	1	0	0	0	0	0	0	1	1	0	0
Other private, total	10	5	2	1	2	1	2	1	2	1	2	1
Other private, city	4	2	0	0	1	0	1	1	1	0	1	1
Other private, suburban	4	2	0	0	1	1	1	0	1	1	1	0
Other private, town	1	0	1	0	0	0	0	0	0	0	0	0
Other private, rural	1	1	1	1	0	0	0	0	0	0	0	0

b. Student Frames and Samples

All sampled schools will be contacted and asked to upload their student lists to a secure website to serve as sampling frames for student samples. Moreover, a backup option will allow schools to provide their student lists via e-mail of zipped/password-protected files. If the school cannot provide electronic lists, paper lists will be requested to be faxed to a fax machine in a locked room. For data security reasons, it will be requested that paper lists not be mailed. Each sample school will be asked to provide the following information for each eligible student:

- student ID number;
- full name;
- sex;
- race (White; Black; Asian; Native Hawaiian or Other Pacific Islander; American Indian or Alaska Native);
- ethnicity (Hispanic indicator, regardless of race); and
- whether an Individualized Education Program (IEP) has been filed for the student (yes, no).

Race/ethnicity will be needed to guide oversampling of minority students. Moreover, race/ethnicity along with gender and IEP indicators often serve as effective variables for nonresponse adjustments in the full-scale study.

No students will be excluded from the sampling frame because of disabilities or language problems. Specifically, the HSLS:09 field test and full-scale study will include students with severe mental disabilities, those with limited command of the English language for understanding the survey materials, and students with physical or emotional problems. Schools will identify such students from those sampled and identify possible accommodations for these students to complete the survey and assessment. Students who cannot complete the survey or cognitive tests will be excused from doing so; however, contextual information about such students will be collected from teachers, principals, high school counselors, and parents.

The student lists will be reviewed for quality, and schools whose lists fail the quality checks will be recontacted by the school recruiter to resolve observed discrepancies.[2] Selecting sample students will proceed when confirmation has been obtained that the list received is correct or when corrected list is received. Students will be sampled on a flow basis as student lists are received. The lists will be stratified by race/ethnicity and select a systematic sample of students from the resulting lists. For schools that provide paper lists, a two-stage process will be used that has been used effectively to select systematic samples from paper lists. This simple, yet scientific, method eliminates the need for data entry of the entire list of students when such lists are provided on paper. Instead, only information for sampled students will be data-entered.

Field Test Student Sample

A random sample of 29 students from the 9th grade and 30 students from the 12th grade will be selected in each of the 55 sample schools, for a total of 1,595 (or, 55×29) students in 9th grade and 1,650 (55 x 30) students in 12th grade. Based on the target eligibility and response rates of 95 and 92 percent, respectively, this will result in a sample of 1,538 ($1{,}760 \times 0.95 \times 0.92$) and 1,442 (1,650 x 0.95. x 0.92) responding students in 9th and 12th grade, respectively. This sample has grown from the original design of 50 schools and 25 students per grade to ensure that the sample size is adequate for needs of the field test math assessment and has been further increased from 55 schools and 27 students per grade to ensure adequate sample size. The estimated field test yield is 1,318 (1,595 x 0.95 x 0.87) responding students in 9th grade and 1,332 (1,650 x 0.95 x 0.85) responding students in 12th grade. Table 12 shows an allocation of the sample and responding students for each grade, by school and student characteristics, overall and for each of the five participating states based on the original proportion of 50 schools and 25 students per grade. The five additional schools were apportioned across state and school type accordingly, with four more public schools and one more private school. During the recruitment process, we will ask schools when their student lists will be ready; however, requesting lists and drawing student samples will occur on a flow basis for the field test between August and November 2008.

[2] Inevitably, there will be inconsistencies between student counts obtained from the sample schools and CCD/PSS. When the relative magnitude of an observed discrepancy exceeds 25 percent, such cases will call for further examinations. For instance, for public schools this measure will be the absolute value of (List – CCD)/List.

