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

# *National Assessment of Educational Progress (NAEP) 2019 and 2020 Update Long-Term Trend (LTT) 2020 Update Emergency Clearance*

## *Appendices A-C*

*Appendix A: External Advisory Committees*

*Appendix B1: NAEP 2013 Weighting Procedures*

*Appendix B2: Long-Term Trend (LTT) 2012 Weighting Procedures*

*Appendix C1: NAEP 2019 Sampling Memo*

*Appendix C2: LTT 2020 Sampling Memo*

*OMB# 1850-0928 v.14*

NAEP
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

March 2019

# Table of Contents

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

# *National Assessment of Educational Progress (NAEP) 2019 and 2020*

# *Appendix A*

## *External Advisory Committees*

*OMB# 1850-0928 v.14*

September 2018
No changes since v.10

## Appendix A-1: NAEP Design and Analysis Committee

| Name | Affiliation |
|---|---|
| Betsy Becker | Florida State University, FL |
| Peter Behuniak | University of Connecticut, CT |
| Lloyd Bond | University of North Carolina, Greensboro, NC (Emeritus)/Carnegie Foundation (retired) |
| Derek Briggs | University of Colorado, CO |
| Steve Elliott | Arizona State University, AZ |
| Ben Hansen | University of Michigan, MI |
| Matthew Johnson | Columbia University, NY |
| Brian Junker | Carnegie Mellon University, PA |
| David Kaplan | University of Wisconsin-Madison, WI |
| Kenneth Koedinger | Carnegie Mellon University, PA |
| Sophia Rabe-Hesketh | University of California, Berkeley, CA |
| Michael Rodriguez | University of Minnesota, MN |
| S.Lynne Stokes | Southern Methodist University, TX |
| Chun Wang | University of Minnesota, MN |

## Appendix A-2: NAEP Validity Studies Panel

| Name | Affiliation |
|---|---|
| Peter Behuniak | University of Connecticut, CT |
| George Bohrnstedt | American Institutes for Research, Washington, DC |
| Jim Chromy | RTI International (Emeritus Fellow), Raleigh, NC |
| Phil Daro | Strategic Education Research (SERP) |
| Richard Duran | University of California, Berkeley, CA |
| David Grissmer | University of Virginia, VA |
| Larry Hedges | Northwestern University, IL |
| Sami Kitmitto | American Institutes for Research, San Mateo, CA |
| Ina Mullis | Boston College, MA |
| Scott Norton | Council of Chief State School Officers, Washington, DC |
| Jim Pellegrino | University of Illinois at Chicago/Learning Sciences Research Institute, IL |
| Gary Phillips | American Institutes for Research, Washington, DC |
| Lorrie Shepard | University of Colorado at Boulder, CO |
| Fran Stancavage | American Institutes for Research, San Mateo, CA |
| David Thissen | University of North Carolina at Chapel Hill, NC |
| Sheila Valencia | University of Washington, WA |
| Ting Zhang | American Institutes for Research, Washington, DC |

## Appendix A-3: NAEP Quality Assurance Technical Panel

| Name | Affiliation |
|---|---|
| Jamal Abedi | University of California, Davis, CA |
| Chuck Cowan | Analytic Focus LLC, San Antonio, TX |
| Gail Goldberg | Gail Goldberg Consulting, Ellicott City, MD |
| Brian Gong | National Center for the Improvement of Educational Assessment, Dover, NH |
| Richard Luecht | University of North Carolina-Greensboro, NC |
| Jim Pellegrino | University of Illinois at Chicago/Learning Sciences Research Institute, IL |
| Mark Reckase | Michigan State University, MI |
| Michael (Mike) Russell | Boston College, MA |
| Phoebe Winter | Consultant, Chesterfield, VA |
| Richard Wolfe | University of Toronto (Emeritus), Ontario, Canada |

## Appendix A-4: NAEP National Indian Education Study Technical Review Panel

| Name | Affiliation |
|---|---|
| Doreen E. Brown | ASD Education Center, Anchorage, AK |
| Robert B.Cook | Native American Initiative/Teach for America, Summerset, SD |
| Steve Andrew Culpepper | University of Illinois at Urbana-Champaign, IL |
| Susan C. Faircloth | University of North Carolina Wilmington, NC |
| Jeremy MacDonald | Rocky Boy Elementary, Box, Elder, MT |
| Holly Jonel Mackey | University of Oklahoma, OK |
| Jeannette Muskett Miller | Tohatchi High School, Tohatchi, NM |
| Sedelta Oosahwee | National Education Association, DC |
| Debora Norris | Salt River Pima-Maicopa Indian Community |
| Martin Reinhardt | Northern Michigan University, MI |
| Tarajean Yazzie-Mintz | Wakanyeja ECE Initative/American Indian College Fund, Denver, CO |

