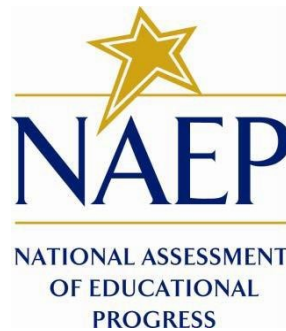*NATIONAL CENTER FOR EDUCATION STATISTICS*
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

National Assessment of Educational Progress (NAEP) 2024

Appendix C
2024 Draft Sampling Memo

OMB# 1850-0928 v.30

June 2023

.

Date:	February 28, 2023	**Memo:**	2024-1.1A/1.1B/1.1D/1.1E/1.1I

To:

William Ward, NCES	Greg Binzer
Ed Kulick, ETS	Marcie Hickman
David Freund, ETS	Kavemuii Murangi
Amy Dresher, ETS	Dwight Brock
Yue Jia, ETS	Yiting Long
Pat Stearns, Pearson	Sabrina Zhang
Lauren Byrne	Amy Lin
Lisa Rodriguez	Natalia Weil
Rick Rogers	Véronique Lieber
Rob Dymowski	Sarah Shore
William Wall	

From: Lloyd Hicks

Reviewers: Leslie Wallace, Tom Krenzke, Keith Rust

Subject: Sample Design for 2024 NAEP – DRAFT

I.   Introduction

For 2024, the NAEP assessment involves the following components:

A.   National assessments in reading and mathematics at grades 4, 8, and 12 and in science at grade 8;

B.   State-by-state and Trial Urban District Assessment (TUDA) assessments in reading and mathematics for public schools at grades 4 and 8;

C.   An assessment of mathematics in Puerto Rico for public schools at grades 4 and 8;

D.   Pilot assessments for mathematics and reading frameworks and for reading router at grades 4 and 8;

E.   Pilot assessments for Puerto Rico mathematics at grades 4 and 8;

F.   The National Indian Education Study (NIES) at grades 4 and 8; and

G.   The High School Transcript Study (HSTS) at grade 12.

H.   School-based Equipment Proof of Concept (SBE POC) in
       mathematics and reading at grades 4 and 8 (discussed only at the
       end of section III; decision from NCES pending)

Below is a summary list of the features of the 2024 sample design.

1.   The alpha[1] samples for grades 4 and 8 public schools and the delta
       samples for grades 4 and 8 private schools will be used for the
       operational assessments in reading and mathematics. The alpha
       samples will also include the Puerto Rico mathematics operational
       assessments.

2.   The beta public school samples and the epsilon private school samples
       will be used for the national science assessments at grade 8 and for the
       national mathematics and reading assessments at grade 12.

As in past NAEP studies, each Trial Urban District Assessment (TUDA) sample will form part of the
corresponding state sample, and each state sample will form part of the national sample. There are twenty-
seven Trial Urban District Assessment (TUDA) participants.

Twenty-six of the twenty-seven participated in 2022. A new TUDA participant this year is Orange County,
FL. Fresno, CA dropped out of the TUDA program in 2022. The population of schools for a TUDA district
consists of those public schools, charter and non-charter, for which the district is responsible for academic
accountability.

3.   All operational and pilot assessments will be administered using tablets.
       Each operational assessment will be conducted using a combination of
       Surface Pro tablets and Chromebooks as part of a bridge study. Each pilot
       test assessment will be conducted on only Chromebooks.

4.   The school and student sample sizes for the alpha samples in each
       state will be similar in size to 2022, which is considerably smaller than
       earlier NAEP assessments.

5.   There will be no samples in U.S. territories other than for Puerto Rico at
       grades 4 and 8.

6.   The Department of Defense (DoDEA) schools are expected to be
       reported as a single jurisdiction.

7.   The National Indian Education Study will take place again in 2024. All
       Bureau of Indian Education (BIE) schools and students will be included in
       the operational samples at grades 4 and 8. Having all BIE students in
       sample is designed to provide detailed national results for American
       Indian and Alaskan Native (AIAN) students in reading and mathematics,
       as part of the National Indian Education Study (NIES).

8.   Also, as part of NIES, in nine states (Arizona, Minnesota, Nebraska, North
       Carolina, Oregon, Utah, Washington, Wisconsin, and Wyoming) the
       public-school sample at grades 4 and 8 will be increased somewhat, in an
       attempt to give publishable results for AIAN students, for reading and
       mathematics. This will affect school sampling only. There will be no
       special student sampling procedures for this purpose. This will be
       achieved by increasing, by an appropriate factor, the measures of size of

schools with a relatively high proportion of AIAN students.

---

[1] The terminology of alpha, beta, delta, epsilon, and pi is defined in Section III.

9.   The sampling rates of private schools at grades 4 and 8 will be similar to those of 2022. Response rates permitting, this will allow separate reporting for reading and mathematics for Catholic and non-Catholic schools at grades 4 and 8, but no further breakdowns by private school type.

10.   The sample sizes of assessed students for these various components are shown in Table 1 (which also shows the approximate numbers of participating schools).