Table 12. Illustrative student sample allocation and expected yields for 9th- and 12th-graders (field test HSLS:09)

	Total		Hispanic		Asian		Black		Other	
School stratum	Sample	Respondent	Sample	Respondent	Sample	Respondent	Sample	Respondent	Sample	Respondent
Total	1,485	1,298	274	235	59	59	260	235	892	769
Public, city	351	309	64	56	14	14	63	56	210	183
Public, suburban	459	394	85	71	19	19	85	73	270	231
Public, town	135	118	22	21	5	5	21	19	87	73
Public, rural	243	214	44	39	10	10	43	39	146	126
Catholic, city	81	71	16	13	3	3	13	13	49	42
Catholic, suburban	54	48	11	9	2	2	9	9	32	28
Catholic, rural	27	24	5	4	1	1	4	4	17	15
Other private, city	54	48	11	9	2	2	9	9	32	28
Other private, suburban	54	48	11	9	2	2	9	9	32	28
Other private, rural	27	24	5	4	1	1	4	4	17	15
New York	297	258	55	46	12	12	51	46	179	154
Public, city	81	71	15	13	3	3	14	13	49	42
Public, suburban	108	93	20	16	4	4	21	18	63	55
Public, town	27	24	5	4	2	2	4	3	16	15
Public, rural	27	24	5	5	1	1	4	4	17	14
Catholic, city	27	23	5	4	1	1	4	4	17	14
Other private, rural	27	23	5	4	1	1	4	4	17	14
California	297	258	55	47	11	11	53	47	178	153
Public, city	81	71	15	13	3	3	15	13	48	42
Public, suburban	81	70	15	13	3	3	16	13	47	41
Public, town	27	23	4	4	1	1	4	4	18	14
Public, rural	54	48	11	9	2	2	10	9	31	28
Catholic, city	27	23	5	4	1	1	4	4	17	14
Other private, suburban	27	23	5	4	1	1	4	4	17	14
Florida	297	261	55	48	12	12	52	47	178	154
Public, city	81	71	15	13	3	3	15	13	48	42
Public, suburban	81	70	15	14	3	3	15	13	48	40
Public, town	27	24	5	4	1	1	4	4	17	15
Public, rural	54	48	10	9	2	2	10	9	32	28
Catholic, suburban	27	23	5	4	1	1	4	4	17	14
Other private, city	27	25	5	4	2	2	4	4	16	15
Illinois	297	260	54	46	12	12	52	47	179	155
Public, city	81	71	15	13	3	3	15	13	48	42
Public, suburban	81	70	15	12	3	3	15	14	48	41
Public, town	27	23	4	4	1	1	4	3	18	15
Public, rural	54	48	10	9	2	2	10	9	32	28
Catholic, rural	27	24	5	4	1	1	4	4	17	15
Other private, suburban	27	24	5	4	2	2	4	4	16	14

Table 12. Illustrative student sample allocation and expected yields for 9th and 12th graders (field test HSLS:09)—Continued

	Total		Hispanic		Asian		Black		Other	
School stratum	Sample	Respondent	Sample	Respondent	Sample	Respondent	Sample	Respondent	Sample	Respondent
Texas	297	261	55	48	12	12	52	48	178	153
Public, city	54	48	10	9	2	2	10	9	32	28
Public, suburban	81	70	15	13	3	3	15	13	48	41
Public, town	27	23	4	4	1	1	4	4	18	14
Public, rural	54	48	10	9	2	2	10	9	32	28
Catholic, suburban	54	48	11	9	2	2	9	9	32	28
Other private, city	27	24	5	4	2	2	4	4	16	14

Field Test Teacher, High School Counselor, and Parent Samples

One math and one science teacher will be selected for each ninth-grade student. Where sample students have more than one math or science teacher in fall 2008, one of the teachers will be randomly sampled. On the other hand, a number of sample students may not have any math and/or science teachers—a possible reflection of block scheduling—so such students will have no sample teacher. Moreover, for each sample school there will be one sample high school counselor. Where there is more than one counselor at the school, the lead/head/senior counselor will be selected to be in the sample. Experience with this procedure in previous NCES studies, such as the HS&B Administrator and Teacher Survey, suggests that the senior counselors are the most familiar with the school's counseling infrastructure. If this counselor declines to respond, a different counselor, if available, will be substituted. Lastly, for each sample student there will be one sample parent. In two-parent households, NELS:88/ELS:2002 procedures will be followed in asking the parents to identify the parent most knowledgeable about the student's school situation and experience.