## Appendix A-5: Geography Standing Committee

| Name | Affiliation |
|---|---|
| Sarah Bednarz | Texas A&M University, TX |
| Osa Brand | National Council for Geographic Education, Washington, DC |
| Seth Dixon | Rhode Island College, RI |
| Charlie Fitzpatrick | ESRI Schools, Arlington, VA |
| Ruth Luevanos | Pacoima Middle School, Pacoima, CA |
| Joe Stoltman | Western Michigan University, MI |
| Kelly Swanson | Johnson Senior High, St. Paul, MN |

## Appendix A-6: NAEP Civics Standing Committee

| Name | Affiliation |
|---|---|
| Patricia Avery | University of Minnesota, MN |
| Christopher Elnicki | Cherry Creek School District, Greenwood Village, CO |
| Fay Gore | North Carolina Public Schools, Raleigh, NC |
| Barry Leshinsky | Challenger, NC Middle School, Huntsville, AL |
| Peter Levine | CIRCLE (Center for Information & Research on Civic Learning and Engagement), Medford, MA |
| Clarissa Peterson | DePauw University, IN |
| Terri Richmond | Golden Valley High School, Bakersfield, CA |
| Jackie Viana | Miami-Dade County Schools, Miami, FL |

## Appendix A-7 NAEP Economics Standing Committee

| Name | Affiliation |
|---|---|
| Kris Bertelsen | Little Rock Branch-Federal Reserve Bank of St. Louis, Little Rock, AR |
| William Bosshardt | Florida Atlantic University, FL |
| Stephen Buckles | Vanderbilt University, TN |
| Andrea Caceres-Santamaria | Seminole Ride Community High School, FL |
| Steven L. Cobb | University of North Texas, TX |
| Kristen S. McDaniel | Wisconsin Dept. of Public Instruction, WI |
| Richard MacDonald | St. Cloud State University, MN |
| Kevin Smith | Renaissance High School, Detroit, MI |
| William Walstad | University of Nebraska–Lincoln, NE |

## Appendix A-8: NAEP Mathematics Standing Committee

| Name | Affiliation |
|---|---|
| Scott Baldridge | Louisiana State University, LA |
| Carl Cowen | Indiana University–Purdue University, IN |
| Kathleen Heid | Pennsylvania State University, PA |
| Mark Howell | Gonzaga College High School, Washington, DC |
| Carolyn Maher | Rutgers University, NJ |
| Michele Mailhot | Maine Department of Education, Augusta, ME |
| Matthew Owens | Spring Valley High School, Columbia, SC |
| Carole Philip | Alice Deal Middle School, Washington, DC |
| Kayonna Pitchford | University of North Carolina, NC |
| Melisa M. Ramos Trinidad | Educación Bilingüe Luis Muñoz Iglesias, Cidra, PR |
| Allan Rossman | College of Science and Mathematics-CalPoly, CA |
| Carolyn Sessions | Louisiana Department of Education, LA |
| Lya Snell | Georgia Department of Education, GA |
| Ann Trescott | Stella Maris Academy, La Jolla, CA |
| Vivian Valencia | Espanola Public Schools, NM |

## Appendix A-9: NAEP Reading Standing Committee

| Name | Affiliation |
|---|---|
| Peter Afflerbach | University of Maryland, MD |
| Patricia Alexander | University of Maryland, MD |
| Alison Bailey | University of California, LA, CA |
| Katrina Boone | Kentucky Department of Education, KY |
| Margretta Browne | Richard Montgomery High School, Silver Spring, MD |
| Julie Coiro | University of Rhode Island, RI |
| Bridget Dalton | University of Colorado Boulder, CO |
| Jeanette Mancilla-Martinez | Vanderbilt University, TN |
| Pamela Mason | Harvard Graduate School of Education, MA |
| P. David Pearson | University of California, Berkeley, CA |
| Frank Serafini | Arizona State University, AZ |
| Kris Shaw | Kansas State Department of Education, KS |
| Diana Townsend | University of Nevada, Reno, NV |
| Victoria Young | Texas Education Agency, Austin, TX |