11.   In the beta public samples, there will be moderate oversampling of schools with relatively moderate-to-high proportions of Black, Hispanic, and AIAN students.

12.   There will be no oversampling of students within schools for any of the samples.

Table 1.   Target sample sizes of assessed students, and expected number of participating schools, for 2024 NAEP

	Spiral[1]	Jurisdictions		Students		
	Spiral Indic.	States[2]	Urban districts	Public school students	Private school students	Total
Grade 4						
National/state reading	DS	52	27	114,500	2,350	116,850
National/state math	DS	52	27	114,500	2,350	116,850
Puerto Rico	DP	1		4,000		4,000
Total – alpha	2			233,000		233,000
Total – delta	1				4,700	4,700
Typical max. no. students/school				50	50	
Average assessed students/school				38	20	
Total schools – alpha, delta				6,072	235	6,307
Reading Pilot (incl reading router)	NP			11,500		11,500
Mathematics Pilot	NP			10,000		10,000
Puerto Rico Pilot	MP			500		500
Total – pi	2			22,000		22,000
Typical max. no. students/school				50		
Average assessed students/school				38		
Total schools – pi				576		576
Total number of students grade 4				255,000	4,700	259,700
Total number of schools grade 4				6,648	235	6,883

Table 1.   Target sample sizes of assessed students, and expected number of participating schools, for 2024 NAEP (Continued)

	Spiral[1]	Jurisdictions		Students		
	Spiral Indic.	States[2]	Urban districts	Public school students	Private school students	Total
Grade 8						
National/state reading	DS	52	27	114,500	2,350	116,850
National/state math	DS	52	27	114,500	2,350	116,850
Puerto Rico	DP	1		4,000		4,000
Total – alpha	2			233,000		233,000
Total – delta	1				4,700	4,700
Typical max. no. students/school				50	50	
Average assessed students/school				39	22	
Total schools – alpha, delta				5,925	214	6,139
National science	DA			19,800	2,200	22,000
Total - beta	1			19,800		19,800
Total - epsilon	1				2,200	2,200
Typical max. no. students/school				50	50	
Average assessed students/school				39	22	
Total schools - beta, epsilon				495	100	595
Reading pilot (incl reading router)	NP			8,500		8,500
Mathematics pilot	NP			10,000		10,000
Puerto Rico pilot	MP			500		500
Total – pi	2			19,000		19,000
Typical max. no. students/school				50		
Average assessed students/school				39		
Total schools – pi				488		488
Total number of students grade 8				271,800	6,900	278,700
Total number of schools grade 8				6,908	314	7,222

Table 1.   Target sample sizes of assessed students, and expected number of participating schools, for 2024 NAEP (Continued)

| | Spiral[1] | Jurisdictions | | Students | | |
	Spiral Indic.	States[2]	Urban districts	Public school students	Private school students	Total
Grade 12						
National reading	DA			25,200	2,800	28,000
National math	DA			19,800	2,200	22,000
Total – beta	1			45,000		45,000
Total – epsilon	1				5,000	5,000
Typical max. no. students/school				50	50	
Average assessed students/school				35	25	
Total schools – beta, epsilon				1,286	200	1,486
Total number of students grade 12				45,000	5,000	50,000
Total number of schools grade 12				1,286	200	1,486
GRAND TOTAL STUDENTS				571,800	16,600	588,400
GRAND TOTAL SCHOOLS				14,842	749	15,591

[1] See Table 2 for definitions of DS, DA, DP, NP and MP.
[2] Includes BIE, District of Columbia (DC), DoDEA, and Puerto Rico schools.

II.   Assessment Types

The assessment spiral types are shown in Table 2. Four different spirals will be used at grade 4, five different spirals will be used at grade 8, and one spiral will be used at grade 12. Session IDs contain six characters, traditionally. The first two characters identify the assessment "type" (subjects and type of spiral in a general way). Grade is contained in the second pair of characters, and the session sequential number (within schools) in the last two characters. For example, session DS0401 denotes the first grade 4 reading and mathematics operational DBA assessment in a given school.

Table 2.   NAEP 2024 assessment types and IDs

ID	Type	Subjects	Grades	Schools	Comments
DS	Operational/ Bridge	Reading, mathematics	4, 8	Public, Private	All schools in the alpha (except Puerto Rico) and delta samples
DA	Operational/ Bridge	Science	8	Public, Private	All grade 8 schools in the beta and epsilon samples
DA	Operational/ Bridge	Reading, mathematics	12	Public, Private	All grade 12 schools in the beta and epsilon samples
DP	Operational/ Bridge	Puerto Rico mathematics	4, 8	Public	Puerto Rico alpha samples
NP	Pilot	Reading, mathematics, reading router	4, 8	Public	All schools in the pi samples (except Puerto Rico)
MP	Pilot	Puerto Rico mathematics	4, 8	Public	Puerto Rico pi samples

III.   Sample Types and Sizes

In similar fashion to past years, we have identified five different types of school samples: alpha, beta, delta, epsilon, and pi. These distinguish sets of schools that will be conducting distinct portions of the assessment.