Full-Scale Study Student Sample

A sample of 25 students from ninth grade will be randomly selected from the selected 800 schools (600 public and 200 Catholic and other private schools) for a base sample of 20,000 (or, 800 × 25) students. Moreover, this base sample will be augmented by selecting 1,800 additional Asian/Pacific Islander students for a total sample of 21,800 students.[3] This augmentation is required to ensure that this subpopulation meets the minimum sample size needed to achieve the following general precision requirements:

- detect a 15 percent change in proportions across waves of the study;
- detect a 5 percent change in means;
- produce relative standard errors of 10 percent or less for proportion estimates based on data from a single wave of data collection; and

[3] Sample augmentation will not be necessary for Hispanic or Black students, since sufficient sample sizes to support analyses by race/ethnicity will be secured for such students as part of the base sample of 20,000 students.

- produce relative standard errors of 2.5 percent or less for estimated means based on data from a single wave of data collection.

Using student enrollment counts from the CCD/PSS and relying on our experience from the field test, the student sampling rates will be set in advance based on race/ethnicity. Students will be sampled from the student lists received from sample schools, using a stratified, systematic sampling procedure. Sample sizes will be monitored by race/ethnicity and the sampling rates will be adjusted, if necessary, to achieve all sample size goals. While the expectation is to achieve the stated response and eligibility rates, an early identification of low sample yields will be vital in making sure the study can adjust appropriately to reach the target yields. Table 13 shows a possible student sample allocation and yield for the HSLS:09 full-scale study. The anticipated time frame for requesting student lists and drawing student samples on a flow basis is between August and November 2009.

Table 13. Illustrative student sample allocation and expected yields for ninth-graders (full-scale study HSLS:09)

School stratum	Total		Hispanic		Asian		Black		Other	
	Sample	Respondents	Sample	Respondents	Sample	Respondents	Sample	Respondents	Sample	Respondents
Total	21,800	19,053	2,645	2,026	2,419	1,899	2,684	2,039	14,052	13,089
Northeast	3,924	3,430	477	364	434	339	483	367	2,530	2,360
Public, city	681	595	83	63	76	60	84	64	438	408
Public, suburban	1,211	1,058	148	113	135	105	149	113	779	727
Public, town	370	324	45	34	40	33	46	35	239	222
Public, rural	463	405	56	43	51	41	57	43	299	278
Catholic, city	353	310	43	33	39	31	44	33	227	213
Catholic, suburban	292	255	35	27	32	24	36	27	189	177
Catholic, town	90	78	11	8	10	7	11	9	58	54
Catholic, rural	27	24	3	3	3	2	3	3	18	16
Other private, city	164	143	20	15	18	14	20	15	106	99
Other private, suburban	126	109	15	11	14	9	15	11	82	78
Other private, town	38	34	5	4	4	3	5	4	24	23
Other private, rural	109	95	13	10	12	10	13	10	71	65
Midwest	5,477	4,787	665	510	608	478	673	512	3,531	3,287
Public, city	954	835	116	89	105	85	118	88	615	573
Public, suburban	1,460	1,276	177	136	162	126	180	136	941	878
Public, town	447	391	54	41	50	39	55	42	288	269
Public, rural	1,008	881	122	94	112	88	124	94	650	605
Catholic, city	545	476	66	51	61	48	67	51	351	326
Catholic, suburban	313	273	38	29	35	27	38	29	202	188
Catholic, town	96	84	12	9	11	8	11	10	62	57
Other private, city	245	214	30	23	27	21	30	23	158	147
Other private, suburban	250	219	31	23	28	22	31	24	160	150
Other private, town	77	67	9	7	8	7	9	7	51	46
Other private, rural	82	71	10	8	9	7	10	8	53	48
South	8,121	7,096	985	754	902	709	1,000	759	5,234	4,874
Public, city	1,716	1,500	208	159	190	150	211	161	1,107	1,030
Public, suburban	2,170	1,896	264	201	241	190	267	203	1,398	1,302
Public, town	664	580	80	62	74	58	82	62	428	398
Public, rural	1,826	1,595	221	170	203	160	225	171	1,177	1,094
Catholic, city	491	429	60	46	55	41	60	46	316	296
Catholic, suburban	126	109	15	11	15	11	16	11	80	76
Catholic, town	38	34	5	4	4	3	5	4	24	23
Other private, city	463	405	56	43	51	41	57	43	299	278
Other private, suburban	397	347	48	37	44	34	49	37	256	239
Other private, town	121	106	15	11	13	11	15	11	78	73
Other private, rural	109	95	13	10	12	10	13	10	71	65

