NATIONAL CENTER FOR EDUCATION STATISTICS
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

National Assessment of Education Progress (NAEP) Long-Term Trend (LTT) 2025

Appendix B

NAEP 2012 Long -Term Trend (LTT) Weighting Procedures Design (Latest document available)

OMB# 1850-0928 v.32

October 2023
No changes since V. 15 (Appendix B2)

Note: This document 2012Weighting Procedures Design, will be used for LTT 2025, no new design is avaialble.

NAEP Technical Documentation Website

NAEP Technical Documentation Weighting Procedures for the 2012 Long-Term Trend (LTT) Assessment

NAEP assessments use complex sample designs to create student samples that generate population and subpopulation estimates with reasonably high precision. Student sampling weights ensure valid inferences from the student samples to their respective populations. In the 2012 long term trend (LTT) assessments, weights were developed for students sampled at ages 9, 13, and 17 for assessments in mathematics and reading. Each student was assigned a weight to be used for making inferences about students in the target population. This weight is known as the final full-sample student weight, and it contains five major components:

- the student base weight,
- school nonresponse adjustments,
- student nonresponse adjustments,
- school weight trimming adjustments, and
- student weight trimming adjustments.

The student base weight is the inverse of the overall probability of selecting a student and assigning that student to a particular assessment. The sample design that determines the base weights is discussed in the NAEP 2012 LTT sample design section.

The base weight is adjusted for two sources of nonparticipation: school level and student level. These weighting adjustments seek to reduce the potential for bias from such nonparticipation by

- increasing the weights of students from schools similar to those schools not participating, and
- increasing the weights of participating students similar to those students from within participating schools who did not attend the assessment session (or makeup session) as scheduled.

Furthermore, the final weights reflect the trimming of extremely large weights at both the school and student level. These weighting adjustments seek to reduce variances of survey estimates.

In addition to the final full-sample weight, a set of replicate weights was provided for each student. These replicate weights are used to calculate the variances of survey estimates using the jackknife repeated replication method. The methods used to derive these weights were aimed at reflecting the features of the sample design, so that when the jackknife variance estimation procedure is implemented, approximate unbiased estimates of sampling variance are obtained. In addition, the various weighting procedures were repeated on each set of replicate weights to appropriately reflect the impact of the weighting adjustments on the sampling variance of a survey estimate.

Quality control checks were implemented throughout the weighting process to ensure the accuracy of the full-sample and replicate weights. See Quality Control for Weighting Procedures for the various checks implemented and main findings of interest.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Computation of Full-Sample Weights for the 2012 LTT Assessment

The full-sample or final student weight is the sampling weight used to derive NAEP student estimates of population and subpopulation characteristics for a specified age (9, 13, or 17) and assessment subject (mathematics or reading). The full-sample student weight reflects the number of students that the sampled student represents in the population for purposes of estimation. The summation of the final student weights over a particular student group provides an estimate of the total number of students in that group within the population.

The full-sample weight, which is used to produce survey estimates, is distinct from a replicate weight that is used to estimate variances of survey estimates. The full-

sample weight is assigned to participating students and reflects the student base weight after the application of the various weighting adjustments. The full-sample weight for student *k* from school *s* in stratum *j* ($FSTUWGT_{jsk}$) can be expressed as follows:

$$FSTUWGT_{jsk} = STU\_BWT_{jsk} \times SCH\_NRAF_{js} \times STU\_NRAF_{jsk} \times SCH\_TRIM_{js} \times STU\_TRIM_{jsk}$$

where

- STU_BWT_{jsk} is the student base weight;
- SCH_NRAF_{js} is the school-level nonresponse adjustment factor;
- STU_NRAF_{jsk} is the student-level nonresponse adjustment factor;
- SCH_TRIM_{js} is the school-level weight trimming adjustment factor; and
- STU_TRIM_{jsk} is the student-level weight trimming adjustment factor.

School sampling strata for a given assessment varied by school type. See public school strata and private school strata for descriptions of the public and private school stratum definitions.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_comp_full_samp.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Computation of Base Weights for the 2012 LTT Assessment

Every sampled school and student received a base weight equal to the reciprocal of its probability of selection. Computation of a school base weight varies by

School Base Weights

Student Base Weights

- the type of sampled school (original or substitute); and
- the sampling frame (new school frame or not).

Computation of a student base weight reflects

- the student's overall probability of selection accounting for school and student sampling;
- assignment to session type at the school- and student-level; and
- the student's assignment to the mathematics or reading assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_base.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation School Base Weights for the 2012 LTT Assessment

The school base weight for a sampled school is equal to the inverse of its overall probability of selection. The overall selection probability of a sampled school differs by

- type of sampled school (original or substitute); and
- sampling frame (new school frame or not).

Substitute public schools for the 2012 LTT assessments

Substitute private schools for the 2012 LTT assessments

The overall probability of selection of an originally selected school reflects two components:

- the probability of selection of the primary sampling unit (PSU), and
- the probability of selection of the school within the selected PSU from either the NAEP public school frame or the private school frame.

The overall selection probability of a school from the new school frame is the product of two quantities:

- the probability of selection of the school's district into the new-school district sample, and
- the probability of selection of the school into the new school sample.

Substitute schools are preassigned to original schools and take their place if the original schools refuse to participate. For weighting purposes, they are treated as if they were the original schools that they replaced and are assigned the school base weight of the original schools.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_base_wghts_school.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Substitute Public Schools for the 2012 Long-Term Trend (LTT) Assessment

Substitute schools were preselected for the public school samples by sorting the school frame file according to the actual order used in the sampling process (the implicit stratification). For operational reasons, the original selection order was embedded within the sampled primary sampling unit (PSU) and state. Each sampled school had each of its nearest neighbors within the same sampling stratum on the school frame file identified as a potential substitute. When age-eligible enrollment was used as the last sort ordering variable, the nearest neighbors had age enrollment values very close to that of the sampled school. This was done to facilitate the selection of about the same number of students within the substitute as would have been selected from the original sampled school.

Schools were disqualified as potential substitutes if they were already selected in any of the original public school samples or assigned as a substitute for another public school (earlier in the sort ordering). Schools assigned as substitutes for age 17 schools were disqualified as potential substitutes for age 9 and 13 schools, and schools assigned as substitutes for age 13 schools were disqualified as potential substitutes for age 9 schools.

If both nearest neighbors were still eligible to be substitutes, the one with a closer age-eligible enrollment was chosen. If both nearest neighbors were equally distant from the sampled school in their age enrollment (an uncommon occurrence), one of the two was randomly selected.

Of the approximately 1,100 original sampled public schools for the ages 9, 13, and 17 assessments, about 30 schools had a substitute activated because the original eligible school did not participate. Ultimately, about 20 of the activated substitute public schools participated in an assessment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2012/2012_ltt_samp_pub_subs.aspx

NAEP Technical Documentation Website

NAEP Technical Documentation Substitute Private Schools for the 2012 Long-Term Trend (LTT) Assessment

Substitutes were preselected for the private school samples by sorting the school frame file according to the actual order used in the sampling process (the implicit stratification). For operational reasons, the original selection order was embedded within the sampled primary sampling unit (PSU) and state. Each sampled school had each of its nearest neighbors within the same sampling stratum on the school frame file identified as a potential substitute. Since age-specific enrollment was used as the last sort ordering variable, the nearest neighbors had age-specific enrollment values very close to that of the sampled school. This was done to facilitate the selection of about the same number of students within the substitute as would have been selected from the original sampled school.