## Appendix A-10: NAEP Science Standing Committee

| Name | Affiliation |
|---|---|
| Alicia Cristina Alonzo | Michigan State University, MI |
| George Deboer | American Association for the Advancement of Science, Washington, DC |
| Alex Decaria | Millersville University, PA |
| Crystal Edwards | Lawrence Township Public Schools, Lawrenceville, NJ |
| Ibari Igwe | Shrewd Learning, Elkridge, MD |
| Michele Lombard | Kenmore Middle School, Arlington, VA |
| Emily Miller | Consultant, WI |
| Blessing Mupanduki | Department of Defense, Washington, DC |
| Amy Pearlmutter | Littlebrook Elementary School, Princeton, NJ |
| Brian Reiser | Northwestern University, Evanston, IL |
| Michal Robinson | Alabama Department of Education, Montgomery, AL |
| Gloria Schmidt | Darby Junior High School, Fort Smith, AR |
| Steve Semken | Arizona State University, Tempe, AZ |
| Roberta Tanner | Board of Science Education, Longmont, CO |
| David White | Lamoille North Supervisory Union School District, Hyde Park, VT |

## Appendix A-11: NAEP Survey Questionnaires Standing Committee

| Name | Affiliation |
|---|---|
| Angela Duckworth | University of Pennsylvania, PA |
| Hunter Gehlbach | Harvard University, MA |
| Camille Farrington | University of Chicago, Chicago, IL |
| Gerunda Hughes | Howard University, DC |
| David Kaplan | University of Wisconsin-Madison, WI |
| Henry Levin | Teachers College, Columbia University, NY |
| Stanley Presser | University of Maryland, MD |
| Augustina Reyes | University of Houston, Houston, TX |
| Leslie Rutkowski | Indiana University Bloomington, IN |
| Jonathon Stout | Lock Haven University, PA |
| Roger Tourangeau | Westat, Rockville, MD |
| Akane Zusho | Fordham University, NY |

## Appendix A-12: NAEP Technology and Engineering Literacy Standing Committee

| Name | Affiliation |
|---|---|
| Keith Barton | Indiana University Bloomington, IN |
| John Behrens | Pearson eLEADS Center, Mishawaka, IN |
| Brooke Bourdelat-Parks | Biological Sciences Curriculum Study, Colorado Springs, CO |
| Barbara Bratzel | Shady Hill School, Cambridge, MA |
| Lewis Chappelear | James Monroe High School, North Hills, CA |
| Britte Haugan Cheng | SRI International, Menlo Park, CA |
| Meredith Davis | North Carolina State University, NC |
| Chris Dede | Harvard Graduate School of Education, MA |
| Richard Duran | University of California, Santa Barbara, CA |
| Maurice Frazier | Oscar Smith High School, Chesapeake, VA |
| Camilla Gagliolo | Arlington Public Schools, Arlington, VA |
| Christopher Hoadley | New York University, NY |
| Eric Klopfer | Massachusetts Institute of Technology, MA |
| Beth McGrath | Stevens Institute of Technology, NJ |
| Greg Pearson | National Academy of Engineering, Washington, DC |
| John Poggio | University of Kansas, KS |
| Erin Reilly | University of Southern California, CA |
| Troy Sadler | University of Missouri Science Education Center, Columbia, MO |
| Kimberly Scott | Arizona State University, AZ |
| Teh-Yuan Wan | New York State Education Department, Albany, NY |

## Appendix A-13: NAEP U.S. History Standing Committee

| Name | Affiliation |
|---|---|
| Keith Barton | Indiana University Bloomington, IN |
| Michael Bunitsky | Frederick County Public Schools, Frederick, MD |
| Teresa Herrera | Shenandoah Middle School, Miami, FL |
| Cosby Hunt | Center for Inspired Teaching, Washington, DC |
| Helen Ligh | Macy Intermediate School, Monterey, CA |
| Amanda Prichard | Green Mountain High School, Lakewood, CO |
| Kim Rasmussen | Auburn Washburn Unified School District, Topeka, KS |
| Diana Turk | New York University, New York, NY |

## Appendix A-14: NAEP Mathematics Translation Review Committee

| Name | Affiliation |
|---|---|
| Mayra Aviles | Puerto Rico Department of Education, PR |
| David Feliciano | P.S./M.S 29, The Melrose School, Bronx, NY |
| Yvonne Fuentes | Author and Spanish Linguist, Carrollton, GA |
| Marco Martinez-Leandro | Sandia High School, NM |
| Jose Antonio (Tony) Paulino | Nathan Straus Preparatory School, NY |
| Evelisse Rosado Rivera | Teacher, PMB 35 HC, PR |
| Myrna Rosado-Rasmussen | Austin Independent School District, TX |
| Gloria Rosado Vazquez | Teacher, HC-02, PR |
| Enid Valle | Kalamazoo College, Kalamazoo, MI |

