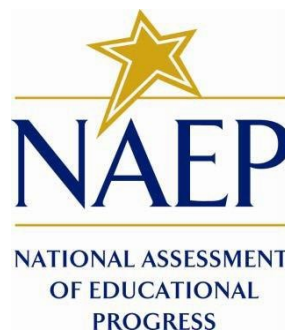


*NATIONAL CENTER FOR EDUCATION STATISTICS
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS*

National Assessment of Educational Progress (NAEP) 2027

*Appendix C
2027 Sampling Memo*

OMB# 1850-0928 v.40



July 2026

Note: The final 2027 Sampling Memo is included in this 2027 Amendment submission.



IMPROVING LIVES THROUGH RESEARCH®

Date:	April 17, 2026		Memo:	2027-m01v02s
To:	Gina Ruehl, NCES	Lisa Rodriguez		
	Enis Dogan, NCES	Lauren Byrne		
	Alex McMillan, NCES	Marcie Hickman		
	Amy Drescher, ETS	Jason Schuknecht		
	Peter Ciemins, ETS	William Wall		
	Yue Jia, ETS	Stephen Bernas		
	Robert Finnegan, ETS	Jay Clark		
	Adrienne Sgammato, ETS	Veronique Lieber		
	Laura Jerry, ETS	Yiting Long		
		Sabrina Zhang		

1. Introduction

For the 2027 NAEP year, the sample design involves a pilot test in science at grade 8. The pilot sample will be used to test new science items. Below is a summary list of the features of the 2027 sample design.

1. The pilot test is a digitally based assessment administered on computer devices: either NAEP-provided devices (Chromebooks, ND) or school devices (SD). It will be administered in schools located in the 50 states and the District of Columbia.
2. The pilot test will employ a three-stage design: selecting PSUs, schools, and students. A sample of 34 PSUs has been selected to reduce overlap with the 2025 Long-Term Trend PSUs and the 2025 Field Test PSUs to the extent possible.
3. The pilot test will be administered using the mixed device model only. This means that the School Technology Survey (STS) will determine if a school qualifies for school devices and if so, school devices will be used. A large majority of all schools is expected to qualify for school devices. The schools that do not qualify will be assessed using NAEP-provided devices.
4. The pilot samples will not employ any oversampling related to Black, Hispanic, or AI/AN students, either at the school level or at the student level.
5. The primary focus of the pilot test is public school students. However, to gain experience with administering the new science items in private schools and to continue correspondence with private organizations, the pilot test will be administered on a small scale in private schools.

The targets of assessed students are shown in Table 1 (which also shows the approximate numbers of participating schools and average number of assessed students per school).

Table 1. Targets of assessed students and expected number of participating schools for 2027 NAEP Pilot Test, grade 8 science

Component	Public school students	Private school students	Total
Assessed students	11,340	660	12,000
Participating schools	291	30	321
Average assessed students/school	39	22	37

2. Assessment Types

The assessment types for 2027 are shown in Table 2.

Table 2. NAEP 2027 assessment types

Type	Subject	Grade	Schools	Comments
Pilot	Science	8	Public, private	Pilot sample

3. Sample Design and Sample Sizes

As mentioned above, NAEP 2027 consists of a pilot assessment in science at grade 8. To supplement the overview presented in sections 1 and 2 above, a summary of major points is included in the sections that follow.

Primary Sampling Units Selection

The sample for the pilot assessments is based on a clustered design using primary sampling units (PSUs) for reasons of operational efficiency. A sample of 34 PSUs was selected from a frame of approximately 980 PSUs (the same frame used to select PSUs for the NAEP 2025 long-term trend and field test samples). A sample of 34 PSUs provides a reasonable workload per PSU in terms of anticipated cooperating schools. It also fits nicely within the structure of our existing PSU sampling frame and stratification scheme, discussed further below in this section.

All sampled schools were drawn from within the sampled PSUs.