Schools were disqualified as potential substitutes if they were already selected in any of the original private school samples or assigned as a substitute for another private school (earlier in the sort ordering). Schools assigned as substitutes for age seventeen schools were disqualified as potential substitutes for age nine and age thirteen schools, and schools assigned as substitutes for age thirteen schools were disqualified as potential substitutes for age nine schools.

If both nearest neighbors were still eligible to be substitutes, the one with a closer age-specific enrollment was chosen. If both nearest neighbors were equally distant from the sampled school in their age-specific enrollment (an uncommon occurrence), one of the two was randomly selected.

Of the 360 original sampled private schools for the long-term trend (LTT) assessment, 107 schools had substitutes activated when the original eligible schools did not participate. Ultimately, 43 of the activated substitute private schools participated.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2012/2012_ltt_samp_priv_subs.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Student Base Weights for the 2012 LTT Assessment

Every sampled student received a student base weight, whether or not the student participated in the assessment. The student base weight is the reciprocal of the probability that the student was sampled to participate in the assessment for a specified subject. The student base weight for student *k* from school *s* in stratum *j (STU_BWT$_{jsk}$)* is the product of seven weighting components and can be expressed as follows:

$$STU\_BWT_{jsk} = SCH\_BWT_{js} \times SCHSESWT_{js} \times WINSCHWT_{js} \times STUSESWT_{jsk} \times SUBJFAC_{jsk} \times SUBADJ_{js} \times YRRND\_AF_{js}$$

where

- *SCH_BWT$_{js}$* is the school base weight;
- *SCHSESWT$_{js}$* is the school-level session assignment weight that reflects the conditional probability, given the school, that the particular session type was assigned to the school;
- *WINSCHWT$_{js}$* is the within-school student weight that reflects the conditional probability, given the school, that the student was selected for the NAEP assessment;
- *STUSESWT$_{jsk}$* is the student-level session assignment weight that reflects the conditional probability, given the particular session type was assigned to the school, that the student was assigned to that session type;
- *SUBJFAC$_{jsk}$* is the subject spiral adjustment factor that reflects the conditional probability, given the student was assigned to a particular session type, that the student was assigned the specified subject;
- *SUBADJ$_{js}$* is the substitution adjustment factor to account for the difference in enrollment size between the substitute and original school; and
- *YRRND_AF$_{js}$* is the year-round adjustment factor to account for students in year-round schools on scheduled break at the time of the NAEP assessment and thus not available for sample.

The within-school student weight (*WINSCHWT$_{js}$*) is the inverse of the student sampling rate in the school.

The subject spiral adjustment factor (*SUBJFAC$_{jsk}$*) adjusts the student weight to account for the spiral pattern used in distributing mathematics or reading booklets to the students. The subject factor varies by sample age, subject, and school type (public/private). It is equal to the inverse of the booklet proportions (mathematics or reading) in the overall spiral for a specific sample.

For cooperating substitutes of nonresponding sampled original schools, the substitution adjustment factor *(SUBADJ$_{js}$)* is equal to the ratio of the estimated age-specific enrollment for the originally sampled school to the estimated age-specific enrollment for the substitute school. The student sample from the substitute school then "represents" the set of age-eligible students from the originally sampled school.

The year-round adjustment factor ($YRRND\_AF_{js}$) adjusts the student weight for students in year-round schools who do not attend school during the time of the assessment. This situation typically arises in overcrowded schools. School administrators in year-round schools randomly assign students to portions of the year in which they attend school and portions of the year in which they do not attend. At the time of assessment, a certain percentage of students (designated as OFF_{js}) do not attend school and thus cannot be assessed. The $YRRND\_AF_{js}$ for a school is calculated as $1/(1-OFF_{js}/100)$.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_base_stud.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation School and Student Nonresponse Weight Adjustments for the 2012 LTT Assessment

School Nonresponse Weight Adjustment

Student Nonresponse Weight Adjustment

Nonresponse is unavoidable in any voluntary survey of a human population. Nonresponse leads to the loss of sample data that must be compensated for in the weights of the responding sample members. This differs from ineligibility, for which no adjustments are necessary. The purpose of the nonresponse adjustments is to reduce the mean square error of survey estimates. While the nonresponse adjustment reduces the bias from the loss of sample, it also increases variability among the survey weights leading to increased variances. However, it is presumed that the reduction in bias more than compensates for the increase in the variance, thereby reducing the mean square error and thus improving the accuracy of survey estimates. Nonresponse adjustments are made in the NAEP surveys at both the school and the student levels: the responding (original and substitute) schools receive a weighting adjustment to compensate for nonresponding schools, and responding students receive a weighting adjustment to compensate for nonresponding students.

The paradigm used for nonresponse adjustment in NAEP is the quasi-randomization approach (Oh and Scheuren 1983). In this approach, school response cells are based on characteristics of schools known to be related to both response propensity and achievement level, such as the locale type (e.g., large principal city of a metropolitan area) of the school. Likewise, student response cells are based on characteristics of the schools containing the students and student characteristics, which are known to be related to both response propensity and achievement level, such as student race/ethnicity, gender, and age.

Under this approach, sample members are assigned to mutually exclusive and exhaustive response cells based on predetermined characteristics. A nonresponse adjustment factor is calculated for each cell as the ratio of the sum of adjusted base weights for all eligible units to the sum of adjusted base weights for all responding units. The nonresponse adjustment factor is then applied to the adjusted base weight of each responding unit. In this way, the weights of responding units in the cell are "weighted up" to represent the full set of responding and nonresponding units in the response cell.

The quasi-randomization paradigm views nonresponse as another stage of sampling. Within each nonresponse cell, the paradigm assumes that the responding sample units are a simple random sample from the total set of all sample units. If this model is valid, then the use of the quasi-randomization weighting adjustment will eliminate any nonresponse bias. Even if this model is not valid, the weighting adjustments will eliminate bias if the achievement scores are homogeneous within the response cells (i.e., bias is eliminated if there is homogeneity either in response propensity or in achievement levels). See, for example, chapter 4 of Little and Rubin (1987).

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_nonresp.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation School Nonresponse Weight Adjustments for the 2012 LTT Assessment

The school nonresponse adjustment procedure inflates the weights of participating schools to account for eligible nonparticipating schools for which no substitute schools participated. The adjustments are computed within nonresponse cells and are based on the assumption that the participating and nonparticipating schools within the same cell are more similar to one another than to schools from different cells. Exactly how nonresponse cells were defined varied for public and private schools.

Development of Initial School Nonresponse Cells

Development of Final School Nonresponse Cells

School Nonresponse Adjustment Factor Calculation

NAEP Technical Documentation Website

NAEP Technical Documentation Development of Initial School Nonresponse Cells for the 2012 LTT Assessment

The cells for nonresponse adjustments are generally functions of the school sampling strata for the individual samples. For NAEP 2012 LTT, school sampling strata were the same for each age and subject sample, but differed by school type (public or private). Assessment subjects that are administered together by way of spiraling have the same school samples and stratification schemes. Subjects that are not spiraled with any other subjects have their own separate school sample. In NAEP 2012 LTT, the mathematics and reading assessments were spiraled together.

The description of the initial nonresponse cells for the NAEP 2012 LTT samples is given below.

Public School Samples

For public school samples, initial weighting cells were formed within each age sample using the following nesting cell structure:

- census region,
- collapsed urbanicity (collapsed urban-centric locale) stratum, and
- race/ethnicity classification.