Table 13. Illustrative student sample allocation and expected yields for ninth-graders (full-scale study HSLS:09)—Continued

School stratum	Total		Hispanic		Asian		Black		Other	
	Sample	Respondents	Sample	Respondents	Sample	Respondents	Sample	Respondents	Sample	Respondents
West	4,278	3,740	518	398	475	373	528	401	2,757	2,568
Public, city	1,199	1,048	145	111	133	105	148	112	773	720
Public, suburban	1,398	1,221	170	129	155	123	172	131	901	838
Public, town	428	374	51	41	48	37	53	40	276	256
Public, rural	354	310	43	33	39	31	44	33	228	213
Catholic, city	191	167	23	18	21	17	24	18	123	114
Catholic, suburban	125	109	15	11	15	11	15	11	80	76
Catholic, town	38	34	5	4	3	3	5	4	25	23
Other private, city	327	286	40	30	36	27	40	31	211	198
Other private, suburban	104	92	12	10	11	9	13	10	68	63
Other private, town	32	28	4	3	4	3	4	3	20	19
Other private, rural	82	71	10	8	10	7	10	8	52	48

Full-Scale Study Teacher, High School Counselor, and Parent Samples

Analogous to the field test sample, one math and one science teacher will be selected for each ninth-grade student. Where sample students have more than one math or science teacher in fall 2009, one of the teachers will be randomly sampled. In addition, for each sample school there will be one sample high school counselor and one sample parent. In two-parent households, the parent most knowledgeable with the student's school situation and experience will be asked to participate.

c. Weighting, Variance Estimation, and Imputation

After data collection, survey data must go through several steps before analysis and reporting tasks can begin. Once data have been compiled and edited, survey weights will be computed, followed by variance estimation and imputation of missing data. This section provides a brief overview of each of these steps for the HSLS:09 full-scale study.

Weighting

Virtually all survey data are weighted before they can be used to produce reliable estimates of population parameters. While reflecting the selection probabilities of sampled units, weighting also attempts to compensate for practical limitations of a sample survey, such as differential nonresponse and undercoverage. Furthermore, by taking advantage of auxiliary information about the target population, weighting can reduce the variability of estimates. The weighting process essentially entails four major steps. The first step consists of the computation of *design* or *base weights*. In the second step, base weights will be adjusted for nonresponse, while in the third step nonresponse-adjusted weights will be further adjusted so that aggregate counts can match reported estimates for the target population. Finally, adjusted weights will go

through a series of quality control checks to detect extreme outliers and to prevent any computational as well as procedural errors.

The HSLS:09 multilevel and multicomponent design introduces significant complexity to the task of weighting. Cognizant of this complexity, every effort will be made to keep the resulting weights as simple and intuitive as possible. A minimum of two sets of weights will be required for the analysis of the HSLS:09 data: school weights and student weights. While the expectation is to secure the stated rates of response, when response rates fall below the accepted limit (both at unit and item levels), detailed nonresponse bias analysis will be conducted to measure the extent of the incurred bias and to identify effective methods for nonresponse adjustment.

Several methods have been suggested for measuring nonresponse bias. In the simplest form, this bias can be approximated temporally by comparing responses obtained from those who respond earlier in the data collection period against late respondents. The incurred bias due to nonresponse can be measured more systematically, however, as the difference between survey estimates and their respective target parameters—the values that would result if a complete census were conducted and all units responded. For instance, when estimating a population mean (μ) based on respondents only ($\bar{y}_R$) nonresponse bias can be expressed as

$$B(\bar{y}_R) = \bar{y}_R - \mu .$$

However, for variables that are available from the sampling frame, μ can be estimated by $\hat{\mu}$ without sampling error, in which case the bias in $\bar{y}_R$ can then be estimated by

$$\hat{B}(\bar{y}_R) = \bar{y}_R - \hat{\mu} .$$

B. Moreover, an estimate of the population mean based on respondents and nonrespondents can be obtained by

$$\hat{\mu} = (1 - \hat{\eta})\,\bar{y}_R + \hat{\eta}\,\bar{y}_{NR} .$$

C. where $\hat{\eta}$ is the weighted unit nonresponse rate, based on design weights prior to nonresponse adjustment. Consequently, the bias in $\bar{y}_R$ can then be estimated by

$$\hat{B}(\bar{y}_R) = \bar{y}_R - \hat{\mu} = \bar{y}_R - \left[(1 - \hat{\eta})\,\bar{y}_R + \hat{\eta}\,\bar{y}_{NR}\right] = \hat{\eta}\,(\bar{y}_R - \bar{y}_{NR}) .$$