1.   *Alpha Samples*

These are public school samples for grades 4 and 8. They will be used for the operational/bridge state-by-state assessments in reading and mathematics and contribute to the national samples for these subjects as well. There will be alpha samples for each state, District of Columbia (DC), Department of Defense Education Activity (DoDEA) schools, Bureau of Indian Education (BIE) schools, and Puerto Rico.

The details of the target student sample sizes for the alpha samples are as follows:

A.   At each grade, the assessed student target sample size is 3,500 per state. The goal in each state (before considering the contribution of TUDA districts) is to roughly assess 1,750 students for math and 1,750 students for reading. The initial target sample size of students after considering attrition is 4,100 for grade 4 and 4,200 for grade 8.

B.   There will be samples for twenty-seven TUDA districts. For the six large TUDA districts (New York, Los Angeles, Chicago, Miami-Dade, Clark County, and Houston) the assessed student target sample sizes are three-quarters the size of a state sample (2,625). The target student sample size after considering attrition is 3,075 for grade 4 and 3,150 for grade 8.

C.   For the remaining twenty-one TUDA districts, the assessed student target sample sizes are half the size of a state sample (1,750). The target student sample size after inflation to account for attrition is 2,050 for grade 4 and 2,100 for grade 8.

D.   Note that, above, there is a conflict between sample size requirements at the state level, and the TUDA district level. This will be resolved as in previous years: the districts will have the target samples indicated in B and C, and reflected in Table 3. For the states that contain one or more of these districts, the target sample size indicated in A (and shown in Table 3) will be used to determine a school sampling rate for the state, which will be applied to the balance of the stateoutside the TUDA district(s). Thus the target student sample sizes, shown in Table 3, for states that contain a TUDA district, are only 'design targets', and are smaller than the final total sample size for the state, but larger than the sample for the balance of the state, exclusive of its TUDA districts.

E.   In Puerto Rico, the target sample size is 4,800 per grade (grades 4 and 8), with the goal of assessing 4,000 students.

As in past state-by-state assessments, schools with fewer than 20 students in the grade in question will be sampled at a moderately lower rate than other schools (at least half, and often higher, depending upon the size of the school). This is in implicit recognition of the greater cost and burden associated with surveying these schools.

As mentioned above, the NAEP 2024 design includes an oversample of relatively high proportion AIAN schools in certain states (as part of the NIES design). These schools will be sampled at higher rates than the

other schools. The NIES oversample will take place in Arizona, Minnesota, Nebraska, North Carolina, Oregon, Utah, Washington, Wisconsin, and Wyoming. Table 3 below shows the thresholds used to define the NIES oversampling strata along with their corresponding oversampling factors. Schools with AIAN percentages that meet or exceed the thresholds will be separately stratified, as shown below, and oversampled by factors ranging from 3 to 6 based on state and grade.

Table 3.   Percent AIAN thresholds and oversampling factors for the NIES school oversample by state and grade

State	Grade 4		Grade 8	
	Percent AIAN thresholds	Oversampling factor	Percent AIAN thresholds	Oversampling factor
Arizona	50	4	50	3
Minnesota	10	6	10	5
North Carolina	5	6	10	6
Nebraska	10	6	10	6
Oregon	5	6	5	6
Utah	4	6	5	6
Washington	6	6	7	6
Wisconsin	8	6	8	6
Wyoming	15	6	25	3

Table 4 shows the target student sample sizes, and the approximate counts of schools to be selected in the alpha samples, along with the school and student frame counts, by state and TUDA districts for grades 4 and 8. The table also identifies the jurisdictions where we take all schools and where we take all students.

Table 5 consolidates the target student (and resulting school) sample size numbers, to show the total target sample sizes in each state, combining the TUDA targets with those for the balance of the state.

Table 4.   Grade 4 and 8 school and student frame counts, expected school sample sizes, and initial target student sample sizes for the 2024 state-by-state and TUDA district assessments (Alpha samples)

Jurisdiction	Grade 4					Grade 8				
	Schools in frame	Schools in sample	Students in frame	Overall target student sample size		Schools in frame	Schools in sample	Students in frame	Overall target student sample size	
Alabama	695	87	55,049	4,100		440	87	56,397	4,200	
Alaska	345	144	8,243	4,100		266	108	8,693	4,200	
Arizona	1,256	90	81,571	4,100		847	91	84,103	4,200	
Arkansas	470	87	35,717	4,100		310	85	37,562	4,200	
Bureau of Indian Education	137	137	2,583	2,583	**	112	112	2,437	2,437	**
California	6,171	87	425,185	4,100		3,088	88	431,483	4,200	
Colorado	1,100	92	61,636	4,100		624	90	64,834	4,200	
Connecticut	554	87	35,879	4,100		287	86	37,515	4,200	
Delaware	119	79	10,276	4,100		70	51	11,048	4,200	
District of Columbia	127	93	6,245	4,100		70	70	5,552	4,200	*
DoDEA Schools	90	90	5,601	5,601	**	55	55	4,427	4,427	**
Florida	2,327	85	208,179	4,100		1,345	88	213,850	4,200	
Georgia	1,254	83	124,867	4,100		590	85	134,510	4,200	
Hawaii	207	89	13,795	4,100		86	55	12,186	4,200	
Idaho	398	92	23,178	4,100		227	87	24,448	4,200	
Illinois	2,114	91	131,275	4,100		1,416	91	140,398	4,200	
Indiana	1,013	86	75,223	4,100		496	86	79,274	4,200	
Iowa	614	92	35,335	4,100		354	88	37,203	4,200	
Kansas	691	97	33,802	4,100		396	94	36,016	4,200	
Kentucky	725	88	46,863	4,100		422	90	49,507	4,200	
Louisiana	737	89	50,543	4,100		502	89	50,568	4,200	
Maine	308	112	12,200	4,100		200	93	12,919	4,200	
Maryland	888	85	64,913	4,100		366	86	65,642	4,200	
Massachusetts	948	87	65,014	4,100		498	86	68,916	4,200	
Michigan	1,664	91	99,728	4,100		1,101	92	103,412	4,200	
Minnesota	990	93	61,527	4,100		729	95	65,001	4,200	
Mississippi	400	87	32,598	4,100		272	85	34,639	4,200	
Missouri	1,156	95	64,682	4,100		713	94	68,508	4,200	