## Appendix A-15: NAEP Science Translation Review Committee

| Name | Affiliation |
|---|---|
| Daniel Berdugo | Teacher, PS 30X Wilton, NY |
| Yvonne Fuentes | Author and Spanish Linguist, Carrollton, GA |
| Myrna Rosado- Rasmussen | Austin Independent School District, Austin, TX |
| Enid Valle | Kalamazoo College, Kalamazoo, MI |

## Appendix A-16: NAEP Grade 8 Social Science Translation Review Committee

| Name | Affiliation |
|---|---|
| Yvonne Fuentes | Author and Spanish Linguist, Carrollton, GA |
| Jose Antonio Paulino | Middle School Teacher, Nathan Strauss Preparatory School, NY |
| Dagoberto Eli Ramierz | Bilingual Education Expert, Palmhurst, TX |
| Enid Valle | Kalamazoo College, Kalamazoo, MI |

## Appendix A17: NAEP Grade 4 and 8 Survey Questionnaires and eNAEP DBA System Translation Committee

| Name | Affiliation |
|---|---|
| Daniel Berdugo | PS 30X Wilton, Bronx, NY |
| Yvonne Fuentes | Carrollton, GA |
| Marco Martinea-Leandro | Sandia High School. Albuquerque, NM |
| Jose Antonio (Tony) Paulino | Nathan Straus Preparatory School, New York, NY |
| Evelisse Rosado Rivera | PMB 36 HC 72, Naranjito, PR |
| Myrna Rosado-Rasmussen | Austin Independent School District, Austin, TX |
| Gloria M. Rosado Vazquez | HC – 02 Barranquitas, PR |
| Enid Valle | Kalamazoo College, Kalamazoo, MI |

## Appendix A-18: NAEP Writing Standing Committee

| Name | Affiliation |
|---|---|
| Margretta Browne | Montgomery County Public Schools, Silver Spring, MD |
| Dina Decristofaro | Scituate Middle School, RI |
| Elyse Eidman-Aadahl | National Writing Project, Berkeley, CA |
| Nikki Elliot-Schuman | Smarter Balanced Assessment Consortium |
| Charles MacArthur | University of Delaware, Newark, DE |
| Michael McCloskey | Johns Hopkins University, Baltimore, MD |
| Norma Mota-Altman | San Gabriel High School, Alhambra, CA |
| Sandra Murphy | University of California, Davis, Walnut Creek, CA |
| Peggy O'Neill | Loyola University Maryland, MD |
| Laura Roop | University of Pittsburgh School of Education, PA |
| Drew Sterner | Tamanend Middle School, Warrington, PA |
| Sherry Swain | National Writing Project, Berkeley, CA |
| Jason Torres-Rangel | University of California, CA |
| Victoria Young | Texas Education Agency, Austin, TX |

## Appendix A-19: NAEP Principals' Panel Standing Committee

| Name | Affiliation |
|---|---|
| David Atherton | Clear Creek Middle School, Gresham, OR |
| Ardith Bates | Gladden Middle School, Chatsworth, GA |
| Williams Carozza | Harold Martin Elementary School, Hopkinton, NH |
| Diane Cooper | St. Joseph's Academy, Clayton, MO |
| Brenda Creel | Alta Vista Elementary School, Cheyenne, WY |
| Rita Graves | Pin Oak Middle School, Bellaire, TX |
| Don Hoover | Lincoln Junior High School, Springdale, AR |
| Stephen Jackson | (Formerly with) Paul Laurence Dunbar High School, Washington, DC |
| Anthony Lockhart | Lake Shore Middle School, Belle Glade, FL |
| Susan Martin | Berrendo Middle School, Roswell, NM |
| Lillie McMillan | Porter Elementary School, San Diego, CA |
| Kourtney Miller | Chavez Prep Middle School, Washington, DC |
| Jason Mix | Howard Lake–Waverly–Winsted High School, Howard Lake, MN |
| Leon Oo-Sah-We | Ch'ooshgai Community School, Tohatchi, NM |
| Sylvia Rodriguez Vargas | Atlanta Girls' School, Atlanta Georgia, GA |

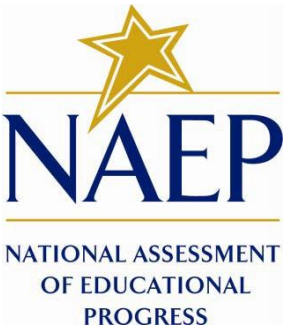
*NATIONAL CENTER FOR EDUCATION STATISTICS*
*NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS*

# *National Assessment of Educational Progress (NAEP) 2019 and 2020*

## *Appendix B1*

## *NAEP 2013 Weighting Procedures*

*OMB# 1850-0928 v.14*




September 2018
No changes since v.10

NAEP Technical Documentation Website

# NAEP Technical DocumentationWeighting Procedures for the 2013 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 2013, weights were developed for students sampled at grades 4, 8, and 12 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 contains the following major components:

- the student base weight;
- school nonresponse adjustments;
- student nonresponse adjustments;
- school weight trimming adjustments;
- student weight trimming adjustments; and
- student raking adjustment.