Private School Samples

For private school samples, initial weighting cells were formed within each age sample using the following nesting cell structure:

- affiliation (Catholic or non-Catholic),
- census region, and
- collapsed urbanicity (collapsed urban-centric locale) stratum.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_nonresp_schl_initial.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Development of Final School Nonresponse Cells for the 2012 LTT Assessment

Limits were placed on the magnitude of cell sizes and adjustment factors to prevent unstable nonresponse adjustments and unacceptably large nonresponse factors. All initial weighting cells with fewer than six cooperating schools or adjustment factors greater than 3.0 for the full sample weight were collapsed with suitable adjacent cells. Simultaneously, all initial weighting cells for any replicate with fewer than four cooperating schools or adjustment factors greater than the maximum of 3.0 (or two times the full sample nonresponse adjustment factor) were collapsed with suitable adjacent cells. Initial weighting cells were generally collapsed in reverse order of the cell structure; that is, starting at the bottom of the nesting structure and working up toward the top level of the nesting structure.

Public School Samples

For the public school samples, race/ethnicity classification cells within a collapsed urbanicity (collapsed urban-centric locale) stratum and census region were collapsed first. If further collapsing was required after all levels of race/ethnicity cells were collapsed, collapsed-urbanicity strata within census region were combined next. Cells were never collapsed across census region.

Private School Samples

For the private school samples, collapsed-urbanicity strata within a census region and affiliation type were collapsed first. If further collapsing was required, census region cells within an affiliation type were collapsed. Cells were never collapsed across affiliation.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_nonresp_schl_final.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation School Nonresponse Adjustment Factor Calculation for the 2012 LTT Assessment

In each final school nonresponse adjustment cell c, the school nonresponse adjustment factor SCH_NRAF_c was computed as follows:

$$SCH\_NRAF_c = \frac{\sum_{s \in S_c} SCH\_BWT_s \times SCH\_TRIM_s \times SCHSESWT_s \times X_s}{\sum_{s \in R_c} SCH\_BWT_s \times SCH\_TRIM_s \times SCHSESWT_s \times X_s}$$

where

- S_c is the set of all eligible sampled schools (cooperating original and substitute schools and refusing original schools with noncooperating or no assigned substitute) in cell c,
- R_c is the set of all cooperating schools within S_c,
- SCH_BWT_s is the school base weight,
- SCH_TRIM_s is the school-level weight trimming factor,
- $SCHSESWT_s$ is the school-level session assignment weight, and
- X_s is the estimated age-specific enrollment corresponding to the original sampled school.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_nonresp_schl_factor.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Website

NAEP Technical DocumentationStudent Nonresponse Adjustment Factor Calculation for the 2012 LTT Assessment

In each final student nonresponse adjustment cell c for a given sample, the student nonresponse adjustment factor STU_NRAF_c was computed as follows:

$$STU\_NRAF_c = \frac{\sum_{k \in S_c} STU\_BWT_k \times SCH\_TRIM_k \times SCH\_NRAF_k / SUBJFAC_k}{\sum_{k \in R_c} STU\_BWT_k \times SCH\_TRIM_k \times SCH\_NRAF_k / SUBJFAC_k}$$

where

- S_c is the set of all eligible sampled students in cell c for a given sample,
- R_c is the set of all assessed students within S_c,
- STU_BWT_k is the student base weight for a given student k,
- SCH_TRIM_k is the school-level weight trimming factor for the school associated with student k,
- SCH_NRAF_k is the school-level nonresponse adjustment factor for the school associated with student k, and
- $SUBJFAC_k$ is the subject factor for a given student k.

The student weight used in the calculation above is the adjusted student base weight, without regard to subject, adjusted for school weight trimming and school nonresponse.

Nonresponse adjustment procedures are not applied to excluded students because they are not required to complete an assessment. In effect, excluded students were placed in a separate nonresponse cell by themselves and all received an adjustment factor of 1. While excluded students are not included in the analysis of the NAEP scores, weights are provided for excluded students in order to estimate the size of this group and its population characteristics.

NAEP Technical Documentation School and Student Weight Trimming Adjustments for the 2012 LTT Assessment

Weight trimming is an adjustment procedure that involves detecting and reducing extremely large weights. "Extremely large weights" generally refer to large sampling weights that were not anticipated in the design of the sample. Unusually large weights are likely to produce large sampling variances for statistics of interest, especially when the large weights are associated with sample cases reflective of rare or atypical characteristics. To reduce the impact of these large weights on variances, weight reduction methods are typically employed. The goal of weight reduction methods is to reduce the mean square error of survey estimates. While the trimming of large weights reduces variances, it also introduces some bias. However, it is presumed that the reduction in the variances more than compensates for the increase in the bias, thereby reducing the mean square error and thus improving the accuracy of survey estimates (Potter 1988). NAEP employs weight trimming at both the school and student levels.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_trimming_adjustments.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Trimming of School Base Weights for the 2012 LTT Assessment

Large school weights can occur for schools selected from the NAEP new-school sampling frame and for private schools. New schools that are eligible for weight trimming are schools with a disproportionately large student enrollment in a particular grade from a school district that was selected with a low probability of selection. The school base weights for such schools may be large relative to what they would have been if they had been selected as part of the original sample.

To detect extremely large weights among new schools, a comparison was made between a new school's school base weight and its ideal weight (i.e., the weight that would have resulted had the

school been selected from the original school sampling frame). If the school base weight was more than three times the ideal weight, a trimming factor was calculated for that school that scaled the base weight back to three times the ideal weight. The calculation of the school-level trimming factor for a new school s is expressed in the following formula:

$$SCH\_TRIM_s = \begin{cases} \dfrac{3 \times EXP\_WT_s}{SCH\_BWT_s}, & \text{if } \dfrac{SCH\_BWT_s}{EXP\_WT_s} > 3 \\ 1, & \text{otherwise} \end{cases}$$

where

- EXP_WT_s is the ideal base weight the school would have received if it had been on the NAEP public school sampling frame, and
- SCH_BWT_s is the actual school base weight the school received as a sampled school from the new school frame.

No new schools in any of the NAEP 2012 LLT samples had their weights trimmed.

Private schools eligible for weight trimming were Private School Universe Survey (PSS) nonrespondents who were found subsequently to have either larger enrollments than assumed at the time of sampling, or an atypical probability of selection given their affiliation, the latter being unknown at the time of sampling. For private school *s*, the formula for computing the school-level weight trimming factor SCH_TRIM_s is identical to that used for new schools. For private schools,

- EXP_WT_s is the ideal base weight the school would have received if it had been on the NAEP private school sampling frame with accurate enrollment and known affiliation, and
- SCH_BWT_s is the actual school base weight the school received as a sampled private school.

No private schools in any of the NAEP 2012 LTT samples had their weights trimmed.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_base_schtrim.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Trimming of Student Weights for the 2012 LTT Assessment

Large student weights generally come from compounding nonresponse adjustments at the school and student levels with artificially low first-stage selection probabilities, which can result from inaccurate enrollment data on the school frame used to define the school size measure. Even though measures are in place to limit the number and size of excessively large weights—such as the implementation of adjustment factor size constraints in both the school and student nonresponse procedures and the use of the school trimming procedure—large student weights can still occur.

The student weight trimming procedure uses a multiple median rule to detect excessively large student weights. Any student weight within a given trimming group greater than a specified multiple of the median weight value of the given trimming group has its weight scaled back to that threshold. Trimming groups were defined by age, subject, region, and Black/Hispanic strata (age 17 only) for public schools, and affiliation (Catholic/non-Catholic) for private schools.