The PSU frame was constructed for NAEP 2025 as follows. The PSUs were created from aggregates of counties. Data on counties were obtained from the 2020 Census Demographic and Housing Characteristics File (DHC). Each Metropolitan Statistical Area (MeSA) constitutes a PSU, except that MeSAs that cross Census region boundaries were split into separate PSUs according to Census region boundaries. Non-metropolitan PSUs were formed by aggregating counties into geographic units of sufficient minimum size to provide enough schools to constitute a workload of about 1% of the total sample. These PSUs were made of contiguous counties where possible, and nearly contiguous counties (separated by MeSA counties) otherwise. Thirty-two PSUs were identified as certainties¹. The remaining PSUs on the frame were assigned to 76 noncertainty strata using an index that distinguished PSUs based on homogeneity. The index was generated by a LASSO regression that

¹ Certainty PSUs were identified assuming a target sample size of 105 PSUs, which was our large PSU sample size from 2013-2024. 25 certainties were identified this way. Then to allow for population growth over time, PSUs with measures of size within 90% of the size of the smallest certainty were also identified as certainties. Finally, Honolulu was identified as a certainty due to its unique characteristics as had been done in the past.

identified PSU characteristics related to NAEP achievement in past assessments among a set of variables representative of NAEP reporting subgroups.

For NAEP 2027, a sample of 34 PSUs was selected for the pilot samples. The four largest MeSAs were selected with certainty as determined analytically in accordance with probability proportional to size (PPS) methodology. The certainty PSU cutoff was computed as the total MOS (youth population) divided by the number of PSUs to select (34). An arbitrary factor of about 0.9 was used to allow for population growth, similar to past years. The certainty PSU cutoff was 1,913,657 youth. The remaining sample was a stratified PPS sample, where the size measure was a function of the number of children as given in the 2020 Decennial Census. 11 noncertainty PSUs were selected from the 28 remaining PSUs that were identified as certainties in 2025. 19 PSU strata were selected from the 76 noncertainty PSU strata that were formed for 2025, and then 1 PSU was selected from each of the 19 strata. The 19 PSU strata were a predetermined random quarter-sample of the 76 PSU strata to be used in the event that smaller PSU samples were desired.

Overlap Control

A desired feature of the NAEP program is to avoid overlap with other samples. Consider potential overlap at the school level between NAEP 2027 and NAEP 2025. When selecting the 19 PSUs, care was taken to reduce overlap between them and the noncertainty PSUs selected for either long-term trend or field test for NAEP 2025. Within the noncertainty PSU strata, overlap control is managed from year to year by offsetting the random start within each stratum. For example, in 2025, the LTT PSUs were selected first. Then when selecting the 2025 FT PSUs, the random start was shifted by 0.5 in each stratum. In stratum 1, if the random start was 0.25 for LTT, then for FT it was 0.75, and if it was 0.85 for LTT, then it was 0.35 for FT. In this way we are not likely to hit the same PSU twice, which eliminates school overlap by definition. For 2027, the random starts were shifted by 0.25 (placed halfway between what they had been for LTT and FT).

This procedure was only mildly successful because both PSU samples in 2025 were large (108 PSUs each), and because in some strata, some PSUs have large probabilities of selection relative to other PSUs in the same stratum. The 2027 PSU sample has 23 PSUs in common with 2025—the 15 PSUs that were certainties in 2025, plus 8 additional PSUs. The expectation is that PSUs in common to both years are large enough, and the 2027 school sample small enough, that little to no overlap at the school level will occur in the successive PSU-based samples in 2025 and 2027. However, a large state-based sample was selected for NAEP 2026. It is difficult to avoid school overlap with a state NAEP year because the sampling rates in some states and Trends in Urban Districts Assessment (TUDA) districts are quite high. To address undue burden on certain schools, any schools selected for 2025 (either LTT or FT), 2026, and 2027, i.e., selected in all three years, will be dropped from the 2027 sample. There were eight such schools.

Stratification and Oversampling

The plan for the pilot sample design is to draw separate public and private school samples, which has proven advantages:

- 1) it permits the timing of sample selection to vary between public and private schools, should this prove necessary;
- 2) it allows us to readily assume different response and eligibility rates for public and private schools;
- 3) it makes it easier to use different sort variables used in the selection of public schools and private schools; and
- 4) it allows for the possibility of a late change of mind concerning the sample sizes that differ between public and private schools.