Table 4.   Grade 4 and 8 school and student frame counts, expected school sample sizes, and initial target student sample sizes for the 2024 state-by-state and TUDA district assessments (Alpha samples) (Continued)

Jurisdiction	Grade 4					Grade 8				
	Schools in frame	Schools in sample	Students in frame	Overall target student sample size		Schools in frame	Schools in sample	Students in frame	Overall target student sample size	
Montana	390	129	11,327	4,100		273	100	11,852	4,200	
Nebraska	515	103	23,441	4,100		296	95	24,000	4,200	
Nevada	448	86	35,609	4,100		196	85	36,338	4,200	
New Hampshire	269	103	12,197	4,100		150	78	12,790	4,200	
New Jersey	1,353	88	92,373	4,100		799	87	97,658	4,200	
New Mexico	445	97	22,486	4,100		246	90	24,103	4,200	
New York	2,537	87	182,623	4,100		1,578	87	188,477	4,200	
North Carolina	1,522	87	111,202	4,100		805	87	117,697	4,200	
North Dakota	266	121	9,067	4,100		186	90	8,841	4,200	
Ohio	1,662	89	120,052	4,100		1,101	89	125,713	4,200	
Oklahoma	832	95	49,132	4,100		586	94	50,449	4,200	
Oregon	766	95	39,389	4,100		425	92	42,428	4,200	
Pennsylvania	1,536	85	122,312	4,100		883	86	128,573	4,200	
Puerto Rico	534	199	18,978	4,800		354	196	18,085	4,800	
Rhode Island	166	86	9,914	4,100		64	64	10,239	4,200	*
South Carolina	660	85	56,842	4,100		329	85	59,147	4,200	
South Dakota	317	121	10,493	4,100		256	99	10,925	4,200	
Tennessee	1,009	87	73,203	4,100		609	88	73,138	4,200	
Texas	4,693	85	385,177	4,100		2,402	87	409,871	4,200	
Utah	662	85	50,111	4,100		271	86	52,437	4,200	
Vermont	205	135	5,715	4,100		114	87	5,624	4,200	
Virginia	1,109	85	90,057	4,100		380	84	95,430	4,200	
Washington	1,280	89	77,521	4,100		631	89	81,131	4,200	
West Virginia	385	104	17,514	4,100		191	88	18,892	4,200	
Wisconsin	1,078	96	56,271	4,100		652	92	58,944	4,200	
Wyoming	176	101	6,927	4,100		90	63	7,439	4,200	

Table 4.   Grade 4 and 8 school and student frame counts, expected school sample sizes, and initial target student sample sizes for the 2024 state-by-state and TUDA district assessments (Alpha samples) (Continued)

Jurisdiction	Grade 4					Grade 8				
	Schools in frame	Schools in sample	Students in frame	Overall target student sample size		Schools in frame	Schools in sample	Students in frame	Overall target student sample size	
Albuquerque	103	46	5,703	2,050		49	35	5,719	2,100	
Atlanta	55	40	3,979	2,050		26	26	3,712	2,100	*
Austin	77	42	5,597	2,050		20	20	5,121	2,100	*
Baltimore City	111	46	6,216	2,050		85	46	5,184	2,100	
Boston	71	49	3,270	2,050		46	46	3,298	2,200	*
Charlotte	114	42	10,283	2,050		48	33	10,675	2,100	
Chicago	458	73	23,660	3,075		455	74	24,367	3,150	
Clark County, NV	240	62	22,140	3,075		75	55	23,031	3,150	
Cleveland	65	55	2,489	2,050		64	64	2,437	2,437	**
Dallas	152	42	10,470	2,050		47	35	9,728	2,100	
Denver	110	45	6,318	2,050		61	39	6,390	2,100	
Detroit	73	47	3,736	2,050		60	46	3,094	2,100	
Duval County, FL	127	42	10,046	2,050		60	35	9,103	2,100	
Fort Worth	85	43	5,316	2,050		31	31	5,390	2,100	*
Guilford County, NC	74	44	4,912	2,050		31	31	5,268	2,100	*
Hillsborough County, FL	191	42	17,237	2,050		106	42	17,074	2,100	
Houston	175	63	14,589	3,075		57	44	11,958	3,150	
Jefferson County, KY	100	42	6,981	2,050		42	27	7,106	2,100	
Los Angeles	496	63	34,356	3,075		129	60	30,031	3,150	
Miami	294	64	23,438	3,075		198	65	25,322	3,150	
Milwaukee	110	50	4,968	2,050		79	45	4,481	2,100	
New York City	823	66	58,855	3,075		513	65	59,688	3,150	
Orange County, FL	171	43	15,119	2,050		84	41	15,561	2,100	
Philadelphia	148	45	9,011	2,050		120	43	8,069	2,100	
San Diego	122	45	7,519	2,050		38	38	6,210	2,100	*
Shelby County, TN	117	43	8,246	2,050		65	38	7,517	2,100	
District of Columbia PS	79	48	3,792	2,050		25	25	2,771	2,100	*