The procedure computes the median of the nonresponse-adjusted student weights in the trimming group g for a given grade and subject sample. Any student k with a weight more than M times the median (where $M = 3.5$ for public and private schools) received a trimming factor calculated as follows:

$$STU\_TRIM_{gk} = \begin{cases} \dfrac{M \times MEDIAN_g}{STUWGT_{gk}}, & \text{if } STUWGT_{gk} > M \times MEDIAN_g \\ 1, & \text{otherwise} \end{cases}$$

where

- M is the trimming multiple,
- $MEDIAN_g$ is the median of nonresponse-adjusted student weights in trimming group g,and
- $STUWGT_{gk}$ is the weight after student nonresponse adjustment for student k in trimming group g.

In the NAEP 2012 LTT assessments, relatively few students had weights considered excessively large. Out of the approximately 53,500 students included in the combined 2012 LTT assessment samples, only 22 students had their weights trimmed.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_studtrim.aspx

NAEP Technical Documentation Website

NAEP Technical Documentation Computation of Replicate Weights for Variance Estimation for the 2012 LTT Assessment

In addition to the full-sample weight, a set of 62 replicate weights was provided for each student. These replicate weights are used in calculating the sampling variance of estimates obtained from the data, using the jackknife repeated replication method. The method of deriving these weights was aimed at reflecting the features of the sample design appropriately for each sample, so that when the jackknife variance estimation procedure is implemented, approximate unbiased estimates of sampling variance are obtained. This section gives the specifics for generating the replicate weights for the 2012 LTT assessment samples. The theory that underlies the jackknife variance estimators used in NAEP studies is discussed in the section Replicate Variance Estimation.

For each sample, replicates were formed in two steps. First, each school was assigned to one or more of 62 replicate strata. In the next step, a random subset of schools (or, in some cases, students within schools) in each replicate stratum was excluded. The remaining subset and all schools in the other replicate strata then constituted one of the 62 replicates.

A replicate weight was calculated for each of the 62 replicates using weighting procedures similar to those used for the full-sample weight. Each replicate base weight contains an additional component, known as a replicate factor, to account for the subsetting of the sample to form the replicate. By repeating the various

weighting procedures on each set of replicate base weights, the impact of these procedures on the sampling variance of an estimate is appropriately reflected in the variance estimate.

Each of the 62 replicate weights for student *k* in school *s* and stratum *j* can be expressed as follows:

$$FSTUWGT_{jsk}(r) = STU\_BWT_{jsk}(r) \times SCH\_NRAF_{js}(r) \times STU\_NRAF_{jsk}(r) \times SCH\_TRIM_{js} \times STU\_TRIM_{jsk}$$

where

- $STU\_BWT_{jsk}(r)$ is the student base weight for replicate *r*;
- $SCH\_NRAF_{js}(r)$ is the school-level nonresponse adjustment factor for replicate *r*;
- $STU\_NRAF_{jsk}(r)$ is the student-level nonresponse adjustment factor for replicate *r*;
- SCH_TRIM_{js} is the school-level weight trimming adjustment factor; and
- STU_TRIM_{jsk} is the student-level weight trimming adjustment factor.

Specific school and student nonresponse adjustment factors were calculated separately for each replicate, thus the use of the index (*r*), and applied to the replicate student base weights. Computing separate nonresponse adjustment factors for each replicate allows resulting variances from the use of the final student replicate weights to reflect components of variance due to these various weight adjustments.

School and student weight trimming adjustments were not replicated, that is, not calculated separately for each replicate. Instead, each replicate used the school and student trimming adjustment factors derived for the full sample. Statistical theory for replicating trimming adjustments under the jackknife approach has not been developed in the literature. Due to the absence of a statistical framework, and since relatively few school and student weights in NAEP require trimming, the weight trimming adjustments were not replicated.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_rep_var_est.aspx

NAEP Technical Documentation Website

NAEP Technical DocumentatioDefining Replicate Strata and Forming Replicates for the 2012 LTT Assessment

In the NAEP 2012 LTT assessment, replicates were formed separately for each sample indicated by age (9, 13, 17), and school type (public, private). The first step in forming replicates was to assign each first-stage sampling unit in a primary stratum to a replicate stratum. In 2012, the formation of replicate strata varied by noncertainty and certainty primary sampling units (PSUs). For noncertainty PSUs, the first-stage units were PSUs, and the primary stratum was the combination of region and metropolitan status (MSA or non-MSA). For certainty PSUs, the first-stage units were schools, and the primary stratum was school type (public or private).

For noncertainty PSUs, where only one PSU was selected per PSU stratum, replicate strata were formed by pairing sampled PSUs with similar stratum characteristics within the same primary stratum (region by metropolitan status). This was accomplished by first sorting the 38 sampled PSUs by PSU stratum number and then grouping adjacent PSUs into 19 pairs. The values for a PSU stratum number reflect region and metropolitan status, as well as socioeconomic characteristics such as percent Black and percent children below poverty (those eligible for free/reduced-price school lunch). The formation of these 19 replicate strata in this manner models a design of selecting two PSUs with probability proportional to size with replacement from each of 19 strata.

For certainty PSUs, the first stage of sampling is at the school level, and the formation of replicate strata must reflect the sampling of schools within the certainty PSUs. Replicate strata were formed by sorting the sampled schools in the 29 certainty PSUs by their order of selection within a primary stratum (school type) so that the sort order reflected the implicit stratification (region, locality type, race/ethnicity classification, and student enrollment for public schools; and region, private school type, and student enrollment size for private schools) and systematic sampling features of the sample design.

The first-stage units were then paired off into 43 preliminary replicate strata. Within each primary stratum with an even number of first-stage units, all of the preliminary replicate strata were pairs, and within primary strata with an odd number of first-stage units, one of the replicate strata was a triplet (the last one), and all others were pairs.

If there were more than 43 preliminary replicate strata within a primary stratum, the preliminary replicate strata were grouped to form 43 replicate strata. This grouping effectively maximized the distance in the sort order between grouped preliminary replicate strata. The first 43 preliminary replicate strata, for example, were assigned to 43 different final replicate strata in order (1 through 43), with the next 43 preliminary replicate strata assigned to final replicate strata 1 through 43, so that, for example, preliminary replicate stratum 1, preliminary replicate stratum

44, preliminary replicate stratum 87 (if there were that many), etc., were all assigned to the first final replicate stratum. The final replicate strata for the schools in the certainty PSUs were 1 through 43.

Within each pair of preliminary replicate stratum, the first first-stage unit was assigned as the first variance unit and the second first-stage unit as the second variance unit. Within each triplet preliminary replicate stratum, the three schools were assigned variance units 1 through 3.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_repwts_strata.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Defining Replicate Strata and Forming Replicates for the 2012 LTT Assessment

In the NAEP 2012 LTT assessment, replicates were formed separately for each sample indicated by age (9, 13, 17), and school type (public, private). The first step in forming replicates was to assign each first-stage sampling unit in a primary stratum to a replicate stratum. In 2012, the formation of replicate strata varied by noncertainty and certainty primary sampling units (PSUs). For noncertainty PSUs, the first-stage units were PSUs, and the primary stratum was the combination of region and metropolitan status (MSA or non-MSA). For certainty PSUs, the first-stage units were schools, and the primary stratum was school type (public or private).

For noncertainty PSUs, where only one PSU was selected per PSU stratum, replicate strata were formed by pairing sampled PSUs with similar stratum characteristics within the same primary stratum (region by metropolitan status). This was accomplished by first sorting the 38 sampled PSUs by PSU stratum number and then grouping adjacent PSUs into 19 pairs. The values for a PSU stratum number reflect region and metropolitan status, as well as socioeconomic characteristics such as percent Black and percent children below poverty (those eligible for free/reduced-price school lunch). The formation of these 19 replicate strata in this manner models a design of selecting two PSUs with probability proportional to size with replacement from each of 19 strata.

