

# High School and Beyond 2021 (HS&B:21) Base-Year Full-Scale Study Data Collection

OMB# 1850-0944 v.7

## Supporting Statement Part B

Submitted by  
National Center for Education Statistics  
U.S. Department of Education

February 2020  
Revised August 2020

### Table of Contents

---

B.1 Respondent Universe.....	1
B.2 Procedures for the Collection of Information.....	2
B.3 Methods to Secure Cooperation, Maximize Response Rates, and Deal with Nonresponse.....	9
B.4 Tests of Methods and Procedures.....	15
B.5 Reviewing Statisticians and Individuals Responsible for Study Design and Conduct.....	16

## B. Collection of Information Employing Statistical Methods

Part B of this submission presents information on the statistical methods employed for the HS&B:21 base-year full-scale (BYFS) study sampling techniques and recruitment.

### B.1 Respondent Universe

The High School and Beyond 2021 study (HS&B:21) will follow a nationally-representative sample of ninth-grade students from the start of high school in the fall of 2021 to the spring of 2025 when most will be in twelfth grade. The study sample will be freshened in 2025 to create a nationally representative sample of twelfth-graders. A high school transcript collection and additional follow-up data collections beyond high school are also planned. The sample of ninth-grade students selected in the fall of 2021 is referred to the ninth-grade cohort while the sample of students enrolled in twelfth-grade in spring of 2025 is referred to as the twelfth-grade cohort.

The target populations for BYFS consists of ninth-grade students in public and private schools in the 50 United States and District of Columbia as of fall 2021<sup>1</sup>. Excluded from the target universe are special education schools, area vocational schools that do not enroll students directly, Department of Defense (DoD) schools outside of the US, and schools associated with temporary housing such as correctional facilities and treatment centers.

BYFS will be conducted during the 2021-22 school year, with recruitment initiated in August 2019<sup>2</sup> and data collection to begin in September 2021. The BYFS is designed to select a nationally representative sample of schools offering grade 9 instruction and a nationally representative sample of students enrolled in grade 9. The BYFS school population consists of regular public schools, including state department of education schools, that include 9th grade; Bureau of Indian Education schools that include 9th grade; and Catholic and other private schools that include 9th grade. It excludes the following types of schools:

- DoD Education Activity schools outside of the United States,
- Schools associated with correctional facilities, treatment facilities, hospitals, and other temporary housing facilities,
- Area vocational schools that do not enroll students directly, and
- Special education schools.<sup>3</sup>

The HS&B:21 BYFS employs a multi-stage sampling design with schools selected in the first stage and students to be selected, within schools, at the second stage. Schools were selected using probability proportional to size sampling within school sampling strata.

Students will be selected using simple random sampling within student sampling strata within schools. The school sampling frame was constructed from the 2017-18 Common Core of Data (CCD 2017-18) and the 2015-16 Private School Universe Survey (PSS 2015-16)<sup>4</sup> and includes 28,688 schools that report offering ninth-grade instruction to at least 1 student. An initial sample of 1,373 schools was selected with the goal of achieving 920 participating schools. A sample of approximately 26,000 students from an estimated 920 participating schools is estimated to produce 20,995 participating students enrolled in grade 9.

The 28,688 schools in the sampling frame were explicitly stratified using combinations of the categories defined by the cross-classification of the following characteristics<sup>5</sup>:

<sup>1</sup> While the BYFS sample will include only ninth-grade students, the base-year field test sample included both ninth- and twelfth-grade students to prognosticate the progression that will be observed when reassessing the sample ninth-grade students three years later, when most will be in 12th grade.

<sup>2</sup> The main study data collection period was originally scheduled for fall 2020. Due to COVID-19, the data collection was postponed to fall 2021. Recruitment for the main study data collection began in August 2019 for data collection in fall 2020. Upon the decision to postpone data collection, recruitment paused and will resume in October 2020 to recruit schools for the data collection in fall 2021.

<sup>3</sup> A special education school is a public elementary/secondary school that focuses on educating students with disabilities and adapts curriculum, materials, or instruction for the students served.

<sup>4</sup> A new PSS file from the 2017-2018 period was released after the selection of the HS&B:21 full-scale school sample. Sample freshening, an attempt to add new Catholic schools appearing in the 2017-2018 PSS to the existing HS&B:21 full-scale school sample, was considered but a determination was made that the value of sample freshening was minimal and did not offset the added sampling complexity required to implement the freshening. Based on a similar sample freshening experience associated with the base-year full-scale Middle Grades Longitudinal Study of 2017 school sample, we believe that few schools would be added via the sample freshening process. Furthermore, as the weight adjustment process will use the 2017-2018 PSS to create control totals for private schools and students attending private schools, the analysis weights for the participating Catholic schools from the current school sample will be adjusted so that the participating Catholic schools and students attending those schools represent the Catholic schools and students represented in the 2017-2018 PSS.

<sup>5</sup> Some of the characteristics were combined to produce the 27 school sampling strata.

- school type (public, Catholic, other private),
- region (Northeast, Midwest, South, West),
- locale (city, suburb, town, rural), and
- public school type (Charter, Magnet, Virtual, not Charter/Magnet/Virtual).

The distributions of the numbers of schools in the school sampling frame, the initial school sample, and school participation goals are shown by school sampling strata in table 1.

**Table 1. HS&B:21 BYFS School Sample Allocation**

School Type	Census Region	Locale	Participation Goals	School Frame Count	Total Selected School Sample	Initial School Sample	School Reserve Sample
<b>Total</b>			<b>920</b>	<b>28,688</b>	<b>2,654</b>	<b>1,373</b>	<b>1,281</b>
Public-Charter			67	2,466	198	99	99
Public-Magnet			67	971	198	99	99
Public-Virtual			67	623	150	99	51
Public-Not charter, magnet, or virtual	Northeast	City	17	717	51	26	25
Public-Not charter, magnet, or virtual	Northeast	Suburb	30	1,162	89	45	44
Public-Not charter, magnet, or virtual	Northeast	Town	8	247	24	12	12
Public-Not charter, magnet, or virtual	Northeast	Rural	16	775	48	24	24
Public-Not charter, magnet, or virtual	Midwest	City	20	713	59	30	29
Public-Not charter, magnet, or virtual	Midwest	Suburb	29	984	86	43	43
Public-Not charter, magnet, or virtual	Midwest	Town	16	928	48	24	24
Public-Not charter, magnet, or virtual	Midwest	Rural	28	2,697	83	42	41
Public-Not charter, magnet, or virtual	South	City	32	1,097	95	48	47
Public-Not charter, magnet, or virtual	South	Suburb	44	1,250	130	65	65
Public-Not charter, magnet, or virtual	South	Town	24	1,004	71	36	35
Public-Not charter, magnet, or virtual	South	Rural	50	2,857	148	74	74
Public-Not charter, magnet, or virtual	West	City	36	960	107	54	53
Public-Not charter, magnet, or virtual	West	Suburb	34	950	101	51	50
Public-Not charter, magnet, or virtual	West	Town	24	632	60	36	24
Public-Not charter, magnet, or virtual	West	Rural	32	1,311	95	48	47
Catholic	Northeast		34	265	101	51	50
Catholic	Midwest		32	336	95	48	47
Catholic	South		30	258	89	45	44
Catholic	West		32	172	80	48	32
Other private	Northeast		30	1,058	89	45	44
Other private	Midwest		34	854	101	51	50
Other private	South		55	2,442	163	82	81
Other private	West		32	959	95	48	47

## B.2 Procedures for the Collection of Information

HS&B:21 will collect data from high school students and their parents, math teachers, guidance counselors, and school administrators. Data will be collected from ninth graders in the fall of 2021 as they begin high school and again in the spring of 2025 when most students in the sample will be seniors at the end of their high school career. Collecting data at these time points from students, parents, teachers, counselors, and administrators, with high school transcripts collected after high school, will culminate in a rich data set that will provide educators, policymakers, and researchers with information about transitions, outcomes, and experiences in multiple contexts.