That is, the estimate of the nonresponse bias is the difference between the mean for respondents and the mean for nonrespondents, multiplied by the weighted nonresponse rate, using the design weights prior to nonresponse adjustment. This basic approach will be used to measure bias in key survey estimates by relying on data that will be available for both respondents and nonrespondents.

As an attempt to reduce some of the bias due to nonresponse, when appreciable bias is detected at any level, design weights will be adjusted within cells indexed by variables that are deemed strong predictors of response status. To identify such variables, which typically include sampling stratification variables and indicators that can efficiently partition units into homogenous segments, classification procedures such as CHAID (Chi-square automatic interaction detection method) will be relied upon. CHAID is a hierarchical clustering algorithm that successively partitions units according to a categorical characteristic. The algorithm begins with all sample units as a whole and cycles over each predictor to find the optimal partition of

the units. The most significant predictor is identified, resulting in partitioning of units into smaller subsets. Next, the algorithm is applied to each partitioned subset of units to find further partitions using the remaining predictors. The process stops after a specified number of partitioning steps or if none of the partitions at a given step is found to be significant.

For HSLS:09 all weight adjustments—including those for nonresponse and poststratification—will be calculated using RTI International's generalized exponential model (GEM) software.[4] GEM is a raking procedure that is a generalization of the logic-type model, which has been proven to produce weights with less variability than what is achievable via traditional methods. GEM is superior to standard raking methods in two regards. First, it allows a much larger set of variables and their interactions to be used during the model development for nonresponse and raking adjustments, hence enabling the weighted data to mimic the distribution of the target universe with respect to a more comprehensive set of indices. Second, this desirable property is achieved while preventing the adjusted weights from becoming too extreme. That is, GEM produces study estimates that better represent the target universe without increasing variance of estimates significantly, which would otherwise reduce the power of statistical tests.

Variance Estimation

For variance estimation, sets of 200 balanced repeated replication (BRR) weights will be created for school and student samples. The BRR weights are appropriate for use in NCES's Data Analysis System (DAS) and do not affect the analysis weights used for point estimation. The BRR weighting process will replicate the full weighting process and will use procedures developed for a number of other studies, including ELS:2002 and the National Study of Postsecondary Faculty. In addition, analysis strata and primary sampling units (PSUs) created from the sampling PSUs will be included on the electronic code book for analysts wanting to use Taylor series variance estimation rather than BRR weights.

Imputation of Missing Data

Missing values due to item nonresponse will be imputed after the data are edited. Imputation will be performed for items commonly used to define analysis domains, items that are frequently used in crosstabulations, and items needed for weighting. Items from HSLS:09 that are subject to imputation will be imputed using a weighted sequential hot deck procedure.[5] By incorporating the sampling weights, this method of imputation takes into account the unequal probabilities of selection in the original sample while controlling the expected number of times a particular respondent's answer will be used as a donor.

3. Methods for Maximizing Response Rates

Procedures for maximizing response rates at the institution and respondent levels are based on successful experiences on predecessor and other similar studies. In this section methods for maximizing response rates for school recruitment as well as for students, parents, and school staff are discussed.

---

[4] Folsom, R.E., and A.C. Singh (2000). "The Generalized Exponential Model for Sampling Weight Calibration for Extreme Values, Nonresponse, and Poststratification." *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, pp. 598–603.

[5] Iannacchione, V.G. (1982). "Weighted Sequential Hot Deck Imputation Macros." *In Proceedings of the Seventh Annual SAS User's Group International Conference* (pp.759–763). Cary, NC: SAS Institute, Inc.

