A. 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. The needed respondent samples will be selected from all public and private schools with 9th and 12th 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 the handicapped, 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

RTI plans to use NCES' latest Common Core of Data (CCD:2005–2006) as the public school sampling frame and Private School Survey (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 9th 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 12th 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 probability-proportionate-to-size (PPS) methodology for which a composite size measure methodology developed by RTI statisticians (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), maintains near equal sampling weights for students within each domain, and results 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 will be 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 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. Should enrollment information be unavailable for certain schools, RTI will impute the needed enrollment counts to the median value of the enrollment for ninth graders within race/ethnicity categories in each school stratum. We expect to select the full-scale and field test samples of schools in January 2008, 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 for RTI to include in this investigation. Should such a copy be made available to RTI, it will be used for sample refreshing and related quality control activities.

Obviously, a sample size larger than 800 schools will be necessary to compensate for the anticipated nonresponse and ineligibility. As per NCES standards, we will target a weighted response rate of at least 70 percent at the school level. 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 our 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, $1{,}143 \times 1.04$) schools. Moreover, based on ELS:2002 response rates, we expect 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.

We will closely monitor the school recruitment activities and release additional schools as needed to ensure that we reach our goal of 800 participating schools. 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, RTI will use data from QED 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 us to conduct 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 50 participating schools for the field test study. This sample will be divided into 40 public and 10 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 11, we will select a slightly larger sample of 84 schools 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	84	50	17	10	17	10	16	10	17	10	17	10
Public, total	67	40	14	8	13	8	13	8	14	8	13	8
Public, city	19	11	4	3	3	2	4	2	4	2	4	2
Public, suburban	25	15	5	3	5	3	5	3	5	3	5	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	8	5	1	1	2	1	2	1	1	1	2	1
Catholic, city	4	2	1	1	1	1	1	0	0	0	1	0
Catholic, suburban	2	2	0	0	0	0	1	1	0	0	1	1
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	9	5	2	1	2	1	1	1	2	1	2	1
Other private, city	4	2	0	0	1	0	1	1	1	0	1	1
Other private, suburban	3	2	0	0	1	1	0	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, we will ask for paper lists to be faxed to a fax machine in a locked room at RTI. For data security reasons, we will request that paper lists not be mailed. RTI will ask each sample school 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 or not 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.

As requested by NCES, 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, and we will work with the schools to determine if any accommodations can be made 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] We will proceed with selecting sample students when we have either confirmed that the list received is correct or received a corrected list. Students will be sampled on a flow basis as student lists are received. We will stratify the lists by race/ethnicity and select a systematic sample of students from the resulting lists. For schools that provide paper lists, RTI will use a two-stage process that we have 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 25 students from the 9th grade and 25 students from the 12th grade will be selected in each of the 50 sample schools, for a total of 1,250 (or, 50×25) students in each grade. Based on the ELS:2002 eligibility and response rates of 95 and 92 percent, respectively, this will result in a sample of 1,093 ($1{,}250 \times 0.95 \times 0.92$) responding students in each grade. 12 shows an expected allocation of the sample and responding students for each grade, by school and student characteristics, overall and for

[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.

each of the five participating states. During the recruitment process, we will ask schools when their student lists will be ready; however, we anticipate requesting lists and drawing student samples on a flow basis for the field test between August and November of 2008.

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

School Stratum	Total		Hispanic		Asian		Black		Other	
	Sample	Respondent	Sample	Respondent	Sample	Respondent	Sample	Respondent	Sample	Respondent
Total	1,250	1,093	231	198	50	50	219	198	750	647
Public, city	275	242	50	44	11	11	49	44	165	143
Public, suburban	375	323	70	59	15	15	70	60	220	189
Public, town	125	110	20	19	5	5	20	18	80	68
Public, rural	225	198	41	36	9	9	40	36	135	117
Catholic, city	50	44	10	8	2	2	8	8	30	26
Catholic, suburban	50	44	10	8	2	2	8	8	30	26
Catholic, rural	25	22	5	4	1	1	4	4	15	13
Other private, city	50	44	10	8	2	2	8	8	30	26
Other private, suburban	50	44	10	8	2	2	8	8	30	26
Other private, rural	25	22	5	4	1	1	4	4	15	13
New York	250	218	47	39	10	10	43	39	150	130
Public, city	75	66	14	12	3	3	13	12	45	39
Public, suburban	75	64	14	11	3	3	14	12	44	38
Public, town	25	22	4	4	1	1	4	3	16	14
Public, rural	25	22	5	4	1	1	4	4	15	13
Catholic, city	25	22	5	4	1	1	4	4	15	13
Other private, rural	25	22	5	4	1	1	4	4	15	13
California	250	219	46	40	10	10	44	40	150	129
Public, city	50	44	9	8	2	2	9	8	30	26
Public, suburban	75	65	14	12	3	3	14	12	44	38
Public, town	25	22	4	4	1	1	4	4	16	13
Public, rural	50	44	9	8	2	2	9	8	30	26
Catholic, city	25	22	5	4	1	1	4	4	15	13
Other private, suburban	25	22	5	4	1	1	4	4	15	13
Florida	250	219	46	40	10	10	44	40	150	129
Public, city	50	44	9	8	2	2	9	8	30	26
Public, suburban	75	65	14	13	3	3	14	12	44	37
Public, town	25	22	4	3	1	1	4	4	16	14
Public, rural	50	44	9	8	2	2	9	8	30	26
Catholic, suburban	25	22	5	4	1	1	4	4	15	13
Other private, city	25	22	5	4	1	1	4	4	15	13
Illinois	250	218	46	39	10	10	44	39	150	130
Public, city	50	44	9	8	2	2	9	8	30	26
Public, suburban	75	64	14	11	3	3	14	12	44	38
Public, town	25	22	4	4	1	1	4	3	16	14
Public, rural	50	44	9	8	2	2	9	8	30	26
Catholic, rural	25	22	5	4	1	1	4	4	15	13
Other private, suburban	25	22	5	4	1	1	4	4	15	13