### Full-Scale School Sample

Prior to selection of the school sample, schools were sorted by locale (city, suburban, town, rural), school grade configuration, and school size measure within each of the explicit school strata so that approximate proportionality across locale and school configuration was preserved.

Declining response rates are a concern for any research study, and some of the recent school-based NCES longitudinal studies achieved response rates lower than a desired target of 75 percent. For example, the school response rate for the High School Longitudinal Study of 2009 (HSL:09) was 56 percent and the school response rate for the Early Childhood Longitudinal Study Kindergarten Class of 2010-11 (ECLS-K:2011) was 63 percent.

Nevertheless, to be conservative, the HS&B:21 BYFS sampling plan is designed to be flexible so that the study can better achieve school participation targets even if eligibility and response rates are lower than anticipated. The school sampling process is designed to achieve 920 participating schools (641 public, 128 Catholic, and 151 other private) distributed over 27 school sampling strata. We selected 2,654 schools using stratified probability proportional to size sampling, from which an initial simple random sample of 1,373 schools was selected within school strata. This subset of 1,373 schools formed the initial set of schools that are being pursued for recruitment, starting in August 2019, into the BYFS. The remaining schools were designed provide a reserve sample from which additional schools could be sampled and pursued for recruitment.

The ability to recruit schools for the field test was assessed in November of 2019, approximately 5 weeks prior to the end of the field test. The overall school participation in the field test was below 50%. This assessment led to a recommendation to select schools from the full-scale school reserve sample and release them for recruitment for the full-scale study. In January of 2020, a stratified random sample of 467 schools was selected from the 1,281 schools in the school reserve sample. The additional sample of 467 schools were added to the set of schools being pursued for recruitment into the full-scale study. The distribution of the sample of 467 schools across the school sampling strata is provided in table 2.

**Table 2. HS&B:21 BYFS School Sample Allocation Updated to Include Sample from Reserve**

School Type	Census Region	Locale	Total Selected School Sample	Initial School Sample	School Reserve Sample	January 2020 Sample from Reserve	Total Sample In Recruitment
<b>Total</b>			<b>2,654</b>	<b>1,373</b>	<b>1,281</b>	<b>467</b>	<b>1,840</b>
Public-Charter			198	99	99	35	134
Public-Magnet			198	99	99	35	134
Public-Virtual			150	99	51	35	134
Public-Not charter, magnet, or virtual	Northeast	City	51	26	25	8	34
Public-Not charter, magnet, or virtual	Northeast	Suburb	89	45	44	15	60
Public-Not charter, magnet, or virtual	Northeast	Town	24	12	12	4	16
Public-Not charter, magnet, or virtual	Northeast	Rural	48	24	24	8	32
Public-Not charter, magnet, or virtual	Midwest	City	59	30	29	10	40
Public-Not charter, magnet, or virtual	Midwest	Suburb	86	43	43	15	58
Public-Not charter, magnet, or virtual	Midwest	Town	48	24	24	8	32
Public-Not charter, magnet, or virtual	Midwest	Rural	83	42	41	14	56
Public-Not charter, magnet, or virtual	South	City	95	48	47	16	64
Public-Not charter, magnet, or virtual	South	Suburb	130	65	65	23	88
Public-Not charter, magnet, or virtual	South	Town	71	36	35	12	48
Public-Not charter, magnet, or virtual	South	Rural	148	74	74	26	100
Public-Not charter, magnet, or virtual	West	City	107	54	53	18	72
Public-Not charter, magnet, or virtual	West	Suburb	101	51	50	17	68
Public-Not charter, magnet, or virtual	West	Town	60	36	24	12	48
Public-Not charter, magnet, or virtual	West	Rural	95	48	47	16	64
Catholic	Northeast		101	51	50	17	68
Catholic	Midwest		95	48	47	16	64
Catholic	South		89	45	44	15	60
Catholic	West		80	48	32	16	64
Other private	Northeast		89	45	44	15	60
Other private	Midwest		101	51	50	17	68
Other private	South		163	82	81	28	110
Other private	West		95	48	47	16	64

In September of 2020, the remaining 814 schools in the reserve were added to the set of schools being pursued for recruitment into the full-scale study. It was deemed necessary to release the remaining schools due to the changing nature of school attendance in the 2020-21 school year due to COVID-19 and the need to have ample time to recruit all possible schools to achieve the 920 yield for the fall 2021 data collection. The distribution of the sample of 814 schools across the school sampling strata is provided in table 3.

**Table 3. HS&B:21 BYFS School Sample Allocation Updated to Include Samples from Reserve**

School Type	Census Region	Locale	Total Selected School Sample	Initial School Sample	School Reserve Sample	January 2020 Sample from Reserve	September 2020 Sample from Reserve	Total Sample In Recruitment
<b>Total</b>			<b>2,654</b>	<b>1,373</b>	<b>1,281</b>	<b>467</b>	<b>814</b>	<b>2,654</b>
Public-Charter			198	99	99	35	64	198
Public-Magnet			198	99	99	35	64	198
Public-Virtual			150	99	51	35	16	150
Public-Not charter, magnet, or virtual	Northeast	City	51	26	25	8	17	51
Public-Not charter, magnet, or virtual	Northeast	Suburb	89	45	44	15	29	89
Public-Not charter, magnet, or virtual	Northeast	Town	24	12	12	4	8	24
Public-Not charter, magnet, or virtual	Northeast	Rural	48	24	24	8	16	48
Public-Not charter, magnet, or virtual	Midwest	City	59	30	29	10	19	59
Public-Not charter, magnet, or virtual	Midwest	Suburb	86	43	43	15	28	86
Public-Not charter, magnet, or virtual	Midwest	Town	48	24	24	8	16	48
Public-Not charter, magnet, or virtual	Midwest	Rural	83	42	41	14	27	83
Public-Not charter, magnet, or virtual	South	City	95	48	47	16	31	95
Public-Not charter, magnet, or virtual	South	Suburb	130	65	65	23	42	130
Public-Not charter, magnet, or virtual	South	Town	71	36	35	12	23	71
Public-Not charter, magnet, or virtual	South	Rural	148	74	74	26	48	148
Public-Not charter, magnet, or virtual	West	City	107	54	53	18	35	107
Public-Not charter, magnet, or virtual	West	Suburb	101	51	50	17	33	101
Public-Not charter, magnet, or virtual	West	Town	60	36	24	12	12	60
Public-Not charter, magnet, or virtual	West	Rural	95	48	47	16	31	95
Catholic	Northeast		101	51	50	17	33	101
Catholic	Midwest		95	48	47	16	31	95
Catholic	South		89	45	44	15	29	89
Catholic	West		80	48	32	16	16	80
Other private	Northeast		89	45	44	15	29	89
Other private	Midwest		101	51	50	17	33	101
Other private	South		163	82	81	28	53	163
Other private	West		95	48	47	16	31	95

The desired numbers of participating schools by the margins of the school stratification characteristics are shown in table 4.