School Recruitment. Achieving high school participation rates on voluntary research studies has proven increasingly difficult in recent years. Recent experience has shown that many schools already feel burdened by mandated "high stakes" testing and, at the same time, are hampered by fiscal and staffing constraints. Moreover, there will be roadblocks not only at the school, but also at the district level, where research studies must sometimes comply with stringent requirements to submit formal and detailed applications similar to those one would submit to an IRB before individual schools can even be contacted. The keystone of the plan to work with school districts and schools is to demonstrate the importance of the study while maintaining flexibility in negotiations with school districts and schools.

Immediately after drawing the sample, recruitment for the field test will commence. Sample materials to be sent to states, districts, and schools are provided in appendix A. Succinct yet compelling advance materials will be sent to the school districts and schools to introduce the study. Within a few days of receiving the materials, a trained recruiter will contact the school district or school to discuss their participation in the study. Recruiters are hired for their knowledge, skill, and articulation with the proven ability to develop relationships with district and school contacts that will foster participation and persist throughout the in-school follow-ups for the longitudinal study.

As much as possible, burden will be shifted from the school to research staff. Possible ways of shifting the burden include scheduling contacts or survey administrations to best fit the school calendar, mailing consent forms directly to parents, providing compensation for time/help completing forms, offering a session administrator to come to the school to compile sampling information, and having a session administrator coordinate all aspects of survey day (e.g., posting reminders, processing consents, and gathering students). These options have proven helpful on similar studies to gain cooperation in schools that expressed scheduling, burden, or staffing concerns.

One of the key factors to a successful recruitment period is time. A task force convened in 2004 to help NCES brainstorm ways to improve school response rates in their international studies recommended that all recruitment activities begin at least 1 year prior to the start of data collection. Though there will not be a full year to recruit schools for the field test, the request for approval to begin recruitment for both the field test and the main study will afford the benefit of having sufficient time to recruit for the main study.

It is worth noting that the proposed sample design will not cluster schools at the district level. This will mitigate the undesirable situation of losing clusters of schools from sample districts that opt not to participate in this study.

An incentive experiment is proposed at the school level for the field test to help offset some of the challenges associated with obtaining school cooperation. A successful incentive program can greatly reduce labor costs associated with school recruitment and refusal conversion efforts. For the field test, the experiment compares the effect of a $500 technology allowance against no incentive. All schools within a given district would receive the same incentive. The technology allowance would be in the form of a check written to the school that can be used at the school's discretion, though HSLS:09 field staff will recommend that it be used toward technology for the school to align with the focus of the study.

The small number of responding schools that will be involved in the field test, 50 to 55 schools, calls for an uncomplicated design protocol. As such, a simple design is proposed

whereby sample schools in each of the five states are randomly partitioned into control and experimental groups. All schools within the experimental group will be offered an incentive of a $500 technology allowance for participation in HSLS:09. Schools in the control group, however, will be offered no incentive for their participation in the field test.

Student. Ensuring a high student response at each school begins several weeks prior to the student session. Session administrators will work closely with the school coordinators to coordinate the logistics of the sessions and notify students about the sessions. Because the sampled students are not selected by classroom and are dispersed across multiple classes, there is a heavy burden on the school coordinator to inform students about the session, distribute parental consent materials, and ensure that the students arrive at the prescribed location at the scheduled date and time. Session administrators will assume as much of this burden as is possible and permissible by the school.

From past experience, ensuring that students are made aware of the session is the most critical aspect of making sure they arrive at the session at the scheduled time. Despite receiving the consent form to take home, students do not necessarily distinguish the form from other materials they take home, and they often forget about the session without frequent reminders. To help remind students about the sessions, the study will implement options such as distributing postcard reminders a day or two prior to the session, notifying the teachers of selected students, asking the school coordinator to make an announcement on the PA system, and having the session administrator visit a few days prior to the session and convene a brief meeting of the student sample members to encourage participation. Parent contacting information will be collected from each school from which the parent survey will be conducted. If phone numbers are provided, the session administrator will contact parents a day or two prior to the session to remind the students when they should arrive.

Each week, project staff will conduct group strategy calls with the session administrators to discuss the status of the schools with test dates scheduled for the coming 2 weeks. The purpose of these conference calls is to learn about the preparedness of each school for the student session, identify any concerns about anticipated response rate or computer capabilities at the school, provide a forum for brainstorming solutions to anticipated problems, and share success stories and lessons learned from other schools. Project staff will follow up frequently with SAs who report problems or concerns with the preparations for student sessions at particular schools.