For certainty PSUs, the first stage of sampling is at the school level, and the formation of replicate strata must reflect the sampling of schools within the certainty PSUs. Replicate strata were formed by sorting the sampled schools in the 29 certainty PSUs by their order of selection within a primary stratum (school type) so that the sort order reflected the

implicit stratification (region, locality type, race/ethnicity classification, and student enrollment for public schools; and region, private school type, and student enrollment size for private schools) and systematic sampling features of the sample design.

The first-stage units were then paired off into 43 preliminary replicate strata. Within each primary stratum with an even number of first-stage units, all of the preliminary replicate strata were pairs, and within primary strata with an odd number of first-stage units, one of the replicate strata was a triplet (the last one), and all others were pairs.

If there were more than 43 preliminary replicate strata within a primary stratum, the preliminary replicate strata were grouped to form 43 replicate strata. This grouping effectively maximized the distance in the sort order between grouped preliminary replicate strata. The first 43 preliminary replicate strata, for example, were assigned to 43 different final replicate strata in order (1 through 43), with the next 43 preliminary replicate strata assigned to final replicate strata 1 through 43, so that, for example, preliminary replicate stratum 1, preliminary replicate stratum 44, preliminary replicate stratum 87 (if there were that many), etc., were all assigned to the first final replicate stratum. The final replicate strata for the schools in the certainty PSUs were 1 through 43.

Within each pair of preliminary replicate stratum, the first first-stage unit was assigned as the first variance unit and the second first-stage unit as the second variance unit. Within each triplet preliminary replicate stratum, the three schools were assigned variance units 1 through 3.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_repwts_strata.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Computing School-Level Replicate Base Weights for the 2012 LTT Assessment

For the NAEP 2012 LTT assessment, school-level replicate base weights for school s in primary stratum j ($SCH\_BWT_{js}(r)$, $r = 1,..., 62$) were calculated as follows:

$$SCH\_BWT_{js}(r) = \begin{cases} 2 \times SCH\_BWT_{js}, & \text{for } js \in R_{jr}, \quad U_{js} = 1 \\ 0, & \text{for } js \in R_{jr}, \quad U_{js} = 2 \\ SCH\_BWT_{js}, & js \notin R_{jr} \end{cases}$$

where

- SCH_BWT_{js} is the school base weight for school s in primary stratum j,
- R_{jr} is the set of schools within the r-th replicate stratum for primary stratum j, and
- U_{js} is the variance unit (1 or 2) for school s in primary stratum j.

For schools in replicate strata comprising three variance units, two sets of school-level replicate base weights were computed (see replicate variance estimation for details): one for the first replicate r_1 and another for the second replicate r_2. The two sets of school-level replicate base weights $SCH\_BWT_{js}(r_1)$, $r_1 = 1,..., 62$ and $SCH\_BWT_{js}(r_2)$, $r_2 = 1,..., 62$ were calculated as described below.

$$SCH\_BWT_{js}(r_1) = \begin{cases} 1.5 \times SCH\_BWT_{js}, & \text{for } js \in R_{jr_1}, U_{js} = 1 \\ 1.5 \times SCH\_BWT_{js}, & \text{for } js \in R_{jr_1}, U_{js} = 2 \\ 0, & \text{for } js \in R_{jr_1}, U_{js} = 3 \\ SCH\_BWT_{js}, & js \notin R_{jr_1} \end{cases}$$

$$SCH\_BWT_{js}(r_2) = \begin{cases} 1.5 \times SCH\_BWT_{js}, & \text{for } js \in R_{jr_2}, U_{js} = 1 \\ 0, & \text{for } js \in R_{jr_2}, U_{js} = 2 \\ 1.5 \times SCH\_BWT_{js}, & \text{for } js \in R_{jr_2}, U_{js} = 3 \\ SCH\_BWT_{js}, & js \notin R_{jr_2} \end{cases}$$

where

- SCH_BWT_{js} is the school base weight for school s in primary stratum j,
- R_{jr1} is the set of schools within the r_1-th replicate stratum for primary stratum j,
- R_{jr2} is the set of schools within the r_2-th replicate stratum for primary stratum j, and
- U_{js} is the variance unit (1, 2, or 3) for school s in primary stratum j.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_repwts_schl.aspx

NAEP Technical Documentation Website

NAEP Technical Documentation Computing Student-Level Replicate Base Weights for the 2012 LTT Assessment

For the 2012 LTT assessment, the calculation of the student-level replicate base weights for student *k* from school *s* in stratum *j* for each of the 62 replicates, $STU\_BWT_{jsk}(r)$, where $r = 1$ to 62, were calculated as follows:

$$STU\_BWT_{jsk}(r) = SCH\_BWT_{js}(r) \times SCHSESWT_{js} \times WINSCHWT_{js} \times STUSESWT_{jsk} \times SUBJFAC_{js} \times SUBADJ_{js} \times YRRND\_AF_{js}$$

where

- $SCH\_BWT_{js}(r)$ is the replicate school base weight;
- $SCHSESWT_{js}$ is the school-level session assignment weight used in the full-sample weight;
- $WINSCHWT_{js}$ is the within-school student sampling weight used in the full-sample weight;
- $STUSESWT_{jsk}$ is the student-level session assignment weight used in the full-sample weight;
- $SUBJFAC_{js}$ is the subject factor used in the full-sample weight;
- $SUBADJ_{js}$ is the substitute adjustment factor used in the full-sample weight; and
- $YRRND\_AF_{js}$ is the year-round adjustment factor used in the full-sample weight.

These components are described on the Student Base Weights page.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_repwts_stud.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Replicate Variance Estimation for the 2012 Assessment

Variances for NAEP assessment estimates are computed using the paired jackknife replicate variance procedure. This technique is applicable for common statistics, such as means and ratios, as well as for more complex statistics such as Item Response Theory (IRT) scores.

In general, the paired jackknife replicate variance procedure involves pairing clusters of first-stage sampling units to form H variance strata (h = 1, 2, 3, ...,H) with two units per stratum. The first replicate is formed by deleting one unit at random from the first variance stratum, inflating the weight of the remaining unit to weight up to the variance stratum total, and using all other units from the other (H - 1) strata. This procedure is carried out for each variance stratum resulting in H replicates, each of which provides an estimate of the population total.

The jackknife estimate of the variance for any given statistic is given by the following formula:

$$v\left(\hat{t}\right) = \sum_{h=1}^{H} \left(\hat{t}_h - \hat{t}\right)^2$$

where

- $\hat{t}$ represents the full sample estimate of the given statistic, and
- $\hat{t}_h$ represents the corresponding estimate for replicate *h*.

Each replicate undergoes the same weighting procedure as the full sample so that the jackknife variance estimator reflects the contributions to or reductions in variance resulting from the various weighting adjustments.

The NAEP jackknife variance estimator is based on 62 variance strata resulting in a set of 62 replicate weights assigned to each school and student.

The basic idea of the paired jackknife variance estimator is to create the replicate weights so that use of the jackknife procedure results in an unbiased variance estimator for simple totals and means, which is also reasonably efficient (i.e., has a low variance as a variance estimator). The jackknife variance estimator will then produce a consistent (but not fully unbiased) estimate of variance for (sufficiently smooth) nonlinear functions of total and mean estimates such as ratios, regression coefficients, and so forth (Shao and Tu, 1995).

The development below shows why the NAEP jackknife variance estimator returns an unbiased variance estimator for totals and means, which is the cornerstone to the asymptotic results for nonlinear estimators. See for example Rust (1985). This paper also discusses why this variance estimator is generally efficient (i.e., more reliable than alternative approaches requiring similar computational resources).