**Table 4. HS&B:21 BYFS School Participation Goals, by School Stratification Characteristics**

Total Region		Public	Catholic	Other private	Total
		<b>641</b>	<b>128</b>	<b>151</b>	<b>920</b>
	Northeast	71	34	30	135
	Midwest	93	32	34	159
	South	150	30	55	235
	West	126	32	32	190

		<b>Public</b>	<b>Catholic</b>	<b>Other private</b>	<b>Total</b>
<b>Locale</b>	City	105	NA	NA	105
	Suburb	137	NA	NA	137
	Town	72	NA	NA	72
	Rural	126	NA	NA	126
<b>Public-school Type</b>	Charter	67	NA	NA	67
	Magnet	67	NA	NA	67
	Virtual	67	NA	NA	67
	Not charter, magnet, or virtual	440	NA	NA	440

NA: Not Applicable. No explicit participation goals are established for Catholic and other private schools by locale.

The 27 school strata along with the corresponding stratum-specific participation goals, frame counts, total selected school sample (n=2,654), initial school sample (n=1,373), and reserve sample (n=1,281) are shown in table 1. The size measure used for the probability proportional to size selection of 2,654 schools was constructed using the overall sampling rates for students in the following five student categories:

- American Indian or Alaskan Native (AIAN), non-Hispanic
- Asian, non-Hispanic,
- Hispanic,
- Black, non-Hispanic, and
- Other race, non-Hispanic

combined with the total number of students in each of those five categories at a given school. In other words, the size measure for a given school (*i*) in school stratum *h* may be written as follows:

$$S_{h,i} = \sum_{j=1}^5 r_{h,j} N_{h,i,j}$$

Where  $r_{h,j}$  is the sampling rate for the  $j^{\text{th}}$  student category in the  $h^{\text{th}}$  school stratum and  $N_{h,i,j}$  is the number of students in the  $j^{\text{th}}$  category within school *i* in the  $h^{\text{th}}$  school stratum. The sampling rate,  $r_{h,j}$ , equals the number of students to sample from the  $j^{\text{th}}$  category in the  $h^{\text{th}}$  school stratum divided by the number of students in the  $j^{\text{th}}$  category across all schools in the  $h^{\text{th}}$  school stratum. The sampling rates for the five student categories listed above vary across the school strata; for example, a rate of .072 is used for AIAN students attending traditional public schools in the suburban areas in the South while an overall rate of .003 is used for students attending traditional public schools in suburban areas in the South. The student sampling rates by school strata are provided in table 5.

The sampling plan is designed to produce constant weights within each of the five student domains (AIAN non-Hispanic, Asian non-Hispanic, Hispanic, Black non-Hispanic, and other non-Hispanic) within each school stratum. When weights are constant within a given student domain and school stratum, there is no increase in the design effect due to unequal weights for estimates produced for the given student domain and school stratum.

Within participating schools, students will be stratified into the five student categories defined above and a systematic sample of students will be selected from each student sampling stratum. Approximately 28 students will be sampled, on average, from each of the anticipated 920 participating schools. However, the number of students sampled per student stratum will vary by school because the within-school student-stratum sample sizes depend upon the numbers of students in the five student sampling strata. The process of determining the student sample allocation follows the procedure outlined in section 2 of Folsom et al (1987).<sup>6</sup>

**Table 5. Aggregate Student Sampling Rates Used for School Selection**

<b>School Type</b>	<b>Census Region</b>	<b>Locale</b>	<b>Overall</b>	<b>American Indian or Alaskan Native, non-Hispanic</b>	<b>Asian, non-Hispanic</b>	<b>Hispanic</b>	<b>Black, non-Hispanic</b>	<b>Other, non-Hispanic</b>
Public-Charter			0.011	0.103	0.030	0.007	0.009	0.011
Public-Magnet			0.006	0.097	0.025	0.004	0.005	0.005

<sup>6</sup> Folsom, R.E., Potter, F.J., and Williams, S.R. (1987). Notes on a Composite Size Measure for Self-Weighting Samples in Multiple Domains. Proceedings of the Section on Survey Research Methods of the American Statistical Association, 792-796.

School Type	Census Region	Locale	Overall	American Indian or Alaskan Native, non-Hispanic	Asian, non-Hispanic	Hispanic	Black, non-Hispanic	Other, non-Hispanic
Public-Virtual			0.048	0.090	0.075	0.042	0.053	0.047
Public-Not charter, magnet, or virtual	Northeast	City	0.004	0.065	0.009	0.002	0.003	0.003
Public-Not charter, magnet, or virtual	Northeast	Suburb	0.003	0.063	0.007	0.002	0.003	0.002
Public-Not charter, magnet, or virtual	Northeast	Town	0.005	0.061	0.048	0.012	0.020	0.003
Public-Not charter, magnet, or virtual	Northeast	Rural	0.004	0.062	0.032	0.002	0.010	0.002
Public-Not charter, magnet, or virtual	Midwest	City	0.004	0.067	0.012	0.003	0.003	0.002
Public-Not charter, magnet, or virtual	Midwest	Suburb	0.003	0.065	0.010	0.003	0.003	0.002
Public-Not charter, magnet, or virtual	Midwest	Town	0.004	0.066	0.037	0.004	0.009	0.002
Public-Not charter, magnet, or virtual	Midwest	Rural	0.005	0.079	0.031	0.004	0.007	0.002
Public-Not charter, magnet, or virtual	South	City	0.003	0.067	0.011	0.002	0.002	0.002
Public-Not charter, magnet, or virtual	South	Suburb	0.003	0.072	0.007	0.002	0.002	0.002
Public-Not charter, magnet, or virtual	South	Town	0.004	0.076	0.037	0.003	0.003	0.002
Public-Not charter, magnet, or virtual	South	Rural	0.004	0.085	0.016	0.002	0.002	0.002
Public-Not charter, magnet, or virtual	West	City	0.004	0.077	0.006	0.002	0.003	0.002
Public-Not charter, magnet, or virtual	West	Suburb	0.003	0.070	0.006	0.002	0.004	0.002
Public-Not charter, magnet, or virtual	West	Town	0.008	0.082	0.032	0.003	0.021	0.003
Public-Not charter, magnet, or virtual	West	Rural	0.009	0.086	0.025	0.003	0.013	0.003
Catholic	Northeast		0.032	0.064	0.090	0.043	0.049	0.023
Catholic	Midwest		0.028	0.069	0.082	0.047	0.050	0.020
Catholic	South		0.032	0.065	0.079	0.039	0.052	0.025
Catholic	West		0.041	0.073	0.094	0.035	0.065	0.030
Other private	Northeast		0.028	0.063	0.035	0.065	0.059	0.023
Other private	Midwest		0.039	0.064	0.085	0.066	0.057	0.031
Other private	South		0.022	0.070	0.040	0.039	0.036	0.017
Other private	West		0.033	0.072	0.028	0.055	0.067	0.029

Once schools are selected and recruited, students enrolled in grade 9 will be selected from student rosters that schools or school districts will be asked to provide. Whether the school or school district provides the information will be decided by the school or school district. The student sample sizes were determined by the requirement that at least 1,708 students in each of the five student domains<sup>7</sup> participate in a second follow-up of HS&B:21. The 1,708 requirement was determined by evaluating the minimum required sample size that would be able to measure a relative change of 15 percent in proportions between the first- and second follow-up for students in the twelfth-grade as of the first follow-up. Students in twelfth grade during the first follow-up, both those who were in the sample in ninth grade and those who were added through freshening in twelfth grade, form the twelfth-grade cohort.<sup>8</sup> Setting the minimum sample size required to achieve 1,708 participating students in the twelfth-grade cohort as of a second follow-up also ensures that the number of participating students in the second follow-up who were in the ninth-grade in the base year will exceed 1,708.<sup>9</sup> Several assumptions were used to conduct this evaluation, as noted below.