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 2013 sample design section.

The student 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 participating 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.

An additional weighting adjustment was implemented in the state and Trial Urban District Assessment (TUDA) samples so that estimates for key student-level characteristics were in agreement across assessments in reading and mathematics. This adjustment was implemented using a raking procedure.

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, approximately 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. A finite population correction (fpc) factor was incorporated into the replication scheme so that it could be reflected in the variance estimates for the reading and mathematics assessments. See Computation of Replicate Weights for Variance Estimation for details.

Quality control checks were carried out 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.

In the linked pages that follow, please note that Vocabulary, Reading Vocabulary, and Meaning Vocabulary refer to the same reporting scale and are interchangeable.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/naep_assessment_weighting_procedures.aspx

NAEP Technical Documentation Website

# NAEP Technical Documentation Computation of Full-Sample Weights for the 2013 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 grade (4, 8, or 12) and assessment subject (reading or mathematics). 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:

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

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;
- $STU\_TRIM_{jsk}$ is the student-level weight trimming adjustment factor; and
- $STU\_RAKE_{jsk}$ is the student-level raking adjustment factor.

School sampling strata for a given assessment vary by school type and grade. See the links below for descriptions of the school strata for the various assessments.

- Public schools at grades 4 and 8
- Public schools at grade 12
- Private schools at grades 4, 8 and 12

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/computation_of_full_sample_weights_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical Documentation Computation of Base Weights for the 2013 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

- type of sampled school (original or substitute); and
- 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 reading or mathematics assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/computation_of_base_weights_for_the_2013_assessment.aspx

# NAEP Technical Documentation School Base Weights for the 2013 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);
- sampling frame (new school frame or not).

The overall selection probability of an originally selected school in a reading or mathematics sample is equal to its probability of selection from the NAEP public/private school frame.

The overall selection probability of a school from the new school frame in a reading or mathematics sample 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 the place of original schools if they refuse to participate. For weighting purposes, they are treated as if they were the original schools that they replaced; so substitute schools are assigned the school base weight of the original schools.

Learn more about substitute schools for the 2013 private school national assessment and substitute schools for the 2013 twelfth grade public school assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/school_base_weights_for_the_2013_assessment.aspx

# NAEP Technical Documentation Student Base Weights for the 2013 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;
- $SCHSsessionassignmentESWT_{js}$ 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 Stu_bookmarkthe student-level session assignment weight that reflects the conditional probability, given that the particular session type was assigned to the school, that the student was assigned to the session type;
- $SUBJFACsubjfac_{jsk}$ is the subject spiral adjustment factor that reflects the conditional probability, given that 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 to be included in the 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 reading or mathematics booklets to the students. The subject factor varies by grade, subject, and school type (public or private), and it is equal to the inverse of the booklet proportions (reading or mathematics) in the overall spiral for a specific sample.

For cooperating substitutes of nonresponding original sampled schools, the substitution adjustment factor ($SUBADJ_{js}$) is equal to the ratio of the estimated grade enrollment for the original sampled school to the estimated grade enrollment for the substitute school. The student sample from the substitute school then "represents" the set of grade-eligible students from the original 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/2013/student_base_weights_for_the_2013_assessment.aspx

# NAEP Technical Documentation School and Student Nonresponse Weight Adjustments for the 2013 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 of the sample estimates. 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 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/2013/school_and_student_nonresponse_weight_adjustments_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical Documentation School Nonresponse Weight Adjustment

The school nonresponse adjustment procedure inflates the weights of cooperating schools to account for eligible noncooperating schools for which no substitute schools participated. The adjustments are computed within nonresponse cells and are based on the assumption that the cooperating and noncooperating schools within the same cell are more similar to each other than to schools from different cells. School nonresponse adjustments were carried out separately by sample; that is, by

- sample level (state, national),
- school type (public, private), and
- grade (4, 8, 12).



http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/school_nonresponse_weight_adjustment_for_the_2013_assessment.aspx

# NAEP Technical Documentation Development of Initial School Nonresponse Cells

The cells for nonresponse adjustments are generally functions of the school sampling strata for the individual samples. School sampling strata usually differ by assessment subject, grade, and 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 2015, all operational assessments were spiraled together.