Counts for states *do not* reflect the oversampling for their constituent TUDA districts, nor the impact of oversampling for NIES.

Target student sample sizes reflect sample sizes prior to attrition due to exclusion, ineligibility, and nonresponse.

* identifies jurisdictions where all schools (but not all students) for the given grade are included in the NAEP sample. The decision to include all schools in Boston at grade 8 results in a slightly larger overall target student sample size there (2,200 instead of 2,100).

** identifies jurisdictions where all students for the given grade are included in the NAEP sample.

Table 5.   Total sample sizes, combining state and TUDA samples

Jurisdiction	Grade 4					Grade 8				
	Schools in frame	Schools in sample	Students in frame	Overall target student sample size		Schools in frame	Schools in sample	Students in frame	Overall target student sample size	
Alabama	695	87	55,049	4,100		440	87	56,397	4,200	
Alaska	345	144	8,243	4,100		266	107	8,693	4,200	
Arizona	1,256	90	81,571	4,100		847	91	84,103	4,200	
Arkansas	470	87	35,717	4,100		310	85	37,562	4,200	
Bureau Of Indian Education	137	137	2,583	2,583	**	112	112	2,437	2,437	**
California	6,171	186	425,185	8,820		3,088	179	431,483	9,096	
Colorado	1,100	127	61,636	5,728		624	120	64,834	5,885	
Connecticut	554	87	35,879	4,100		287	86	37,515	4,200	
Delaware	119	79	10,276	4,100		70	51	11,048	4,200	
District Of Columbia	127	93	6,245	4,100		70	70	5,552	4,342	*
DoDEA Schools	90	90	5,601	5,601	**	55	55	4,427	4,427	**
Florida	2,327	249	208,179	12,027		1,345	243	213,850	12,332	
Georgia	1,254	120	124,867	6,019		590	108	134,510	6,184	
Hawaii	207	89	13,795	4,100		86	55	12,186	4,200	
Idaho	398	92	23,178	4,100		227	87	24,448	4,200	
Illinois	2,114	146	131,275	6,435		1,416	148	140,398	6,620	
Indiana	1,013	86	75,223	4,100		496	86	79,274	4,200	
Iowa	614	92	35,335	4,100		354	88	37,203	4,200	
Kansas	691	97	33,802	4,100		396	94	36,016	4,200	
Kentucky	725	118	46,863	5,539		422	105	49,507	5,697	
Louisiana	737	89	50,543	4,100		502	89	50,568	4,200	
Maine	308	112	12,200	4,100		200	93	12,919	4,200	
Maryland	888	122	64,913	5,758		366	124	65,642	5,969	
Massachusetts	948	132	65,014	5,945		498	127	68,916	6,199	
Michigan	1,664	134	99,728	5,996		1,101	134	103,412	6,174	
Minnesota	990	93	61,527	4,100		729	95	65,001	4,200	
Mississippi	400	87	32,598	4,100		272	84	34,639	4,200	
Missouri	1,156	95	64,682	4,100		713	94	68,508	4,200	
Montana	390	129	11,327	4,100		273	100	11,852	4,200	

Table 5.   Total sample sizes, combining state and TUDA samples (Continued)