- Two-tailed tests with significance of alpha = 0.05 were used to test differences between means and proportions with required power of 80 percent.
- A proportion of p = .30 was used to calculate sample sizes for tests of proportion.
- Design effect is 2.5.

<sup>7</sup> The five student domains are as follows: AIAN, non-Hispanic; Asian, non-Hispanic; Hispanic; Black, non-Hispanic; and Other race, non-Hispanic.

<sup>8</sup> The twelfth-grade cohort includes those ninth-grade students sampled in the base year who are enrolled in twelfth grade in the first follow-up. These students alone, however, are not representative of all students enrolled in twelfth grade in spring 2025. There will be students who enroll in U.S. schools after ninth grade, as well as students in twelfth grade in the spring of 2025 who were not in ninth grade in the fall of 2021. Additional students enrolled in twelfth grade will be added to the student sample in the first follow-up through sample freshening to produce a sample of students that is representative of all students enrolled in twelfth-grade in spring 2025. Because the number of students to be added via sample freshening is uncontrollable and estimation is imprecise, base-year students sample sizes were established assuming no students would be added via sample freshening.

<sup>9</sup> Some of the students sampled in the ninth grade will be in the twelfth grade as of the first follow-up but some will not. Students sampled in ninth grade will be included in the first follow-up data collection regardless of the grade they are in, but those who are not in twelfth grade will not be part of the twelfth-grade cohort.

- Correlation between waves is 0.6.

McNemar’s test using Connor’s approximation was used to determine the minimum sample size needed to meet the precision requirement under the previously stated assumptions. The Proc Power procedure available in SAS software<sup>10</sup> was used to determine the minimum sample size.

Sample sizes were also established for three other student domains: students attending virtual schools, students attending charter schools, and students attending magnet schools. The minimum sample sizes used for these three student domains were established by using all but one of the same assumptions and requirements as were used for five student race/ethnicity domains. The requirement to measure a relative change of 15 percent in proportions between a first- and second follow-up for students in the twelfth-grade as of the first follow-up was changed to require the ability to measure a 20 percent change in proportions. Under this modified requirement, a minimum of 978 participating students are required as of the end of a second follow-up for students in these three types of public-schools.

The minimum number of students to sample from each of the five student categories and three public-school domains in the BYFS to achieve 1,708 participating students in each of the five student race/ethnicity domains and to achieve 978 participating students in the three public-school domains for the twelfth-grade cohort as of a second follow-up, along with the assumptions used to derive those numbers, are provided in table 6. The assumptions in Tables 6 and 7 are supported and motivated by previous work on HLS:09. The numbers reported in table 6 show, for example, that if 3,500 Hispanic students are sampled in grade 9 then approximately 1,708 will respond in the base year, enter grade 12 in 2025, and respond in both the first- and second-follow-ups.

As can be seen in table 7, if the minimum sample sizes reported in table 6 are used in the BYFS, the estimated numbers of participating students among the BYFS ninth-grade cohort as of a second follow-up exceed the minimum numbers of required participating students of 1,708 and 978, for the five student race/ethnicity domains and three public-school type domains, respectively. The numbers reported in table 7 show, for example, that if 3,500 Hispanic students are sampled in grade 9 then approximately 1,847 will respond in the base year as well as in both the first- and second-follow-ups (regardless of what grade they enter in 2025).

**Table 6. Minimum Sample Sizes and Associated Sample Design Assumptions for Student Sampling Categories for Twelfth-grade Cohort**

Assumption	Each Student Race/Ethnicity Domain <sup>11</sup>	Charter students	Magnet students	Virtual Students
Ninth-grade inflated student sample size for each key student domain	3,500	2003	2003	2003
Ninth-grade student eligibility rate	95%	95%	95%	95%
Ninth-grade student response rate	85%	85%	85%	85%
Ninth-grade respondents	2,826	1,617	1,617	1,617
Percentage of ninth-grade students in twelfth grade as of the first follow-up	90%	90%	90%	90%
Ninth-grade respondents in twelfth grade as of the first follow-up	2,544	1,456	1,456	1,456
Percentage of twelfth-grade students whose transfer schools are known	15%	15%	15%	15%
Percentage of twelfth-grade students whose transfer schools are not known by their base-year school	8%	8%	8%	8%
Percentage of twelfth -grade students who are known to remain in base-year schools	77%	77%	77%	77%
Response rate for twelfth-grade transfer students whose schools are known	60%	60%	60%	60%
Response rate for twelfth-grade transfer students whose schools are unknown	50%	50%	50%	50%
Response rate for twelfth-grade students known to remain in their base-year schools	90%	90%	90%	90%
First follow-up respondents	2,093	1,198	1,198	1,198
Second follow-up eligibility rate	100%	100%	100%	100%
Second follow-up locate rate	96%	96%	96%	96%

<sup>10</sup> SAS Institute Inc. 2008. SAS/STAT® 9.2 User’s Guide. Cary, NC: SAS Institute Inc.

<sup>11</sup> AIAN, non-Hispanic; Asian, non-Hispanic; Hispanic; Black, non-Hispanic; and Other race, non-Hispanic.



Assumption	Each Student Race/Ethnicity Domain	Charter students	Magnet students	Virtual Students
Second follow-up response rate	85%	85%	85%	85%
Second follow-up respondents	1,708	978	978	978

**Table 7. Minimum Sample Sizes and Associated Sample Design Assumptions for Student Sampling Categories for Ninth-grade Cohort**

Assumption	Each Student Race/Ethnicity Domain <sup>8</sup>	Charter students	Magnet students	Virtual Students
Ninth-grade inflated student sample size for each key student domain	3,500	2003	2003	2003
Ninth -grade student eligibility rate	95%	95%	95%	95%
Ninth -grade student response rate	85%	85%	85%	85%
Ninth -grade respondents	2,826	1,617	1,617	1,617
Percentage of Ninth-grade students not in base-year school or whose status is indeterminate <sup>12</sup> as of the first follow-up	30%	30%	30%	30%
Twelfth-grade response rate among students known to be in base-year schools	90%	90%	90%	90%
Twelfth-grade response rate among students not in base-year schools or whose status is indeterminate	57%	57%	57%	57%
First follow-up respondents	2,264	1,296	1,296	1,296
Second follow-up eligibility rate	100%	100%	100%	100%
Second follow-up locate rate	96%	96%	96%	96%
Second follow-up response rate	85%	85%	85%	85%
Second follow-up respondents	1,847	1,057	1,057	1,057

Estimates of the minimum number of students to sample in the BYFS were derived by adjusting the 1,708 and 978 required for a twelfth-grade cohort to account for a variety of factors including estimates of student response in 2021 (grade 9), 2025 (grade 12), and at some point a few years after 2025 representing a second follow-up as well as other factors, including the extent to which BYFS participating schools agree to participate in first and second follow-up studies and the extent to which students are expected to move between schools between grades 9 and 12.

Using the minimum sample sizes reported in the top row of table 7, the minimum required sample size is 17,500 (3,500\*5) students but the sample allocation substantially oversamples AIAN, non-Hispanic students; Asian, non-Hispanic students; students attending virtual schools; and students attending charter schools. The sample allocation substantially undersamples Other, non-Hispanic students and undersamples Hispanic students. In order to reduce the impact of disproportionate sampling on national estimates and estimates that compare or combine estimates across student categories, the sample sizes for the Hispanic and Other, non-Hispanic student domains were increased. The increases in the Hispanic and Other, non-Hispanic student domains were determined by specifying and solving a non-linear optimization problem that sought to determine the total student sample size and student domain sample sizes that would produce design effects of 2.5 or less for estimates of means within each of the five key student race/ethnicity groups. The solution of this non-linear optimization problem indicated a total student sample size of 26,000 students would be required. The allocation of these 26,000 students to each of the five key student/race ethnicity groups was also determined by the solution to the non-linear optimization problem and the allocation is reported in the second row of table 8.