The initial nonresponse cells for the various NAEP 2015 samples are described below.

## Public School Samples for Reading and Mathematics at Grades 4 and 8

For these samples, initial weighting cells were formed within each jurisdiction using the following nesting cell structure:

- Trial Urban District Assessment (TUDA) district vs. the balance of the state for states with TUDA districts,
- urbanicity (urban-centric locale) stratum; and
- race/ethnicity classification stratum, or achievement level, or median income, or grade enrollment.

In general, the nonresponse cell structure used race/ethnicity classification stratum as the lowest level variable. However, where there was only one race/ethnicity classification stratum within a particular urbanicity stratum, categorized achievement, median income, or enrollment data were used instead.

## Public School Sample at Grade 12

The initial weighting cells for this sample were formed using the following nesting cell structure:

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

## Private School Samples at Grades 4, 8 and 12

The initial weighting cells for these samples were formed within each grade using the following nesting cell structure:

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

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/development_of_initial_school_nonresponse_cells_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical Documentation Development of Final School Nonresponse Cells

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 at Grades 4 and 8

For the grade 4 and 8 public school samples, cells with the most similar race/ethnicity classification within a given jurisdiction/Trial Urban District Assessment (TUDA) district and urbanicity (urban-centric locale) stratum were collapsed first. If further collapsing was required after all levels of race/ethnicity strata were collapsed, cells with the most similar urbanicity strata were combined next. Cells were never permitted to be collapsed across jurisdictions or TUDA districts.

## Public School Sample at Grades 12

For the grade 12 public school sample, race/ethnicity classification cells within a given census division stratum and urbanicity stratum were collapsed first. If further collapsing was required after all levels of race/ethnicity classification were collapsed, cells with the most similar urbanicity strata were combined next. Any further collapsing occurred across census division strata but never across census regions.

## Private School Samples at Grades 4, 8, and 12

For the private school samples, cells with the most similar race/ethnicity classification within a given affiliation, census division, and urbanicity stratum were collapsed first. If further collapsing was required after all levels of race/ethnicity strata were collapsed, cells with the most similar urbanicity classification were combined. Any further collapsing occurred across census division strata but never across affiliations.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/development_of_final_school_nonresponse_cells_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical Documentation School Nonresponse Adjustment Factor Calculation

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 grade enrollment corresponding to the original sampled school.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/school_nonresponse_adjustment_factor_calculation_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical Documentation Student Nonresponse Weight Adjustment



The student nonresponse adjustment procedure inflates the weights of assessed students to account for eligible sampled students who did not participate in the assessment. These inflation factors offset the loss of data associated with absent students. The adjustments are computed within nonresponse cells and are based on the assumption that the assessed and absent students within the same cell are more similar to one another than to students from different cells. Like its counterpart at the school level, the student nonresponse adjustment is intended to reduce the mean square error and thus improve the accuracy of NAEP assessment estimates. Also, like its counterpart at the school level, student nonresponse adjustments were carried out separately by sample; that is, by

- grade (4, 8, 12),
- school type (public, private), and
- assessment subject (mathematics, reading, science, meaning vocabulary).

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/student_nonresponse_weight_adjustment_for_the_2013_assessment.aspx

# NAEP Technical Documentation Development of Initial Student Nonresponse Cells for the 2013 Assessment

Initial student nonresponse cells are generally created within each sample as defined by grade, school type (public, private), and assessment subject. However, when subjects are administered together by way of spiraling, the initial student nonresponse cells are created across the subjects in the same spiral. The rationale behind this decision is that spiraled subjects are in the same schools and the likelihood of whether an eligible student participates in an assessment is more related to its school than the subject of the assessment booklet. In NAEP 2013, there was only one spiral, with the reading and mathematics assessments spiraled together. The initial student nonresponse cells for the various NAEP 2013 samples are described below.

Nonresponse adjustment procedures are not applied to excluded students because they are not required to complete an assessment.

## Public School Samples for Reading and Mathematics at Grades 4 and 8

The initial student nonresponse cells for these samples were defined within grade, jurisdiction, and Trial Urban District Assessment (TUDA) district using the following nesting cell structure:

- students with disabilities (SD)/English language learners (ELL) by subject,
- school nonresponse cell,
- age (classified into "older"[1] student and "modal age or younger" student),
- gender, and
- race/ethnicity.