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

School Stratum	Total		Hispanic		Asian		Black		Other	
	Sample	Respondent	Sample	Respondent	Sample	Respondent	Sample	Respondent	Sample	Respondent
Texas	250	219	46	40	10	10	44	40	150	129
Public, city	50	44	9	8	2	2	9	8	30	26
Public, suburban	75	65	14	12	3	3	14	12	44	38
Public, town	25	22	4	4	1	1	4	4	16	13
Public, rural	50	44	9	8	2	2	9	8	30	26
Catholic, suburban	25	22	5	4	1	1	4	4	15	13
Other private, city	25	22	5	4	1	1	4	4	15	13

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, we will randomly sample one of the teachers. 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. Our 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, we will follow the NELS:88/ELS:2002 procedures to ask 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 9th 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% change in proportions across waves of the study;
- detect a 5% change in means;
- produce relative standard errors of 10% 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% 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 RTI will receive 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 we expect to achieve the stated response and eligibility rates, an early identification of low sample yields will be vital in making sure we can adjust appropriately to reach our target yields. 13 shows a possible student sample allocation and yield for the HSLS:09 full-scale study. We anticipate requesting student lists and drawing student samples on a flow basis between August and November of 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 9th-grade student. Where sample students have more than one math or science teacher in fall 2009, we will randomly sample one of the teachers. 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.

We expect that a number of sample students will have the same math and science teachers; however, in most schools the above design can include virtually all eligible teachers. As such, an alternative approach under consideration involves conducting a census of ninth-grade teachers, instead of using a linked student-teacher design. Our survey protocols will be developed in such a way that either approach could be implemented without any ramifications on other aspects of this study.

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. In this section we provide 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, RTI will make every effort 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 we expect to secure the stated rates of response, when response rates fall below the accepted limit (both at unit and item levels) we will carry out detailed nonresponse bias analysis 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

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

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

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

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. In order to identify such variables, which typically include sampling stratification variables and indicators that can efficiently partition units into homogenous segments, we will rely on classification procedures such as CHAID (Chi-square automatic interaction detection method). 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'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, we will create sets of 200 balanced repeated replication (BRR) weights 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 (NSOPF). In addition, analysis strata and primary sampling units (PSUs) created from the sampling PSUs will be included on the electronic code book (ECB) 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 RTI's weighted sequential hot deck procedure.[5] By incorporating the sampling weights, this method of imputation

---

[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.

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

Our procedures for maximizing response rates at the institution and respondent levels are based on our successful experience on predecessor and other similar studies. Because this submission to OMB covers only the sampling and school recruitment components of the HSLS:09, discussions of maximizing response rates among students, parents, teachers, and school administrators will take place in the next submission. However, it is worth noting that our 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.

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.

Obtaining agreement from schools to participate in voluntary research efforts grows increasingly difficult each year as schools are reluctant to add to the state accountability efforts and demands of the "No Child Left Behind" legislation. In addition to roadblocks anticipated at the school level, we also expect the need to comply with stringent requirements at the district level prior to contacting schools as well as the need to deal with reluctance from some districts to participate in HSLS:09. To mitigate these challenges for the field test and main study of HSLS:09, we will build on our experience conducting past studies and will, throughout the recruitment process, look for new and innovative approaches to obtaining school cooperation.

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 we will not have a full year to recruit schools for the field test, our request for approval to begin recruitment for both the field test and the main study will afford us the benefit of having sufficient time to recruit for the main study.

Offering incentives at the school level has not been attempted on the predecessor studies to HSLS:09. However, we believe that testing the effectiveness of a school-level incentive in the field test will determine whether it will help persuade schools to participate in the full-scale study. As described in section A9, we propose an experiment comparing the effect of a $500 technology allowance against no incentive. Our experimental design, an overview of which is provided next, will be developed such that all schools sampled within a given district receive the same incentive treatment.

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

Many of the procedures developed for HSLS:09 field test and full-scale study have been tested in other longitudinal surveys of students and young adults. However, it is important to carefully examine all aspects of the study design and procedures in the field test to ensure that the most efficient and effective sampling and data collection approaches will be used for the full-scale study in 2009. The main focus of the field test is to collect enough assessment data to perform reliable tests of the items.

Special features of the 2008 field test include the use of computer adaptive testing for the first time for the student component. Plans for field test analyses of questionnaire and test data and of procedures can readily be inferred from the field test report preliminary outline, provided in 14.

4. Individuals Consulted on Statistical Design

A number of individuals have consulted with NCES and RTI on the sampling design and recruitment plans for the 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. Section A15 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 Science Assessment
  - 1.2.3 Student Questionnaire
  - 1.2.4 Parent Questionnaire
  - 1.2.5 Teacher Questionnaire
  - 1.2.6 Administrator Questionnaire
  - 1.2.7 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 and Science 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
- 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
Appendices (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) 728-1962
Mansour Fahimi	RTI	(301) 230-4675
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 Ferrara	AIR	(202) 403-5431