**Table 8. Final Student Sample Sizes and Expected Minimum Student Participation**

Assumption	AIAN, non-Hispanic	Asian, non-Hispanic	Hispanic	Black, non-Hispanic	Other, non-Hispanic	Total
Ninth-grade inflated student sample size for each key student domain	3,500	3,500	4,500	3,500	11,000	26,000
Ninth-grade student eligibility rate	95%	95%	95%	95%	95%	
Ninth-grade student response rate	85%	85%	85%	85%	85%	

<sup>12</sup> Students of indeterminate status include those whose whereabouts are not known by their base-year schools. Such students include those who may be transfer students, early graduates, or dropouts.

Assumption	AIAN, non-Hispanic	Asian, non-Hispanic	Hispanic	Black, non-Hispanic	Other, non-Hispanic	Total
Ninth-grade respondents	2,826	2,826	3,634	2,826	8,883	20,995
Percentage of Ninth-grade students not in base-year school or whose status is indeterminate <sup>9</sup> as of the first follow-up	30%	30%	30%	30%	30%	
Twelfth-grade response rate among students known to be in base-year schools	90%	90%	90%	90%	90%	
Twelfth-grade response rate among students not in base-year schools or whose status is indeterminate	57%	57%	57%	57%	57%	
First follow-up respondents	2,264	2,264	2,911	2,264	7,115	16,818
Second follow-up eligibility rate	100%	100%	100%	100%	100%	
Second follow-up locate rate	96%	96%	96%	96%	96%	
Second follow-up response rate	85%	85%	85%	85%	85%	
Second follow-up respondents	1,847	1,847	2,375	1,847	5,806	13,722

Therefore, for the BYFS the plan is to sample 28 students, on average, within each of 920 participating schools for a total of 26,000 sample students and, assuming the grade 9 eligibility and response rates shown in table 8 (an estimated 5 percent of sampled students are expected to be ineligible), to produce approximately 20,995 participating grade 9 students. Thus, to achieve a yield of 20,995 ninth-grade students, the parents of approximately 24,700 ninth-grade students will need to be contacted for consent ( $26,000 \times 0.95 = 24,700$ ). The distribution of the grade 9 student sample and estimates of the number of participating students in the BYFS, first follow-up, and second follow-up are provided in table 8.

**Roster Collection** BYFS rosters will be requested for ninth-grade students beginning in the fall of 2021 (appendix A.10).

The rosters may be provided from the district or from the school, and it will be requested that the roster be provided once the enrollment for the school year has stabilized (which is often approximately 4 weeks into the school year) to increase accuracy. Key information needed for student sampling will be requested, such as: student name; school or district student ID number; date of birth; grade level; gender; race/ethnicity; IEP/504 status; and ELL status. Each of these characteristics is important for sampling purposes, but we will work with schools that are unable to provide all of the information to obtain the key information available. Based on this information, the student sample will be drawn. As part of the roster collection, the study will also request from the school coordinator or designated district personnel the following information for each student eligible for sampling: student's parent and/or guardian contact information (e.g., mailing address; landline phone number; cell phone number; e-mail address) and student's math teacher. Schools and districts often find it easier, and therefore more efficient, to supply all of the desired information one time for all of their students. However, should it be problematic for any school or district to provide the parent and teacher information on the complete roster, the recruitment team will gather that information as a second step for the sampled students only. If the school and/or district is unwilling to provide parent contact information for the sampled students, the team will work with the school and/or district to determine the best way to contact parents (e.g., the school coordinator or designated district personnel would facilitate contacting parents and/or would mail the required materials to parents using the contact information they have on file). Parent contact information is required to conduct the out-of-school student data collection.

The roster request will include a template and secure transfer options to deliver the rosters. The data quality of the student rosters will then be evaluated by:

- reviewing and assessing the quality and robustness of student and parent information available at each school, including contact information for parents;
- reviewing and assessing the quality of the data on student-teacher linkages;
- addressing any incompleteness or irregularities in the roster file;
- requesting additional information as needed from the school coordinator or designated district personnel; and
- (re)verifying that the sampled students are currently in attendance in the school.

The provider of the roster, whether from a school district or school, will receive the \$50 incentive for providing the roster as described in the Supporting Statement Part A, section A.9.

## B.3 Methods to Secure Cooperation, Maximize Response Rates, and Deal with Nonresponse

### School Recruitment

Gaining cooperation from school districts and schools is paramount to the success of this voluntary study. However, recruitment efforts in similar studies have been meeting with increasing challenges that must be carefully mitigated to ensure adequate school participation. For example, in 1998–99 the Early Childhood Longitudinal Study had a weighted school-level response rate of 74 percent,<sup>13</sup> whereas 12 years later, the successor ECLS-K:2011 cohort study had a weighted school-level response rate of 63 percent.<sup>14</sup> Additionally, response rates tend to be lower for schools that serve older students (e.g., the High School Longitudinal Study of 2009 (HLS: 09) had a weighted school-level response rate of 56 percent,<sup>15</sup> and the Middle Grades Longitudinal Study of 2017-18 (MGLS:2017) had an unweighted school-level response rate of 39 percent). Methods to secure cooperation of school districts and schools, maximize response rates, and deal with nonresponse are described in this section.

**Maximizing School Participation.** The success of HS&B:21 hinges on securing the cooperation and maximizing response rates among school districts and schools, and then their students, parents, and staff. Participation among school districts and schools has been declining for voluntary school-based studies. Often, district and school personnel understand the value of the research but that does not offset their reasons for not participating. Reasons cited for not participating in voluntary school-based studies such as MGLS:2017 and HLS: 09 are the high burden associated with participating, lack of direct benefit, over-testing of students, loss of instructional time, lack of parent and teacher support, increased demands on school staff, and a moratorium on outside research, especially in light of so much uncertainty in the 2020-2021 school due to the ongoing pandemic. To mitigate these concerns, HS&B:21 has developed a recruitment plan to maximize school participation that is comprehensive and flexible in its approach and incentive structure to effectively secure cooperation for the study. Strategies recommended to maximize school participation include:

**Outreach.** Study and NCES name recognition add validity to the study when recruiting school districts and schools. Even prior to drawing the sample, outreach activities will be conducted to announce the upcoming study and begin to garner support from states, districts, schools, and stakeholders. Outreach activities will include:

- Contacting school districts that are typically challenging to recruit to discuss their decision-making process about participating in research studies and the benefits they would like to see from participating. This would be done generically as an exercise to help plan the recruitment effort prior to selecting the sample.
- Distribution of a brief video to stakeholders, state, school districts, and schools to explain the importance of the study (Appendix A.6e; video can be viewed at <https://www.youtube.com/watch?v=EolkmqoHoWk>).
- Attendance at conferences attended by stakeholders to introduce and promote the study.

**Compelling recruitment materials.** School districts and school staff are busy. Materials sent to these contacts must be informative, compelling, and brief. In addition, multiple types of materials (e.g., mailings, video, website) will be made available to ensure that decision-makers have options to receive the message in the manner that works best for them. The materials will all be available on the study website for easy access. Reviewing these study materials should provide districts and school administrators with an understanding of the study's value, the importance of HS&B:21, and the data collection activities required as part of the study. A full understanding of these factors will be important both to obtain cooperation and to ensure that schools and districts accept the data collection requests that follow.