Plans for student incentives in the field test were described in section A9. Each participating student will receive $10. Seniors are offered this incentive as a motivation to attend the fall data collection session. Based on the experience of NAEP and other studies, seniors are presumed to be more apathetic toward participation in additional testing activities, and often have competing claims, such as jobs, for their time. At the same time, it seems prudent to treat 9th and 12th graders in the same way within the same school. Student-level incentives also aid in motivating school officials to participate by giving something back to the students.

For the main study, when only 9th grade students will participate, an educational “goodie bag” will be offered. Participating students will receive a drawstring backpack filled with educational items. The total value of the token incentive for the main study will be no more than $5 per unit.

To ensure that there is a sufficient number of students participating in the field test and thus sufficient data to construct assessment parameters precisely, 2 additional students in grade 9 and 3 additional students in grade 12 will be sampled at each school, for a total of 29 students sampled at 9th grade and 30 students sampled at 12th grade at each school. Student response will be closely monitored, and if problems are encountered that would justify further review of the incentive levels, NCES may take the issue of incentive back to OMB.

Parent. There will be several opportunities to interact with parents to encourage their participation in the study. The parental consent form will be sent home with the students several weeks before the student session, and the letter will mention that the parent interview is forthcoming. Parent contacting information will be collected from the school after the student sample is identified. A letter will be sent to the parent via e-mail and Federal Express to initiate the parent interview, providing a URL and credentials for the web instrument and a telephone number that can be used for a telephone interview. If a telephone number is available, the SA will contact the parent to remind him or her of the student session, and will take the opportunity to build a relationship with the parent and encourage participation from both the student and parent. Parents who do not complete the web instrument will be followed up via CATI. Unless the field test (all electronic) conclusively suggests this is not needed, in the main study, paper-and-pencil versions of the questionnaire will be available for parents who do not have a telephone or Internet access. The parent interview will be translated into Spanish to accommodate limited English proficient and nonproficient parents.

There is no precedent for offering an incentive to complete the parent questionnaire. Thus, no parent incentive is included in our budget for the HSLS:09.

School Staff (School Administrators, Counselors, Teachers). School staff will receive a letter to initiate their questionnaire about 3 weeks prior to the student session. The session administrator will work with the school coordinator to prompt school staff to complete their interview. While at the school, the SA will prompt for any outstanding staff questionnaires. If the questionnaires still have not been completed by 1 week after the session(s) are complete in the school, CATI follow-up will commence.

If the field test all-electronic experience warrants it, in the main study, teachers will have an option to complete a paper-and-pencil version of the questionnaire. Past experience has demonstrated the need for a teacher-level incentive to achieve high response rates and many schools have required that teacher compensation be commensurate with their hourly wage. Thus, the proposed teacher incentive is a $25 base with an additional $5 for each additional class on which they are asked to report.

4. Individuals Consulted on Statistical Design

A number of individuals have consulted with NCES on the sampling design and recruitment plans for HSLS:09. Members of the Technical Review Panel are listed in section A8 of this document. In addition, Dr. Laura LoGerfo, Research Scientist, and Dr. Jeffrey Owings, Associate Commissioner for the Elementary/ Secondary and Library Studies Division, at NCES have reviewed and approved the statistical aspects of the study. Other statistical reviewers at NCES include Marilyn Seastrom, Chief Statistician, and the following statistical program staff: John Wirt, Tate Gould, and Michael Ross. Table 14 provides the names of additional consultants on statistical aspects of HSLS:09.

Exhibit 2. Preliminary outline for HSLS:09 Base-Year Field Test Report

Executive Summary
Introduction
Chapter 1 Field Test Preparation: Sampling and Instrumentation
- 1.1 Sample Design and Selection
  - 1.1.1 Selection of the Field Test States
  - 1.1.2 School Sampling
  - 1.1.3 Student Sampling
  - 1.1.4 Sampling Teachers, Administrators, and Counselors
- 1.2 Instrumentation
  - 1.2.1 Mathematics Assessment
  - 1.2.2 Student Questionnaire
  - 1.2.3 Parent Questionnaire
  - 1.2.4 Teacher Questionnaire
  - 1.2.5 Administrator Questionnaire
  - 1.2.6 Counselor Questionnaire