The highest level variable in the cell structure separates students who were classified either as having disabilities (SD) or as English language learners (ELL) from those who are neither, since SD or ELL students tend to score lower on assessment tests than non-SD/non-ELL students. In addition, the students in the SD or ELL groups are further broken down by subject, since rules for excluding students from the assessment differ by subject. Non-SD and non-ELL students are not broken down by subject, since the exclusion rules do not apply to them.

## Public School Samples for Reading and Mathematics at Grade 12

The initial weighting cells for these samples were formed hierarchically within state for the state-reportable samples and the balance of the country for remaining states as follows:

- SD/ELL,
- school nonresponse cell,
- age (classified into "older"[1] student and "modal age or younger" student),
- gender, and
- race/ethnicity.

## Private School Samples for Reading and Mathematics at Grades 4, 8, and 12

The initial weighting cells for these private school samples were formed hierarchically within grade as follows:

- SD/ELL,
- school nonresponse cell,
- age (classified into "older"[1] student and "modal age or younger" student),
- gender, and
- race/ethnicity.

Although exclusion rules differ by subject, there were not enough SD or ELL private school students to break out by subject as was done for the public schools.

[1]Older students are those born before October 1, 2002, for grade 4; October 1, 1998, for grade 8; and October 1, 1994, for grade 12.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/development_of_initial_student_nonresponse_cells_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical Documentation Development of Final Student Nonresponse Cells for the 2013 Assessment

Similar to the school nonresponse adjustment, cell and adjustment factor size constraints are in place to prevent unstable nonresponse adjustments or unacceptably large adjustment factors. All initial weighting cells with either fewer than 20 participating students or adjustment factors greater than 2.0 for the full sample weight were collapsed with suitable adjacent cells. Simultaneously, all initial weighting cells for any replicate with either fewer than 15 participating students or an adjustment factor greater than the maximum of 2.0 or 1.5 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. Race/ethnicity cells within SD/ELL groups, school nonresponse cell, age, and gender classes were collapsed first. If further collapsing was required after collapsing all race/ethnicity classes, cells were next combined across gender, then age, and finally school nonresponse cells. Cells are never collapsed across SD and ELL groups for any sample.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/development_of_final_student_nonresponse_cells_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical DocumentationStudent Nonr esponse Adjustment Factor Calculation

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.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/student_nonresponse_adjustment_factor_calculation_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical Documentation School and Student Weight Trimming Adjustments for the 2013 Assessment

Trimming of School Base Weights

Trimming of Student Weights

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 employing 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/2013/school_and_student_weight_trimming_adjustments_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical DocumentationTrimming of School Base Weights

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.

Thirty-seven (37) schools out of 377 selected from the new-school sampling frame had their weights trimmed: eight at grade 4, 29 at grade 8, and zero at grade 12.

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 had their weights trimmed.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/trimming_of_school_base_weights_for_the_2013_assessment.aspx

# NAEP Technical DocumentationTrimming of Student Weights

Large student weights generally come from compounding nonresponse adjustments at the school and student levels with artificially low school 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 occur due to compounding effects of the various weighting components.

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. Student weight trimming was implemented separately by grade, school type (public or private), and subject. The multiples used were 3.5 for public school trimming groups and 4.5 for private school trimming groups. Trimming groups were defined by jurisdiction and Trial Urban District Assessment (TUDA) districts for the public school samples at grades 4 and 8; by dichotomy of low/high percentage of Black and Hispanic students (15 percent and below, above 15 percent) for the public school sample at grade 12; and by affiliation (Catholic, Non-Catholic) for private school samples at grades 4, 8 and 12.

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 received a trimming factor calculated as follows:

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

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 2013 assessment, relatively few students had weights considered excessively large. Out of the approximately 840,000 students included in the combined 2013 assessment samples, 226 students had their weights trimmed.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/trimming_of_student_weights_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical DocumentationStudent Weight Raking Adjustment for the 2013 Assessment



Weighted estimates of population totals for student-level subgroups for a given grade will vary across subjects even though the student samples for each subject generally come from the same schools. These differences are the result of sampling error associated with the random assignment of subjects to students through a process known as spiraling. For state assessments in particular, any difference in demographic estimates between subjects, no matter how small, may raise concerns about data quality. To remove these random differences and potential data quality concerns, a new step was added to the NAEP weighting procedure starting in 2009. This step adjusts the student weights in such a way that the weighted sums of population totals for specific subgroups are the same across all subjects. It was implemented using a raking procedure and applied only to state-level assessments.