**Determining the “right” amount of time to facilitate participation.** One of schools' primary concerns about participating in outside research is the loss of instructional time. In an effort to balance the need to collect information with the desire to minimize the burden on schools and students, we tested two student session lengths in the field test. Described in section B.4, for the field test, half of the schools were assigned at the outset to a 45-minute session and half to a 90-minute session. If a school in the 90-minutes group was unable to accommodate

<sup>13</sup> Tourangeau, K., Nord, C., Lê, T., Sorongon, A.G., Hagedorn, M.C., Daly, P., and Najarian, M. (2001). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), User's Manual for the ECLS-K Base Year Public-Use Data Files and Electronic Codebook* (NCES 2001-029). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

<sup>14</sup> Tourangeau, K., Nord, C., Lê, T., Sorongon, A.G., Hagedorn, M.C., Daly, P., and Najarian, M. (2012). *Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011), User's Manual for the ECLS-K:2011 Kindergarten Data File and Electronic Codebook* (NCES 2013-061). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

<sup>15</sup> Ingels, S.J., Pratt, D.J., Herget, D.R., Burns, L.J., Dever, J.A., Ottem, R., Rogers, J.E., Jin, Y., and Leinwand, S. (2011). *High School Longitudinal Study of 2009 (HLS:09). Base-Year Data File Documentation* (NCES 2011-328). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

the full 90-minutes, it was offered to partake in a one-class period or 45-minute session with the remainder of the session to be completed by the student outside of school. Schools unable to provide a 45-minute session were asked, as a refusal conversion option, to provide roster information so that students could be contacted to participate outside of school. For BYFS, the 90-minute session will be offered to schools during recruitment and the 45-minute session or out-of-school option will be offered only as part of a non-response and/or refusal conversion strategy.

For teachers and parents, in the field test, we experimented with two survey lengths to determine which was most effective in securing participation. Teachers were offered either the full survey or an abbreviated survey. The mathematics teacher survey consisted of two sections: the teacher-level portion of the survey, which included classroom-level information and took either 16 or 10 minutes, and the student-level report, which took either 4 or 3 minutes per student to complete. For BYFS, teachers will be offered the full-length survey. Because of the small school sample size, no experimentation was explored with guidance counselors or administrators beyond the student session length.

**Varied communication modes.** Prior to the start of recruitment, study team members will review the sample of school districts and schools to determine the appropriate mode of communication for each. Staff who may have connections to a particular area or district may be called upon to make the first contact. That staff person may remain the primary contact to the school district or school, or they may turn the school over to a recruiter for the collection of logistical information. Communication may be conducted via mail, email, phone, and in-person methods.

**Leveraging participation in past studies.** NCES has been conducting the longitudinal studies in schools since the early 1970s. HS&B:21 will leverage the participation information of school districts and schools from prior studies. Information gathered will include whether or not the school district or school participated, whether a research application was required, who in the district or school made the decision about participating, and the reasons for refusal, if applicable. This information will be used only to strategize who to contact and how to respond to previous concerns, if any, to encourage participation in HS&B:21.

**Geospatial Modeling.** The use of geospatial modeling helps overcome a serious challenge in base-year studies – a lack of information on sample members to use in predicting response. We will use geospatial modeling with data from open-access sources (e.g., NCES’s Education Demographic and Geographic Estimates, FBI’s Uniform Crime Reporting statistics) and RTI’s Enhanced Address-based Sampling Frame (Enhanced ABS Frame) to generate a superset of covariates that may help in estimating the likelihood of participation at all levels. From this covariate superset, we will identify a subset of *substantive* covariates that are significant predictors of response. Model coefficients from this geospatial model will then be used to predict, a priori, the likelihood of response for each unit in the HS&B:21 sample if the model provides sufficient ability to predict nonresponse. Interventions will be implemented based on the patterns of information found in the model. If we see that certain areas are underrepresented, we will focus our outreach efforts in those areas.

**Volunteer hours for students.** Many high schools require volunteer hours to be completed prior to graduation. As a token of appreciation for student participation, the U.S. Department of Education will provide a certificate to each participating student to acknowledge 2 hours of volunteer service for participation in the study. For confidentiality purposes, no study-specific information will be included on the certificate.

**Flexible incentive package.** As observed across NCES studies, schools and sample members vary in their motivation to participate in voluntary research, and what incentivizes some does not work for others. We are thus offering incentive choices to enable schools and sample members to determine what works best for them. The incentive structure is provided in section A.9 of the Supporting Statement Part A. To encourage schools to participate, we are also offering school-related incentives to students, such as entry to a school event or credit towards, or an item from, the school store and parents will have the opportunity to donate their incentive to the school. Such incentives further support the schools and can be seen as a benefit to both the school and students. Furthermore, increased flexibility in the items that are offered may encourage participation. In BYFS, a school-level incentive boost from \$200 to \$400 will be offered to schools as a refusal conversion strategy to mitigate nonresponse both at the school-level and at the student item-level. This school-level incentive boost was not offered in the field test due to sample size and its potential impact on the field test session time experiment. A shorter student session in the school may also be offered as the next refusal conversion step.

**Webinars.** School district decisionmakers report that they require studies to provide timely results to participate

in research. Since we are unable to provide results to schools until after the international data are released, we plan to provide topical webinars to school staff as a way to “give back” soon after data collection while they wait to receive results. These low-cost webinars will be delivered on topics such as science, technology, engineering, and mathematics (STEM), socio-emotional learning (SEL), and project-based learning, to be delivered through RTI’s experienced technical assistance providers. The webinar will be live, but a recording will be made available on the study website.

**School Reports.** Schools participating in the national study will receive a report following each of the two data collection points containing aggregate results. Reports will provide comparative data on the school’s reading and math scores related to other “schools like yours” nationally, wherein data are presented for schools that are similar on key characteristics, such as urbanicity, sector (i.e., public/private), grade-levels served, and geographic region (no specific schools are named). Average raw or scale scores will be provided on topics such as mathematics performance, reading performance, growth mindset, belongingness, race relations, student-counselor interaction, educational attainment, and financial literacy.

Schools will need to have a minimum participation threshold to receive results for their school (50% participation and at least 18 participating students) but all schools will receive data on schools like theirs. While similar data may already be collected by districts, HS&B:21 will afford the opportunity to build upon local efforts and compare results with state and national findings. Links to resources on each topic area will also be provided so schools may gather additional information, if desired. Reports will be provided for each year the school participates, which will enable schools to see change over the course of the longitudinal study.

**Certificates of Service for teachers, counselors, and administrators.** Certificates of service will be given to all staff participating in the study. In most states, districts, and schools, the school staff are required to participate in documented professional development activities. If state or district requirements allow, these certificates may be used to apply for professional development credits.

**Avoiding refusals.** HS&B:21 recruiters will be trained to avoid direct refusals by focusing on strategies to solve problems or meet obstacles to participation faced by district or school administrators. They will endeavor to keep the door open while providing additional information and seeking other ways to persuade school districts and schools to participate. As described above, shortening the time for the student session is one tool that will be used for refusal avoidance.

When possible, HS&B:21 session facilitators will meet with students prior to the session at the school and explain the importance of participating in the study. They will emphasize that participating in the study will not affect their grades, and that none of their responses will be shared with their parents or teachers. Session facilitators may also meet with parents at a scheduled parent event at the school to generate excitement about the study, and answer questions parents may have about either their or their child’s participation. Parents that have not returned permission forms or who have initially refused to provide permission for their child to participate will be contacted by session facilitators who will attempt to alleviate any concerns about their child’s participation and answer any questions parents may have about the study. HS&B:21 session facilitators will also prompt school staff to complete the staff surveys and assist school staff with logging in to their survey or answering questions about the study.