For the pilot samples, explicit stratification has taken place at the PSU level. For schools within PSUs, stratification gains will be achieved 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 for public schools. Consider for example the large MeSAs in the Midwest region. The design is aimed primarily at getting the correct balance of city, suburban, town, and rural schools crossed by city size and distance from urbanized areas, as a priority over getting exactly a proportional representation from each MeSA (Chicago, Detroit, Minneapolis), although of course it should be possible to get a high degree of control over both of these characteristics. The sort of the schools will use other variables beyond the type of location variable.

The exact set of variables used in sorting the schools prior to sampling will depend on whether the schools are located in certainty PSUs or noncertainty PSUs and whether the schools are public or private. Public schools in certainty PSUs will be hierarchically sorted by

- 1) census region;
- 2) type of location – within census region, start with city, suburban, town, and rural, but collapse town and rural together or collapse suburban, town, and rural together as needed to yield cells with at least two expected sampled schools;
- 3) race/ethnicity stratum – within type of location nested within census region, sort schools by the percentage of Black, Hispanic, or AI/AN students and estimated enrollment. Then determine cutpoints on the percentage of Black, Hispanic, or AI/AN students and use them to form distinct race/ethnicity strata that will each yield at least two expected sampled schools. No race/ethnicity strata will be formed in census region/type of location cells with fewer than four expected sampled schools; and
- 4) final sort variable – within race/ethnicity stratum nested within type of location nested within census region, the final sort variable will be estimated enrollment if race/ethnicity strata were formed. Otherwise, the final sort variable will be the percentage of Black, Hispanic, or AI/AN students.

Public schools in noncertainty PSUs will be sorted by

- 1) PSU;
- 2) type of location – within PSU, use city, suburban, town, and rural. Not all categories will exist for all PSUs; and
- 3) final sort variable – within type of location nested within PSU, the final sort variable will be the percentage of Black, Hispanic, or AI/AN students.

Private schools in certainty PSUs will be sorted by

- 1) census region;
- 2) school type – within census region, use Catholic, Lutheran, Conservative Christian, other religious, nonsectarian;
- 3) type of Catholic school – within school type nested within census region, use parochial, diocesan, private, unknown; and
- 4) final sort variable – within type of Catholic school nested within school type nested within census region, the final sort variable will be estimated enrollment.

Private schools in noncertainty PSUs will be sorted by

- 1) PSU;
- 2) school type – within PSU, use Catholic, Lutheran, Conservative Christian, other religious, nonsectarian. Not all categories will exist for all PSUs;
- 3) type of Catholic school – within school type nested within PSU, use parochial, diocesan, private, unknown. Not all categories will exist for all PSUs; and
- 4) final sort variable – within type of Catholic school nested within school type nested within PSU, the final sort variable will be estimated enrollment.

In most national assessments, we implement oversampling of certain public schools to 1) increase the likelihood that the results for American Indian/Alaska Native (AI/AN) students can be reported and 2) increase the sample sizes of Black and Hispanic students, without inducing undesirably large design effects on the sample, either overall or for particular subgroups. However, this oversampling will not be conducted for the 2027 pilot assessment because student assessment results will not be reported. This is in line with our approach in past pilot assessments.

The preliminary 2023-24 CCD and the 2023-24 PSS school files serve as the basis for the public and private school frames for the pilot sample.

4. New Schools

In most NAEP assessments, a sample of new schools is selected to compensate for the fact that files used to create the school sampling frame are two years out of date at the time of assessment. However, for the pilot sample, new schools will not be selected because student results will not be reported.

5. Substitute Samples

A portion of the eligible sample schools will choose not to participate in the assessment. In most NAEP assessments, to reduce nonresponse bias, substitute schools are selected and recruited in place of original sample school that refuse.

In keeping with past practice for pilot samples, substitute schools will not be selected for 2027.

6. Deselection Approach

Note: NCES decided against the deselection method described in this section, however the description below remains for documentation purposes. All selected schools will remain in the sample and deselection will not be implemented.

There is some uncertainty regarding the expected school participation rates for the 2027 sample because participation is voluntary and the sample will be drawn in a year when Main NAEP is not being administered. To inform our assumptions, we plan to use the participation and eligibility rates (collectively called a yield rate) observed for the 2025 Long-Term Trend (LTT) Age 13 sample, which achieved an 87 percent public school yield rate.