The development is done for an estimate of a mean based on a simplified sample design that closely approximates the sample design for first-stage units used in the NAEP studies. The sample design is a stratified random sample with *H* strata with population weights W_h, stratum

sample sizes n_h, and stratum sample means $\bar{y}_h$. The population estimator $\hat{\bar{Y}}$ and standard unbiased variance estimator $v\left(\hat{\bar{Y}}\right)$ are:

$$\hat{\bar{Y}} = \sum_{h=1}^{H} W_h \bar{y}_h \qquad v\left(\hat{\bar{Y}}\right) = \sum_{h=1}^{H} W_h^2 \frac{s_h^2}{n_h}$$

with

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} \left(y_{h_i} - \bar{y}_h \right)^2$$

The paired jackknife replicate variance estimator assigns one replicate h=1,…, H to each stratum, so that the number of replicates equals H. In NAEP, the replicates correspond generally to pairs and triplets (with the latter only being used if there are an odd number of sample units within a particular primary stratum generating replicate strata). For pairs, the process of generating replicates can be viewed as taking a simple random sample (J) of size $n_h/2$ within the replicate stratum, and assigning an increased weight to the sampled elements, and a decreased weight to the unsampled elements. In certain applications, the increased weight is double the full sample weight, while the decreased weight is in fact equal to zero. In this simplified case, this assignment reduces to replacing $\bar{y}_h$ with $\bar{y}_h(J)$, the latter being the sample mean of the sampled $n_h/2$ units. Then the replicate estimator corresponding to stratum $r$is

$$\hat{\bar{Y}}(r) = \sum_{h \neq r}^{H} W_h \bar{y}_h + W_r \bar{y}_r (J)$$

The r-th term in the sum of squares for $v_j\left(\hat{\bar{Y}}\right)$ is thus:

$$\left(\hat{\bar{Y}}(r) - \hat{\bar{Y}} \right)^2 = W_r^2 \left(\bar{y}_r (J) - \bar{y}_r \right)^2$$

In stratified random sampling, when a sample of size $n_r/2$ is drawn without replacement from a population of size n_r,, the sampling variance is

$$E\left(\bar{y}_r (J) - \bar{y}_r \right)^2 = \frac{1}{(n_r/2)} \frac{n_r - n_r/2}{n_r} \frac{1}{n_r - 1} \sum_{i=1}^{n_r} \left(y_{r_i} - \bar{y}_r \right)^2 =$$

$$= \frac{1}{n_r (n_r - 1)} \sum_{i=1}^{n_r} \left(y_{r_i} - \bar{y}_r \right)^2 = \frac{s_r^2}{n_r}$$

See for example Cochran (1977), Theorem 5.3, using n_r, as the "population size," $n_r/2$ as the "sample size," and s_r^2 as the "population variance" in the given formula. Thus,

$$E\left\{W_r^2\left(\bar{y}_r(J)-\bar{y}_r\right)^2\right\} = \quad W_r^2\frac{s_r^2}{n_r}$$

Taking the expectation over all of these stratified samples of size $n_r/2$, it is found that

$$E\left(v_j(\hat{\bar{Y}})\right) = v(\hat{\bar{Y}})$$

In this sense, the jackknife variance estimator "gives back" the sample variance estimator for means and totals as desired under the theory.

In cases where, rather than doubling the weight of one half of one variance stratum and assigning a zero weight to the other, the weight of one unit is multiplied by a replicate factor of (1+δ), while the other is multiplied by (1- δ), the result is that

$$E\left(\hat{\bar{y}}(r)-\hat{\bar{y}}\right)^2 = \quad W_r^2\delta^2\frac{s_r^2}{n_r}$$

In this way, by setting δ equal to the square root of the finite population correction factor, the jackknife variance estimator is able to incorporate a finite population correction factor into the variance estimator.

In practice, variance strata are also grouped to make sure that the number of replicates is not too large (the total number of variance strata is usually 62 for NAEP). The randomization from the original sample distribution guarantees that the sum of squares contributed by each replicate will be close to the target expected value.

For triples, the replicate factors are perturbed to something other than 1.0 for two different replicate factors, rather than just one as in the case of pairs. Again in the simple case where replicate factors that are less than 1 are all set to 0, with the replicate weight factors calculated as follows.

For unit *i* in variance stratum *r*

$$w_i(r) = \begin{cases} 1.5w_i & i = \text{ variance unit 1} \\ 1.5w_i & i = \text{ variance unit 2} \\ 0 & i = \text{ variance unit 3} \end{cases}$$

where weight w_i is the full sample base weight.

Furthermore, for $r' = r + 31 \pmod{62}$:

$$w_i(r') = \begin{cases} 1.5w_i & i = \text{ variance unit 1} \\ 0 & i = \text{ variance unit 2} \\ 1.5w_i & i = \text{ variance unit 3} \end{cases}$$

And for all other values r^*, other than r and r', $w_i(r^*) = 1$.

In the case of stratified random sampling, this formula reduces to replacing $\bar{y}_r$ with $\bar{y}_r(J)$ for replicate r and with $\bar{y}_{r'}(J)$ for replicate r'. $\bar{y}_r(J)$ is the sample mean from a "2/3" sample of $2n_r/3$ units from the n_r sample units in the replicate stratum, and $\bar{y}_{r'}(J)$ is the sample mean from another overlapping "2/3" sample of $2n_r/3$ units from the n_r sample units in the replicate stratum.

The r-th and r'-th replicates can be written as:

$$\hat{\bar{Y}}(r) = \sum_{h \neq r}^{H} W_h \bar{y}_h + W_r \bar{y}_r(J)$$

$$\hat{\bar{Y}}(r') = \sum_{h \neq r}^{H} W_h \bar{y}_h + W_r \bar{y}_{r'}(J)$$

From these formulas, expressions for the r-th and r'-th components of the jackknife variance estimator are obtained (ignoring other sums of squares from other grouped components attached to those replicates):

$$\left(\hat{\bar{Y}}(r) - \hat{\bar{Y}}\right)^2 = W_r^2 \left(\bar{y}_r(J) - \bar{y}_r\right)^2$$

$$\left(\hat{\bar{Y}}(r') - \hat{\bar{Y}}\right)^2 = W_r^2 \left(\bar{y}_{r'}(J) - \bar{y}_r\right)^2$$

These sums of squares have expectations as follows, using the general formula for sampling variances:

$$E\left(\bar{y}_r(J) - \bar{y}_r\right)^2 = \frac{1}{(2n_r/3)} \frac{n_r - (2n_r/3)}{n_r} \frac{1}{n_r - 1} \sum_{i=1}^{n_r} \left(y_{r_i} - \bar{y}_r\right)^2$$

$$= \frac{1}{2n_r(n_r - 1)} \sum_{i=1}^{n_r} \left(y_{r_i} - \bar{y}_r\right)^2 = \frac{s_r^2}{2n_r}$$

$$E\left(\bar{y}_{r'}(J)-\bar{y}_r\right)^2 = \frac{1}{(2n_r/3)}\frac{n_r-(2n_r/3)}{n_r}\frac{1}{n_r-1}\sum_{i=1}^{n_r}\left(y_{ri}-\bar{y}_r\right)^2$$

$$= \frac{1}{2n_r(n_r-1)}\sum_{i=1}^{n_r}\left(y_{ri}-\bar{y}_r\right)^2 = \frac{s_r^2}{2n_r}$$

Thus,

$$E\left\{W_r^2\left(\bar{y}_r(J)-\bar{y}_r\right)^2+W_r^2\left(\bar{y}_{r'}(J)-\bar{y}_r\right)^2\right\} = W_r^2\left(\frac{s_r^2}{2n_r}+\frac{s_r^2}{2n_r}\right) = W_r^2\frac{s_r^2}{n_r}$$

as desired again.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_var_est_appdx.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Quality Control on Weighting Procedures for the 2012 LTT Assessment

Given the complexity of the weighting procedures utilized in NAEP, a range of quality control (QC) checks was conducted throughout the weighting process to identify potential problems with collected student-level demographic data or with specific weighting procedures. The QC processes included

- checks performed within each step of the weighting process;
- checks performed across adjacent steps of the weighting process;
- review of response, exclusion, and accommodation rates;
- checking demographic data of individual schools;
- comparisons with 2008 demographic data; and
- nonresponse bias analyses.