**General Recruiting.** The following approach will be implemented to recruit school districts and schools for both the field test and BYFS. This approach was previously approved in December 2018 (OMB# 1850-0944 v.1.).

**Organizational Endorsements.** Support from leading education organizations can, at times, be influential to school districts’ and schools’ decision to participate. Prior to contacting sampled school districts or schools, we will request the endorsement and support from relevant organizations and key stakeholders in secondary education (see appendix A.1). Organizations will be able to provide a letter of support and/or electronic endorsement of the study. Endorsing organizations will be listed on the HS&B:21 study website. Endorsement collection will continue through summer 2021.

**State Endorsements.** As part of our study outreach efforts, all states will be contacted to inform them that the study will be taking place. States will be asked to provide a letter of endorsement to encourage school districts’ and schools’ participation in the study should schools in their state be selected. Letters will be sent to the state superintendent with copies to the state-level director of research and/or director of secondary education, as applicable (see appendix A.2). Once the sample schools are selected, senior HS&B:21 recruitment staff will contact state staff with schools in the sample to discuss the study and secure support. Endorsement letters received by the

state are included in all mailings to districts and schools within the state. Endorsement collection will continue through summer 2021.

**School District and Diocesan Notification and Recruitment.** Once states have been contacted, whether an endorsement letter was received or not, school districts and dioceses that do not require a research application will be notified that schools in their district have been selected for the study. The letter to school districts will state that NCES's contractor, RTI International, will contact the selected school(s) within two weeks and districts may also contact RTI with questions (see appendix A.3). A customized mailing envelope (see appendix A.15b) will be used to send initial study materials. Along with the letter, districts will receive a separate sheet of paper devoid of study name or logo listing names of district schools selected for participation, a study brochure (see appendix A.6c), and FAQs (see appendix A.6b).

Research applications will be prepared for any school districts that require the approval of applications in order to conduct research in schools in their jurisdiction. If a school district notifies us that an application must be submitted, or some other requirement must be fulfilled, study staff will be prepared to respond to such requirements. If a district chooses not to participate, all reasons will be documented to help formulate a strategy for refusal conversion attempts. Participating districts may be asked to provide student roster information on the school's behalf to reduce the burden on the school.

**Public and Catholic School Recruitment.** Two weeks after the district was notified, or after the district provides approval when required, recruitment will commence at the school-level. Schools will receive a colored folder that contains a letter (appendix A.4a), frequently asked questions (FAQs) about the study (appendix A.6b), and a study brochure (appendix A.6c). Materials will be sent via overnight delivery and follow-up will occur within three business days. The first contact will be intentionally assigned based on prior history of working with the schools and school or district characteristics. First contacts may include modes such as a telephone call from recruitment staff, study management staff at RTI, or NCES staff; or an in-person visit to the school. Each of these modes, as well as email communication, may be used throughout the recruitment process as needed.

Once a school agrees to participate in HS&B:21, a recruiter will work with the school to name a member of the school's staff to serve as the school coordinator for the study. The recruiter will work with the school coordinator to schedule study activities at the school, including gathering grade 9 student roster, distributing consent materials to parents of sample students, and arranging the session logistics. Roster instructions will be sent electronically in the fall of 2019 for the field test and the fall of 2021 for BYFS (see Appendix A10) and are also available on the website (see Appendix A7a). If a school is experiencing difficulty with preparing the roster, the district may be asked to provide the roster on the school's behalf.

In early communications, the recruiter will also gather information about the school including: what type of parental consent procedures need to be followed at the school; hours of operation, including early dismissal days, school closures/vacations, and dates for standardized testing; and any other considerations that may impact the scheduling of student sessions (e.g., planned construction periods, school reconfiguration, or planned changes in leadership). The HS&B:21 study recruitment team will meet regularly to discuss recruitment issues and develop strategies for refusal conversion on a school-by-school basis.

As mentioned, for the field test, half of the schools were presented with a 90-minute student session. The other half were presented with a 45-minute (one-class-period session) student session with the remainder of the session to be completed outside of school. As a refusal conversion effort, schools assigned to the 90-minute session were asked to consider a 45-minute instead. If the school still declined, an out-of-school student session was offered. Similarly, as a refusal conversion effort, if a school that was offered the 45-minute student session declined, that school was also offered an out-of-school student session in which the entire battery was completed outside of school. For BYFS, the 90-minute session will be offered to schools during recruitment and the 45-minute session or out-of-school option will be offered only as part of a non-response and/or refusal conversion strategy.

**Private and Charter School Recruitment.** If a private or charter school selected for the base-year field test operated under a higher-level governing body such as a diocese, a consortium of private schools, or a charter school district, we used the district-level recruitment approach with the appropriate higher-level governing body. If a private or charter school selected for the field test did not have a higher-level governing body, the school recruitment approach outlined above was used.

**Out-of-School Data Collection.** Some schools may not permit the data collection to occur in a school-based session. To maximize participation and address school concerns about loss of instructional time, schools declining

to conduct an in-school group session will be offered the possibility to have their students participate outside of school. For these schools, schools will still be asked to provide the student roster, teacher information, and parent contact information. We will ask out-of-school schools to distribute materials to sampled students and parents and help to encourage participation. This could mean sending materials to parents via mail, email, or distribution of materials through students; providing computer access so that students can participate at their convenience in the school; and/or following up with students and parents to encourage participation. The person designated as the school coordinator for the school would receive the coordinator incentive as if the session was happening in school. Parent contact information would be used to contact parents directly to secure student and parent participation. Contacts outside of school would be in addition to prompting by the school coordinator, for schools willing to assist with this activity. Teachers, administrators, and counselors would be asked to participate as if the session were conducted in schools.

**Parent Recruitment.** For schools allowing in-school student sessions, schools will be given the option of one of three types of parental permission letters: notification (appendix A.5a-b), implicit permission (opt out) (appendix A.5c-d), or explicit permission (opt in) (appendix A.5e-f). Each type of consent requires that parents be notified that their children have been selected for the study. With a notification letter, no permission form is sent home since no action is required on the part of the parent. For implicit consent (opt out), the school does not require verbal or written consent for a student to participate in the study – parents are asked only to notify the appropriate person if they do not want their child to participate. With explicit consent (opt in), children may participate only if their parents provide written or oral consent for their children to do so. Proactive parent recruitment will be focused on maximizing the number of parents (1) returning signed explicit consent forms and (2) completing the parent survey. Because implicit consent does not require a verbal or written response from parents, these parents will not be contacted about consent forms. The letter accompanying the parent permission form will let parents know that the students will complete a survey, the math and reading questions, and a hearing and vision assessment. Parents will be told that they will receive results from the hearing and vision assessments.

The letters will be sent to the school for distribution to sampled students. For the field test, students in explicit permission schools will be offered a pizza party or equivalent food event for those who return the form by a designated date, regardless of whether permission is granted. For BYFS, the option of a \$3 voucher to the school cafeteria or a food event will be offered to students in explicit permission schools for returning a signed permission form by a designated date. Also, the permission letter for parents of students in explicit permission schools will include instructions for how to provide permission electronically. The school coordinator will be able to see electronic permission status via their HS&B:21 website, and the session facilitator will walk the school coordinator through that process if applicable.

For schools that only permit out-of-school data collection, all initial contacts with the student will be conducted through the parent. The parent will receive a letter and an enclosed envelope with study information and study instructions. The parent will be asked to give the student the enclosed envelope with study information and study instructions. By giving the envelope to the student, it is implied that the parent consents to the child's participation. The study information and instructions will include a student letter and the URL and login information for the student session. Students participating outside of school will not complete the hearing or vision assessment.