Jurisdiction	Grade 4					Grade 8				
	Schools in frame	Schools in sample	Students in frame	Overall target student sample size		Schools in frame	Schools in sample	Students in frame	Overall target student sample size	
Nebraska	515	103	23,441	4,100		296	95	24,000	4,200	
Nevada	448	96	35,609	4,620		196	88	36,338	4,685	
New Hampshire	269	103	12,197	4,100		150	78	12,790	4,200	
New Jersey	1,353	88	92,373	4,100		799	87	97,658	4,200	
New Mexico	445	119	22,486	5,101		246	104	24,103	5,298	
New York	2,537	124	182,623	5,853		1,578	125	188,477	6,019	
North Carolina	1,522	160	111,202	7,640		805	139	117,697	7,831	
North Dakota	266	121	9,067	4,100		186	90	8,841	4,200	
Ohio	1,662	141	120,052	6,065		1,101	150	125,713	6,556	
Oklahoma	832	95	49,132	4,100		586	94	50,449	4,200	
Oregon	766	95	39,389	4,100		425	92	42,428	4,200	
Pennsylvania	1,536	124	122,312	5,848		883	122	128,573	6,036	
Puerto Rico	534	199	18,978	4,800		354	196	18,085	4,800	
Rhode Island	166	86	9,914	4,100		64	64	10,239	4,200	*
South Carolina	660	85	56,842	4,100		329	85	59,147	4,200	
South Dakota	317	121	10,493	4,100		256	99	10,925	4,200	
Tennessee	1,009	121	73,203	5,688		609	117	73,138	5,868	
Texas	4,693	267	385,177	12,941		2,402	210	409,871	13,319	
Utah	662	85	50,111	4,100		271	86	52,437	4,200	
Vermont	205	135	5,715	4,100		114	87	5,624	4,200	
Virginia	1,109	85	90,057	4,100		380	84	95,430	4,200	
Washington	1,280	89	77,521	4,100		631	89	81,131	4,200	
West Virginia	385	104	17,514	4,100		191	87	18,892	4,200	
Wisconsin	1,078	137	56,271	5,786		652	129	58,944	5,980	
Wyoming	176	101	6,927	4,100		90	63	7,439	4,200	
Total	52,313	6,269	3,561,640	270,092		29,749	5,660	3,711,269	276,155	

Sample sizes for each state reflect the samples in the TUDA districts within the state, but do not reflect the impact of NIES oversampling.

* identifies jurisdictions where all schools (but not all students) for the given grade are included in the NAEP sample.

** identifies jurisdictions where all students for the given grade are included in the NAEP sample.

Stratification

Each state and grade will be stratified separately but using a common approach in all cases. TUDA districts will be separated from the balance of their state, and each part stratified separately. The first level of stratification will be based on urban-centric type of location. This variable has 12 levels (some of which may not be present in a given state or TUDA district), and these will be collapsed so that each of the resulting location categories contains at least 10 percent of the student population (13 percent for large TUDA districts and 20 percent for small TUDA districts).

Within each of the resulting location categories, schools will be assigned a minority enrollment status. This is based on the two race/ethnic groups that are the second and third most prevalent within the location category. If these groups are both low in percentage terms, no minority classification will be used. Otherwise, three (or occasionally four) equal-sized groups (generally high, medium, and low minority) will be formed based on the distribution across schools of the two minority groups.

Within the resulting location and minority group classes (of which there are likely to be from two to twenty-three, depending upon the jurisdiction), schools will be sorted by a measure derived from school level results from the most recent available state achievement tests at the relevant grade. In general, mathematics test results will be used, but where these are not available, reading results will be used. In the few states that do not have math or reading tests at grades 4 and 8 (or where we are unable to match the results to the NAEP school frame), instead of achievement data, schools will be sorted using a measure of socio-economic status. This is the median household income of the 5-digit ZIP Code area where the school is located, based on the 2021 ACS (5-year) data. For BIE and DoDEA schools neither achievement data nor income data are available, and so grade enrollment is used in these cases.

Once the schools are sorted in a serpentine fashion by location class, minority enrollment class, and achievement data (or household income/grade enrollment), a systematic sample of schools will be selected using a random start. Schools will be sampled with probability proportional to size.

2.   *Beta Samples*

The beta samples include the national public-school samples at grades 8 and 12. At grade 8, the beta sample will be used for the national science assessment. At grade 12, the beta sample will be used for the national mathematics and reading assessments. Each of the samples will be nationally representative. The grade 8 sample will be selected to have minimal overlap with the alpha sample schools at grade 8. The number of students targeted for selection per school will be 50.

To increase the likelihood that the results for AIAN students can be reported for these samples, we will oversample public schools with a relatively high-AIAN student population. That is, a public school with at least 5 AIAN students and at least 5 percent AIAN enrollment will be given four times the chance of selection of a public school of the same size with a lower AIAN percentage. For all other schools, whenever there are at least 10 Black or Hispanic students enrolled and the combined Black and Hispanic enrollment at least 15 percent, the school will be given twice the chance of selection of a public school of the same size with a lower percentage of these two groups. This approach is effective in increasing the sample sizes of AIAN, Black, and Hispanic students without inducing undesirably large design effects on the sample, either overall, or for particular subgroups.

Stratification

The beta samples will have an implicit stratification, using a hierarchy of stratifiers and a serpentine sort. The highest level of the hierarchy is high/low AIAN status. The second stratifier is Census division (10 categories: the usual 9 plus California as a separate category). Some of the Census divisions within the high AIAN stratum will be collapsed with neighboring Census divisions (this will occur if the expected school sample size within the cell is less than 4.0).

The next stratifier in the hierarchy is type of location, which has twelve categories. Within the high AIAN stratum, the cells will likely be too small to further stratify by type of location. Within the low AIAN stratum, many of the type of location strata nested within Census divisions will be collapsed with neighboring type of location cells (this will occur if the expected school sample size within the cell is less than 4.0).