Chapter 2 Securing Cooperation
- 2.1 Securing Endorsements
- 2.2 Securing State Cooperation
- 2.3 Securing District, Diocese, and School Cooperation
- 2.4 School-Level Response Results
  - 2.4.1 Analysis of School Response Rates
  - 2.4.2 Responses to Incentives and Burden
- 2.5 Obtaining Parental Consent
- 2.6 Recommendations for Main Study

Chapter 3 Data Collection
- 3.1 Recruitment and Training of Data Collection Staff
  - 3.1.1 Assessors
  - 3.1.2 School Recruiters
  - 3.1.3 Help Desk Staff and Interviewers
- 3.2 In-School Student Survey Procedures and Results
- 3.3 Procedures and Results for Surveys of Other School Populations
  - 3.3.1 Teachers
  - 3.3.2 School Administrators
  - 3.3.3 School Counselors
- 3.4 Parent Survey Procedures and Results
- 3.5 Recommendations for Main Study

Chapter 4 Analysis of Student Survey Results
- 4.1 Mathematics Assessments
  - 4.1.1 Choice of Item Pool
  - 4.1.2 Timing and Completion Rates
  - 4.1.3 Item Performance
    - 4.1.3.1 Classical Item Analysis
    - 4.1.3.2 Item Response Theory
  - 4.1.4 Reliability and Factor Structure
  - 4.1.5 Differential Item Functioning
  - 4.1.6 Selecting Items
    - 4.1.6.1 Measuring Change Over Time
    - 4.1.7.1 Comparing to Prior and Ongoing Studies
- 4.2 Student Questionnaire
  - 4.2.1 Editing and Retrieval of Critical Items
  - 4.2.2 Item Nonresponse
  - 4.2.3 Inter-item Consistency
  - 4.2.4 Logical Consistency of Responses to Filter and Dependent Questions
  - 4.2.5 Response Variation by Item Position in Questionnaire
- 4.3 Recommendations for Main Study

Chapter 5 Analysis of Teacher, School Administrator, and School Counselor Survey Results
- 5.1 Teacher Survey Responses
  - 5.1.1 Item Nonresponse
  - 5.1.2 Inter-Item Consistency
  - 5.1.3 Logical Consistency of Responses to Filter and Dependent Questions
- 5.2 School Administrator Survey Responses
  - 5.2.1 Item Nonresponse
  - 5.2.2 Inter-Item Consistency
  - 5.2.3 Logical Consistency of Responses to Filter and Dependent Questions
- 5.3 School counselor Responses
  - 5.3.1 Item Nonresponse
  - 5.3.2 Inter-Item Consistency
  - 5.3.3 Logical Consistency of Responses to Filter and Dependent Questions
- 5.4 Recommendations for Main Study

Chapter 6 Analysis of Parent Survey Results
- 6.1 Item Nonresponse
- 6.2 Inter-Item Consistency
- 6.3 Logical Consistency of Responses to Filter and Dependent Questions
- 6.4 Comparisons of Parent and Student Responses
- 6.5 Reliability of Parent Responses
- 6.6 Recommendations for Main Study

Chapter 7 Survey Control System and Data Processing
- 7.1 System Design, Development, and Testing
- 7.2 Data Capture
- 7.3 Data Processing and File Preparation
- 7.4 Recommendations for Main Study

Chapter 8 Conclusions
References
Appendixes (Instruments, Sampling Specifications, Mailout Materials and Forms, TRP Membership)
Listing of NCES Working Papers to Date

Table 14. Consultants on statistical aspects of HSLS:09

Name	Affiliation	Telephone
James Chromy	RTI	(919) 541-7019
Steven J. Ingels	RTI	(202) 974-7834
Shijie Chen	RTI	(202) 974-7820
Peter H. Siegel	RTI	(919) 541-5902
Daniel J. Pratt	RTI	(919) 541-6615
John Riccobono	RTI	(919) 541-7006
Deborah Herget	RTI	(919) 485-7793
Gary Phillips	AIR	(202) 403-6916
Steve Leinwand	AIR	(202) 403-6926

Reference

Folsom, Ralph E., Potter, Frank J., and Williams, S. Rick (1987). Notes on a Composite Size Measure for Self-Weighting Samples in Multiple Domains. *Proceedings of the Section on Survey Research Methods* (pp. 792-796). The American Statistical Association.