Parents will also receive an invitation to participate in the parent questionnaire (appendix A.11) at the start of data collection, with parent cases being added to the data collection process on a flow basis as parent contact information is provided by the school or parents provide such information on consent forms. Parent data collection will entail web-based self-administration with nonresponse follow-up by computer-assisted telephone interviewing (CATI). Letters inviting parents to participate in the study will contain a message on the envelope such as "Important, please open!", "Your input is requested", or "Help improve education! Open to find out how."

### ***Data Collection Approach***

The HS&B:21 data collection will consist of a student session (survey, math and reading assessment, and hearing and vision assessment) as well as surveys for students' parents, math teachers, guidance counselors, and school administrators.

Once schools agree to participate, the designated school coordinator will be asked to provide a roster of all students in grade 9. The school coordinator will receive a template of the roster (see appendix B) with instructions to prepare and upload the roster electronically to the secure study website (see appendix A.10). Upon receipt of the roster, RTI statisticians will randomly select about 35 students each from grade 9. About a month prior to the

scheduled student session, a student tracking form listing the selected students will be sent to the school along with parent permission forms to distribute to the students.

Students will be asked to complete a 90-minute session in a group administration in their school. The student surveys and direct assessments will take place in the school setting and be administered using Chromebooks (tablet-like computers with touchscreen capability and an attached keyboard) brought into the school by HS&B:21 staff. The student survey will be offered in both English and Spanish. HS&B:21 staff will also bring the necessary equipment for the hearing and vision assessments. This portion of data collection is referred to as the student session. To administer the survey and direct assessment in schools, study staff will work with schools to identify and utilize locations for administration that minimize distractions for the student and disruption to the school routine. Students will be prompted on screen to report to the hearing and vision station to complete that portion of the session. This component may be turned off if schools decline to have the students' vision and hearing tested.

Schools that decline to participate in the 90-minute session may be asked to complete a 45-minute session in-school. Students from schools completing the 45-minute student session will be asked to complete the remainder of the session outside of school.

Schools that do not have a brick and mortar school or who decline to participate in the in-school session will be asked to allow students to participate outside of school. For these students, schools may be willing to distribute materials to the students to take home, send emails to parents, or mail materials directly to schools. Schools may also be willing to allow the student(s) to complete the session on a school computer at their convenience. In addition, the study team will contact the family directly to obtain student and parent participation in the study. Students from these schools will be contacted through their parent using contact information provided by the school (appendix A.8a). The parent will be asked to give an envelope to the sampled student which contains a letter including login information for the student to complete the session online (appendix A.8b). Students who participate outside of school will not be asked to complete the hearing and vision assessments.

The parent (appendix A.11) survey will have an internet option and a telephone option, while the mathematics teacher (appendix A.12a), school counselor (appendix A.12b), and school administrator (appendix A.12c) surveys will be self-administered via the Web. The parent survey will be offered in both English and Spanish.

An abbreviated version of all survey instruments may be offered during the last few weeks of data collection to mitigate nonresponse.

## **B.4 Tests of Methods and Procedures**

BYFS will collect base-year data from a sample of ninth-grade students in the fall of 2021, their teachers, their parents, their guidance counselors, and their school administrators. For the field test, we implemented a set of experiments to refine the full-scale data collection procedures for the school/students, teachers, and parents. The experiments manipulated the amount of information requested and in what setting it is collected, as well as variations on how to incentivize participation.

### **Field Test Experiments and Results**

**School/Students.** Student data collection largely relies on the schools' willingness to allow in-school data collection. A student survey that averages 90 minutes may require two class periods to administer. To examine the impact of reducing participation burden on the schools, an alternative design will fit the data collection within one class period (approximately 45 minutes), potentially gaining participation from a larger proportion of sample schools.

To better understand the effect of the length of the student session on school recruitment, the 309 schools in the field test sample were randomly assigned (after controlling for the same treatment within school district) to a 90-minute or a 45-minute in-school student administration request.

From the 159 schools assigned to the 45-minute in-school request, 21% agreed to participate, and from the 150 assigned to the 90-minute in-school request, 27% agreed to do so. However, when looking at actual participation (i.e., yielding data from students), the participation rates were only 3 percentage points apart (20% and 23%, respectively). Given the substantial amount of data that is omitted from the 45-minute administration (student survey, reading assessment, and any math items that could not be completed), we decided against the abbreviated in-school administration for BYFS, though leaving the option available as a refusal conversion strategy.



**Teachers.** Teachers have multiple job responsibilities and demands on their time. Reducing the time demand to provide survey data has the potential to be a highly effective design feature to increase participation. Teachers within each participating school were randomly assigned to provide the full teacher background information and the full teacher-student report, or to provide a reduced version of each set of data.

Only 93 teachers were offered the abbreviated teacher background survey and student reports, and 259 were offered the full instruments. The participation rates were 25% and 35%, respectively—a counter-intuitive finding that we attribute to sampling variance. Absent of any evidence against the use of the full instruments, we decided to continue with them for the BYFS. We will, however, offer an abbreviated instrument toward the end of data collection to mitigate nonresponse.

**Parents.** Although the parent survey will not be administered in an institutional setting, the length of the survey may still be an important factor in gaining participation. In addition, there are multiple ways to incentivize parent participation. We plan to offer a total of \$20 to each sample parent. Part of this \$20 will be offering a \$5 prepaid incentive. A prepaid incentive can be an effective way to encourage participation, with the remainder provided upon completion of the survey. However, we examined the most effective way to provide the remaining \$15 of the total \$20 incentive, in monetary or non-monetary options. We also tested the use of an abbreviated instrument.

We used a 2x2 factorial design, with a treatment on survey length and a treatment on incentive format. Sample parents were randomly assigned to a 30-minute survey or to a 15-minute survey. All parents received the \$5 prepaid incentive. Then, the survey time treatment (30 minutes versus 15-minutes) was crossed with assignment to be offered \$15 in cash (or check) upon completion, or to choose a different incentive with a value of approximately \$15, such as: school tickets to ball game, donation to the school for their participation, college preparation materials, movie tickets for family (2), a family board game, or donation to a charity.

The participation rates for the full and abbreviated surveys were 31.1% and 34.6%, respectively. Given the large amount of information that is sacrificed in the abbreviated survey, we decided against the abbreviated instrument for BYFS. However, an abbreviated instrument will be offered in the last few weeks of data collection to mitigate nonresponse.

The participation rates to the \$15 incentive and nonmonetary incentives were 34.4% and 31.2%, respectively. In addition to the lower participation rate, the nonmonetary incentives imposed substantial burden on the schools and could not always be provided to the respondents. For example, tickets to a school event required interaction with the school and requests were not always answered. Some schools did not even have ticketed events. Given the lower participation rate and substantial operational challenges and school burden, we decided against nonmonetary incentives for BYFS.

## **B.5 Reviewing Statisticians and Individuals Responsible for Study Design and Conduct**

The following individuals at the National Center for Education Statistics (NCES) are responsible for HS&B:21: Elise Christopher, Gail Mulligan, Chris Chapman, and Marilyn Seastrom. The following individuals at RTI are responsible for the study: Dan Pratt, Debbie Herget, Colleen Spagnardi, David Wilson, and Laura Fritch.

### **References**

Folsom, R.E., Potter, F.J., Williams, S.R. (1987) Notes on a Composite Size Measure for Self-Weighting Samples in Multiple Domains, Research Triangle Institute.  
[http://ww2.amstat.org/sections/srms/Proceedings/papers/1987\\_141.pdf](http://ww2.amstat.org/sections/srms/Proceedings/papers/1987_141.pdf)