These geographic strata will be subdivided into high/low Black and Hispanic substrata. If the expected initial sample size in a Black/Hispanic substratum is less than 8.0, it will be left as is. If the expected sample size is greater than or equal to 8.0, then it will be subdivided into up to four substrata (two for expected sample size up to but less than 12.0, three for expected sample size up to but less than 16.0, and four for expected sample size greater than or equal to 16.0). For the high Black/Hispanic substrata, the subdivision will be by percentage Black and Hispanic. For the low Black/Hispanic substrata, the subdivision will be by state or groups of contiguous states.

Within these substrata, the schools are to be sorted by school type (public, BIE, DoDEA) and median household income from the 2021 5-year ACS (using a serpentine sort within the school type substrata).

3. *Delta Samples*

These are the private school samples at grades 4 and 8 for conducting the operational/bridge study assessments in reading and mathematics. The sample sizes are large enough to report results by Catholic and non-Catholic at grades 4 and 8 (of course participation rate standards must be met). Approximately half the sample at each grade will be from Catholic schools. The grade 8 delta samples will be selected to have minimum overlap with the grade 4 delta sample. The number of students targeted per school will be 50 at each grade.

Stratification

The private schools are to be explicitly stratified by private school type (Catholic/Other). Within each private school type, implicit stratification will be by Census region (4 categories), type of location (12 categories), race/ethnicity composition, and enrollment size. In general, where there are few or no schools in a given stratum, categories will be collapsed together, always preserving the private school type.

4. *Epsilon Samples*

With regard to subjects and grades assessed, the epsilon samples are analogous to the beta samples, but for private schools. However, in contrast to the beta samples, there will be no oversampling of high AIAN or high Black and Hispanic schools. The same stratification variables will be used as for the delta

samples. The epsilon sample schools at grade 8 will have minimum overlap with the delta grade 8 sample schools which, given the respective sample sizes, means that no schools will be selected for both the delta and epsilon samples at grade 8. The number of students targeted per school will be 50.

5. *Pi Samples*

These are the public-school samples at grades 4 and 8 for conducting the pilot test assessments in reading and mathematics, including the reading router pilot. The pi sample will also include the pilot assessments for Puerto Rico mathematics at grades 4 and 8. These samples will be selected from a sample of 66 geographic Primary Sampling Units (PSUs) plus Puerto Rico. Schools selected in the alpha and beta samples for grades 4 and 8 will be avoided. The number of students targeted per school will be 50 for schools in the reading and mathematics pilot and 25 for schools in the Puerto Rico mathematics pilot.

Primary Sampling Unit (PSU) Selection

The samples for the pilot test assessments will be based on a clustered design using PSUs. The design will be based on a sample of 66 PSUs selected from a frame of approximately 1,000 PSUs. In this design, the 29 largest MeSAs will be selected with certainty, and the remaining sample will be selected with probability proportional to size (PPS) where the size measure is a function of the number of children as given in the most recent population estimates prepared by the U.S. Census Bureau. For the stratification, 76 minor strata will be formed and paired, and a single PSU will be selected from one stratum in each of the 38 pairs for a total of 67 PSUs. To keep field costs down and because there will be no weights for the pilot tests, the plan is to exclude the PSU in Honolulu, HI. Therefore, there will be a 66-PSU sample for the pilot test samples.

For Puerto Rico, there is not a PSU selection stage.

Stratification

For the pilot test samples, explicit stratification will take place at the PSU level. Although a nationally representative sample is not required, stratification will be used for the purpose of having a good variety of area characteristics, as well as school and student characteristics to serve the purposes of the pilot tests. For schools within PSUs, stratification will occur by sorting the school file prior to systematic selection. As in past national samples, the expectation is that, within the set of certainty MeSA PSUs within a census region, PSU will not necessarily be the highest level sort variable. Thus, type of location will be used as the primary sort variable. The design is aimed primarily at getting the correct balance of city, suburban, town, and rural schools, as a priority over getting exactly a proportional representation from each MeSA. The sort of the schools will use other variables beyond the type of location variable, such as a race/ethnicity percentage variable. No oversampling of schools or students will occur in the pilot test samples. Schools will be selected with probabilities proportionate to size.

For Puerto Rico, the sample will not be clustered. The sampling frame of schools will be stratified using the type of location variable. No oversampling of schools or students will occur in the pilot tests in Puerto Rico. Schools will be selected with probabilities proportionate to size.

The preliminary 2021-22 CCD serves as the basis for the public-school frames for the pilot test samples. The PSS file is not needed since the pilot test samples do not involve private schools.

6.   *School-based Equipment Proof of Concept*

The SBE POC will be fielded shortly after the main NAEP data collection in 2024. It will consist of 1,000 assessed students in each of grades 4 and 8. There will be a target of 26 participating schools in grade 4 and 25 in grade 8. The sample of schools will be a non-probability sample. The details of the sampling for the SBE POC is to be developed further.

IV.   New Schools

To compensate for the fact that files used to create the NAEP school sampling frames are at least two years out of date at the time of frame construction, we will supplement the alpha, beta, delta, and epsilon samples with new school samples at each grade.