To help ensure that we meet the target of 12,000 completed assessments, we will initially target 20 percent more public school students than we would under typical sampling procedures. Based on our projections, we anticipate needing 321 participating schools.

In August or September 2026, once we have recruited most public schools for 2027, we will assess whether the number of cooperating schools is likely to exceed 321. If so, we may deselect some schools before the process of submitting student lists begins. Any deselection would occur by ordering the schools according to the original selection sort order and then removing the necessary number of schools using a systematic random sampling method. This process simply involves updating school disposition codes for the deselected schools, and revising the selection probabilities of the remaining schools to reflect the smaller sample. Deselection, if necessary, will apply only to public schools.

Table 3 shows the attrition and inflation factors used in the sampling calculations. The initial target student sample size is determined by starting with the target completed assessments (Row label = “F”,

multiplying it by the inflation factor, and then dividing by the assumed overall yield rate. The resulting overall initial school sample size is generated through the school sampling program and is based on this initial target student sample size (Row label = “A”) as input. In practice, the actual targets used in the sampling program are rounded up slightly, (e.g., 19,000 for public and 2,400 for private). Details on NAEP school sampling methodology on which the school sampling program is based are available in the technical documentation: [Sampling of Public Schools for the 2022 Long-Term Trend Assessment](#). The 2022 long-term trend page was chosen because long-term trend, although age-based instead of grade-based, is also a PSU-based sample, like the 2027 pilot.

Table 3. Derivation of initial target sample sizes

Row Label	Description	Public	Private	Total	Calculation
A	Initial target student sample size	18,935	2,394	21,329	F*E/D
B	School yield rate	0.869	0.308		
C	Student yield rate	0.827	0.895		
D	Overall yield rate	0.719	0.276		B*C
E	Extra sample size inflation rate	1.200	1		
F	Target completed assessments	11,340	660	12,000	
G	Ratio of assessed students to participating schools	39	22		
H	Number of participating schools	291	30	321	F/G

Considerations for Mac OS schools

There are ongoing discussions regarding support for administering the assessment on Mac OS devices that may be provided by schools. If this option becomes available, results from the School Technology Survey (STS) will allow us in July to identify sampled schools that use Mac OS. These schools will initially be assigned NAEP-provided devices. Following OMB approval on September 11, NSSC will develop a communication—delivered by NSCs—informing Mac OS schools of the new option and offering them the opportunity to switch to using school-provided devices.

Because we will already be in direct communication with these schools, and to avoid the disruption of deselecting schools after outreach has begun, SW proposes treating Mac OS schools identified through the STS as certainty selections (within PSUs). Under this approach, the end-of-September deselection process would apply only to non-Mac OS schools. This approach is statistically sound: school base weights and preliminary weights will appropriately reflect schools’ probabilities of selection, preserving the representativeness of the sample.

Current NPD estimates suggest fewer than 30 Mac OS schools, so the impact on the overall sample is expected to be minimal. If July results indicate a meaningfully larger number of Mac OS schools, we would revisit this approach to ensure the sample is not unduly skewed. Any such adjustment would be minor and made in the best interest of the program.

7. Student Sampling

For the pilot sample, students within the sampled schools will be selected with equal probability. In addition to this, student sample sizes within each school are determined as the combined result of two primary factors:

1. We wish to take all students in relatively small schools.

2. We do not wish to overburden the schools with unduly large student samples.

The plan below reflects the design that results from considering both of these factors and balancing them. Prior to sampling, students will be sorted by sex, race, last name, first name, and middle name.

In all participating schools, the target sample size is 50 students. We will select all grade 8 students in schools with 50 students or fewer. In schools with more than 50 such students, we will select 50 students².

8. Weighting Requirements

As is standard practice for pilot assessments, preliminary weights will be provided. The sample weights will reflect probabilities of selection but no adjustments for nonresponse, trimming, etc.

² For nonoperational assessments like this pilot, schools can request that an additional student or students be assessed if their enrollment is slightly more than 50 and they prefer it. In practice, this is applicable to a very small number of schools, if any.