Main QC Findings of Interest

Participation, Exclusion, and Accommodation Rates

Nonresponse Bias Analysis

To validate the weighting process, extensive tabulations of various school and student characteristics at different stages of the process were conducted. The school-level characteristics included in the tabulations were enrollment by race/ethnicity and urban-centric locale. At the student level, the tabulations included race/ethnicity, gender, categorized grade, students with disability (SD) status, English language learners (ELL) status, and participation status in National School Lunch Program (NSLP).

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_qc_procedures.aspx

---

NAEP Technical Documentation Website

NAEP Technical Documentation Participation, Exclusion and Accommodation Rates for the 2012 LTT Assessment

Final participation, exclusion, and accommodation rates were presented in quality control tables for each age and subject by reporting group. School-level participation rates were calculated as they had been calculated for previous assessments and according to National Center for Education Statistics (NCES) standards.

Age 9 Mathematics
Age 9 Reading

Age 13 Mathematics
Age 13 Reading

Age 17 Mathematics
Age 17 Reading

School-level participation rates were below 85 percent for private schools at all three ages. Student-level participation rates were all above 85 percent. As required by NCES standards, nonresponse bias analyses were conducted on each reporting group falling below the 85 percent participation threshold.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_ltt_part_exclusion_acc_rates.aspx

---

Participation, Exclusion, and Accommodation Rates for Age 9 Mathematics for the 2012 LTT Assessment

The following table displays the school-level participation rates and student-level participation, exclusion, and accommodation rates for the age 9 long-term trend (LTT) mathematics assessment. Various weights were used in the calculation of the rates, as indicated in the column headings of the table.

The participation rates reflect the participation of the original sampled schools only and do not reflect any effect of substitution. The rates weighted by the school base weight and enrollment show the approximate proportion of the student population in the domain that is represented by the responding schools in the sample. The rates weighted by just the base weight show the proportion of the school population that is represented by the responding schools in the sample. These rates differ because schools differ in size.

Participation, exclusion, and accommodation rates for age 17 long-term trend mathematics assessment, by geographic region and school type: 2012

Geographic region and school type	Number of schools in original sample	School participation rates (percent) before substitution (weighted by school base weight and enrollment)	School participation rates (percent) before substitution (weighted by school base weight only)	Number of students sampled	Weighted percent of students excluded	Weighted student participation rates (percent) after makeups	Weighted percent of students accommodated
National all	**482**	**83.82**	**80.26**	**10,900**	**1.74**	**88.06**	**9.57**
Northeast all	81	92.27	74.44	2,000	2.55	85.59	13.29
Midwest all	97	90.74	90.45	2,100	1.46	88.15	10.59
South all	184	82.17	78.53	4,100	1.49	89.96	8.06
West all	120	72.76	75.82	2,600	1.72	87.18	7.91
National public	389	85.58	87.57	10,000	1.86	88.22	9.61
National private	93	62.51	60.45	833	0.13	85.87	9.11
Catholic	16	88.18	86.99	378	0.25	86.80	5.97
Non-Catholic	77	40.30	50.18	455	0.00	84.42	12.30

NOTE: National all includes national public, national private, Bureau of Indian Education (BIE), and Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) that are located in the United States. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2012 Mathematics Long-Term Trend Assessment.

Participation, Exclusion, and Accommodation Rates for Age 9 Reading for the 2012 LTT Assessment

The following table displays the school-level participation rates and student-level participation, exclusion, and accommodation rates for the age 9 long-term trend (LTT) reading assessment. Various weights were used in the calculation of the rates, as indicated in the column headings of the table.

The participation rates reflect the participation of the original sampled schools only and do not reflect any effect of substitution. The rates weighted by the school base weight and enrollment show the approximate proportion of the student population in the domain that is represented by the responding schools in the sample. The rates weighted by just the base weight show the proportion of the school population that is represented by the responding schools in the sample. These rates differ because schools differ in size.

Participation, exclusion, and accommodation rates for age 9 long-term trend reading assessment, by geographic region and school type: 2012

Geographic region and school type	Number of schools in original sample	School participation rates (percent) before substitution (weighted by school base weight and enrollment)	School participation rates (percent) before substitution (weighted by school base weight only)	Number of students sampled	Weighted percent of students excluded	Weighted student participation rates (percent) after makeups	Weighted percent of students accommodated
National all	**484**	**86.64**	**81.54**	**9,800**	**1.68**	**94.94**	**10.46**
Northeast all	83	93.39	77.87	1,500	1.54	94.55	13.30
Midwest all	100	90.82	86.94	1,800	1.50	95.10	12.64
South all	186	84.18	76.81	4,200	2.31	94.99	10.36
West all	115	82.22	84.85	2,300	0.96	95.00	6.71
National public	347	89.03	89.93	8,900	1.79	95.03	11.15
National private	137	61.16	58.60	918	0.44	93.80	2.18
Catholic	32	95.06	92.80	392	0.00	97.52	2.04
Non-Catholic	105	37.71	44.77	526	0.77	89.86	2.29

NOTE: National all includes national public, national private, Bureau of Indian Education (BIE), and Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) that are located in the United States. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2012 Reading Long-Term Trend Assessment.

Participation, Exclusion, and Accommodation Rates for Age 13 Mathematics for the 2012 LTT Assessment

The following table displays the school-level participation rates and student-level participation, exclusion, and accommodation rates for the age 13 long-term trend (LTT) mathematics assessment. Various weights were used in the calculation of the rates, as indicated in the column headings of the table.

The participation rates reflect the participation of the original sampled schools only and do not reflect any effect of substitution. The rates weighted by the school base weight and enrollment show the approximate proportion of the student population in the domain that is represented by the responding schools in the sample. The rates weighted by just the base weight show the proportion of the school population that is represented by the responding schools in the sample. These rates differ because schools differ in size.

Participation, exclusion, and accommodation rates for age 13 long-term trend mathematics assessment, by geographic region and school type: 2012

Geographic region and school type	Number of schools in original sample	School participation rates (percent) before substitution (weighted by school base weight and enrollment)	School participation rates (percent) before substitution (weighted by school base weight only)	Number of students sampled	Weighted percent of students excluded	Weighted student participation rates (percent) after makeups	Weighted percent of students accommodated
National all	**505**	**87.87**	**80.75**	**10,000**	**1.17**	**93.03**	**10.61**
Northeast all	85	94.87	66.98	1,600	0.61	91.14	14.78
Midwest all	106	90.38	91.73	1,900	1.12	94.70	10.96
South all	189	87.69	78.36	4,100	1.56	92.26	10.07
West all	125	81.27	80.68	2,400	1.00	93.90	8.21
National public	375	89.94	89.99	9,000	1.27	92.85	11.04
National private	130	68.63	62.72	995	0.16	95.10	6.03
Catholic	37	91.61	91.70	489	0.34	95.43	3.22
Non-Catholic	93	49.13	50.95	506	0.00	94.70	8.49

NOTE: National all includes national public, national private, Bureau of Indian Education (BIE), and Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) that are located in the United States. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2012 Mathematics Long-Term Trend Assessment.