The new school samples will be drawn using a two-stage design. At the first stage, a minimum of ten school districts (in states with at least ten districts) will be selected from each state for public schools, and ten Catholic dioceses will be selected nationally for the private schools. The sampled districts and dioceses will be asked to review lists of their respective schools and identify new schools. Frames of new schools will be constructed from these updates, and new schools will be drawn with probability proportional to size using the same sample rates as their corresponding original school samples.

The school sample sizes in the above tables do not reflect new school samples.

V.   Substitute Samples

Substitute samples will be selected for each of the beta, delta, and epsilon samples. The substitute school for each original will be the next "available" school on the sorted sampling frame, with the following exceptions:

A.   Schools selected for any NAEP samples will not be used as substitutes.

B.   Private schools whose school affiliation is unknown will not be used as substitutes. Also, unknown affiliated private schools in the original samples will not get substitutes.

C.   New schools will not get substitutes.

D.   A school can be a substitute for one and only one sample. (If a school is selected as a substitute school for grade 12, for example, it cannot be used as a substitute for either grade 4 or grade 8.)

E.   A public-school substitute will always be in the same state as its original school.

F.   A Catholic school substitute will always be a Catholic school, and the same for non-Catholic schools.

VI.   Contingency Samples

The districts that are taking part in the TUDA program are volunteers. Thus it is possible that at some point over the next few months, a given district might choose to opt out of the TUDA program for 2024. However, it is not acceptable for all schools in such a district to decline NAEP, as then the state estimates will be adversely affected. Thus to deal with this possibility, in each TUDA district, subsamples of the alpha sample schools will be identified as contingency samples. In the event that the district withdraws from the TUDA program prior to the selection of the student sample, all alpha sampled schools from that district will be dropped from the sample, with the exception of those selected in the contingency sample. The contingency sample will provide a proportional representation of the district, within the aggregate state sample. Student sampling in those schools will then proceed in the same way as for the other schools within the same state.

VII.   Student Sampling

Students within the sampled schools will be selected with equal probability. The student sampling parameters vary by sample type (alpha, beta, delta, epsilon, and pi) and grade as described below.

Alpha Sample, Grades 4 and 8 Schools (Except Puerto Rico)

A.   All students, up to 52, will be selected.

B.   If the school has more than 52 students, a systematic sample of 50 students will be selected. In some schools, the school may be assigned more than one 'hit' in sampling. In these schools we will select a sample of size 50 times the number of hits, taking all students if this target is greater than or equal to 50/52 of the total enrollment.

Alpha Sample, Puerto Rico Grades 4 and 8

A.   All students, up to 26, will be selected.

B.   If the school has more than 26 students, a systematic sample of 25 students will be selected.

Delta Samples, Grades 4 and 8

A.   All students, up to 52, will be selected.

B.   If the school has more than 52 students, a systematic sample of 50 students will be selected.

Beta Sample, Grades 8 and 12

A.   All students, up to 52, will be selected.

B.   If the school has more than 52 students, a systematic sample of 50 students will be selected.

Epsilon Sample, Grades 8 and 12

A.   All students, up to 52, will be selected.

B.   If the school has more than 52 students, a systematic sample of 50 students will be selected.

Pi Samples, Grades 4 and 8 (Except Puerto Rico)

A.   All students, up to 52, will be selected.

B.   If the school has more than 52 students, a systematic sample of 50 students will be selected.

Pi Samples, Puerto Rico Grades 4 and 8

A.   All students, up to 26, will be selected.

B.   If the school has more than 26 students, a systematic sample of 25 students will be selected.

VIII.   Weighting Requirements

The Operational Reading and Mathematics Assessments, Grades 4 and 8

The exact weighting requirements for these samples have yet to be determined. One likely possibility is that three sets of student weights will be required – for assessments conducted on Surface Pros (SP) alone, on Chromebooks (CB) alone, and SP/CB combined. The samples will have student weights for each subject (reading and math) applied to reflect probabilities of selection, school and student nonresponse, any trimming, and the random assignment to the particular subject. There will be separate replication schemes by grade and public/private. Weights will also be derived for the Puerto Rico assessment at grades 4 and 8.

The Operational Science Assessment, Grade 8

The samples will have a single set of student weights for science applied to reflect probabilities of selection, school and student nonresponse, any trimming. There will be separate replication schemes by public/private.

The Operational Reading and Mathematics Assessments, Grade 12

The samples will have a single set of student weights for each subject (reading and math) applied to reflect probabilities of selection, school and student nonresponse, any trimming, and the random assignment to the particular subject. There will be separate replication schemes by public/private.

Pilot Tests for Reading and Mathematics, Grades 4 and 8

As is standard practice, only preliminary weights will be provided for these assessments. The sample weights will reflect probabilities of selection, and the random assignment to the particular subject.

No weights will be provided for the SBE POC.

School Weights

In addition to student weights, each sample described above will have a set of school weights to provide secondary users a means to analyze data at the school level. Each sample will have a single set of school weights for each subject (reading, math, or science) applied to reflect probabilities of selection, school nonresponse, any trimming, and a small-school adjustment to account for schools too small to do both subjects associated to their respective samples. There will be separate replication schemes by public/private.