Participation, Exclusion, and Accommodation Rates for Age 13 Reading for the 2012 LTT Assessment

The following table displays the school-level participation rates and student-level participation, exclusion, and accommodation rates for the age 13 long-term trend (LTT) reading assessment. Various weights were used in the calculation of the rates, as indicated in the column headings of the table.

The participation rates reflect the participation of the original sampled schools only and do not reflect any effect of substitution. The rates weighted by the school base weight and enrollment show the approximate proportion of the student population in the domain that is represented by the responding schools in the sample. The rates weighted by just the base weight show the proportion of the school population that is represented by the responding schools in the sample. These rates differ because schools differ in size.

Participation, exclusion, and accommodation rates for age 13 long-term trend reading assessment, by geographic region and school type: 2012

Geographic region and school type	Number of schools in original sample	School participation rates (percent) before substitution (weighted by school base weight and enrollment)	School participation rates (percent) before substitution (weighted by school base weight only)	Number of students sampled	Weighted percent of students excluded	Weighted student participation rates (percent) after makeups	Weighted percent of students accommodated
National all	**505**	**87.87**	**80.75**	**10,000**	**1.89**	**93.19**	**10.14**
Northeast all	85	94.87	66.98	1,600	1.60	92.23	14.57
Midwest all	106	90.38	91.73	1,900	1.43	94.97	11.48
South all	189	87.69	78.36	4,200	2.42	92.45	8.84
West all	125	81.27	80.68	2,400	1.74	93.21	7.71
National public	375	89.94	89.99	9,000	2.03	93.13	10.69
National private	130	68.63	62.72	986	0.38	93.94	4.16
Catholic	37	91.61	91.70	484	0.21	96.42	2.01
Non-Catholic	93	49.13	50.95	502	0.53	91.05	6.16

NOTE: National all includes national public, national private, Bureau of Indian Education (BIE), and Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) that are located in the United States. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2012 Reading Long-Term Trend Assessment.

Participation, Exclusion, and Accommodation Rates for Age 17 Mathematics for the 2012 LTT Assessment

The following table displays the school-level participation rates and student-level participation, exclusion, and accommodation rates for the age 17 long-term trend (LTT) mathematics assessment. Various weights were used in the calculation of the rates, as indicated in the column headings of the table.

The participation rates reflect the participation of the original sampled schools only and do not reflect any effect of substitution. The rates weighted by the school base weight and enrollment show the approximate proportion of the student population in the domain that is represented by the responding schools in the sample. The rates weighted by just the base weight show the proportion of the school population that is represented by the responding schools in the sample. These rates differ because schools differ in size.

Participation, exclusion, and accommodation rates for age 13 long-term trend reading assessment, by geographic region and school type: 2012

Geographic region and school type	Number of schools in original sample	School participation rates (percent) before substitution (weighted by school base weight and enrollment)	School participation rates (percent) before substitution (weighted by school base weight only)	Number of students sampled	Weighted percent of students excluded	Weighted student participation rates (percent) after makeups	Weighted percent of students accommodated
National all	**505**	**87.87**	**80.75**	**10,000**	**1.89**	**93.19**	**10.14**
Northeast all	85	94.87	66.98	1,600	1.60	92.23	14.57
Midwest all	106	90.38	91.73	1,900	1.43	94.97	11.48
South all	189	87.69	78.36	4,200	2.42	92.45	8.84
West all	125	81.27	80.68	2,400	1.74	93.21	7.71
National public	375	89.94	89.99	9,000	2.03	93.13	10.69
National private	130	68.63	62.72	986	0.38	93.94	4.16
Catholic	37	91.61	91.70	484	0.21	96.42	2.01
Non-Catholic	93	49.13	50.95	502	0.53	91.05	6.16

NOTE: National all includes national public, national private, Bureau of Indian Education (BIE), and Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) that are located in the United States. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2012 Reading Long-Term Trend Assessment.

Participation, Exclusion, and Accommodation Rates for Age 17 Reading for the 2012 LTT Assessment

The following table displays the school-level participation rates and student-level participation, exclusion, and accommodation rates for the age 17 long-term trend (LTT) reading assessment. Various weights were used in the calculation of the rates, as indicated in the column headings of the table.

The participation rates reflect the participation of the original sampled schools only and do not reflect any effect of substitution. The rates weighted by the school base weight and enrollment show the approximate proportion of the student population in the domain that is represented by the responding schools in the sample. The rates weighted by just the base weight show the proportion of the school population that is represented by the responding schools in the sample. These rates differ because schools differ in size.

Participation, exclusion, and accommodation rates for age 17 long-term trend reading assessment, by geographic region and school type: 2012

Geographic region and school type	Number of schools in original sample	School participation rates (percent) before substitution (weighted by school base weight and enrollment)	School participation rates (percent) before substitution (weighted by school base weight only)	Number of students sampled	Weighted percent of students excluded	Weighted student participation rates (percent) after makeups	Weighted percent of students accommodated
National all	**482**	**83.82**	**80.26**	**11,300**	**1.96**	**88.29**	**8.92**
Northeast all	81	92.27	74.44	2,000	2.68	84.55	13.83
Midwest all	97	90.74	90.45	2,200	1.39	89.18	10.13
South all	184	82.17	78.53	4,300	2.29	90.17	6.94
West all	120	72.76	75.82	2,700	1.43	87.90	6.96
National public	389	85.58	87.57	10,400	2.10	88.34	8.90
National private	93	62.51	60.45	858	0.13	87.64	9.18
Catholic	16	88.18	86.99	362	0.28	88.10	7.27
Non-Catholic	77	40.30	50.18	496	0.00	87.01	10.84

NOTE: National all includes national public, national private, Bureau of Indian Education (BIE), and Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) that are located in the United States. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2012 Reading Long-Term Trend Assessment.

NAEP Technical Documentation Nonresponse Bias Analysis for the 2012 LTT Assessment

NCES Statistical standards call for a nonresponse bias analysis to be conducted for a sample with a response rate below 85 percent at any stage of sampling. Weighted school response rates for the 2012 assessment indicate a need for school nonresponse bias analyses for private school samples for ages 9, 13, and 17. No student nonresponse bias analyses were necessary since the student-level participation rates for all groups were above the 85 percent participation threshold. The school-level analyses were conducted separately at each age. Thus, three separate school- level analyses were conducted.

The procedures and results from these analyses are summarized briefly below. The analyses conducted consider only certain characteristics of schools and students. They do not directly consider the effects of the nonresponse on student achievement, the primary focus of NAEP.

Thus, these analyses cannot be conclusive of either the existence or absence of nonresponse bias for student achievement. For more details, please see the NAEP 2012 LTT NRBA report (337KB).

Each school-level analysis was conducted in three parts. The first part of the analysis looked for potential nonresponse bias that was introduced through school nonresponse. The second part of the analysis examined the remaining potential for nonresponse bias after accounting for the mitigating effects of substitution. The third part of the analysis examined the remaining potential for nonresponse bias after accounting for the mitigating effects of both school substitution and school-level nonresponse weight adjustments. The characteristics examined were census region, reporting subgroup (private school type), urban-centric locale, size of school (categorical), size of school (continuous), and race/ethnicity enrollment percentages.

Based on the school characteristics available, for the private school samples at ages 13 and 17, there does not appear to be evidence of substantial potential bias resulting from school substitution or school nonresponse. However, the analyses suggest that a potential for nonresponse bias remains for the age 9 private school samples for school percentage race/ethnicity characteristics. Please see the full report for more details.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_nonresponse_bias_analysis.asp x