Raking is a weighting procedure based on the iterative proportional fitting process developed by Deming and Stephan (1940) and involves simultaneous ratio adjustments to two or more marginal distributions of population totals. Each set of marginal population totals is known as a dimension, and each population total in a dimension is referred to as a control total. Raking is carried out in a sequence of adjustments. Sampling weights are adjusted to one marginal distribution and then to the second marginal distribution, and so on. One cycle of sequential adjustments to the marginal distributions is called an iteration. The procedure is repeated until convergence is achieved. The criterion for convergence can be specified either as the maximum number of iterations or an absolute difference (or relative absolute difference) from the marginal population totals. More discussion on raking can be found in Oh and Scheuren (1987).

For NAEP 2013, the student raking adjustment was carried out separately in each state for the reading and mathematics public school samples at grades 4 and 8, and in the 13 states with state-reportable samples for the reading and mathematics public school samples at grade 12. The dimensions used in the raking process were National School Lunch Program (NSLP) eligibility, race/ethnicity, SD/ELL status, and gender. The control totals for these dimensions were obtained from the NAEP student sample weights of the reading and mathematics samples combined.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/student_weight_raking_adjustment_for_the_2013_assessment.aspx

# NAEP Technical Documentation Development of Final Raking Dimensions

The raking procedure involved four dimensions. The variables used to define the dimensions are listed below along with the categories making up the initial raking cells for each dimension.

- National School Lunch Program (NSLP) eligibility

    1. Eligible for free or reduced-price lunch
    2. Otherwise

- Race/Ethnicity

    1. White, not Hispanic
    2. Black, not Hispanic
    3. Hispanic
    4. Asian
    5. American Indian/Alaska Native
    6. Native Hawaiian/Pacific Islander
    7. Two or More Races

- SD/ELL status

    1. SD, but not ELL
    2. ELL, but not SD
    3. SD and ELL
    4. Neither SD nor ELL

- Gender

    1. Male
    2. Female

In states containing districts that participated in Trial Urban District Assessments (TUDA) districts at grades 4 and 8, the initial cells were created separately for each TUDA district and the balance of the state. Similar to the procedure used for school and student nonresponse adjustments, limits were placed on the magnitude of the cell sizes and adjustment factors to prevent unstable raking adjustments that could have resulted in unacceptably large or small adjustment factors. Levels of a dimension were combined whenever there were fewer than 30 assessed or excluded students (20 for any of the replicates) in a category, if the smallest adjustment was less than 0.5, or if the largest adjustment was greater than 2 for the full sample or for any replicate.

If collapsing was necessary for the race/ethnicity dimension, the following groups were combined first: American Indian/Alaska Native with Black, not Hispanic; Hawaiian/Pacific Islander with Black, not Hispanic; Two or More Races with White, not Hispanic; Asian with White, not Hispanic; and Black, not Hispanic with Hispanic. If further collapsing was necessary, the five categories American Indian/Alaska Native; Two or More Races; Asian; Native Hawaiian/Pacific Islander; and White, not Hispanic were combined. In some instances, all seven categories had to be collapsed.

If collapsing was necessary for the SD/ELL dimension, the SD/not ELL and SD/ELL categories were combined first, followed by ELL/not SD if further collapsing was necessary. In some instances, all four categories had to be collapsed.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/development_of_final_raking_dimensions_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical DocumentationRaking Adjustment Control Totals for the 2013 Assessment

The control totals used in the raking procedure for NAEP 2013 grades 4, 8, and 12 were estimates of the student population derived from the set of assessed and excluded students pooled across subjects. The control totals for category c within dimension d were computed as follows:

$$TOTAL_{c(d)} = \sum_{R_{c(d)} \cup E_{c(d)}} \frac{STU\_BWT_k \times SCH\_TRIM_k \times SCH\_NRAF_k \times STU\_NRAF_k}{SUBJFAC_k}$$

where

- $R_{c(d)}$ is the set of all assessed students in category c of dimension d,
- $E_{c(d)}$ is the set of all excluded students in category c of dimension d,
- $STU\_BWT_k$ 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,
- $STU\_NRAF_k$ is the student-level nonresponse adjustment factor for student k, and
- $SUBJFAC_k$ is the subject factor for student k.

The student weight used in the calculation of the control totals above is the adjusted student base weight, without regard to subject, adjusted for school weight trimming, school nonresponse, and student nonresponse. Control totals were computed for the full sample and for each replicate independently.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/raking_adjustment_control_totals_for_the_2013_assessment.aspx

# NAEP Technical DocumentationRaking Adjustment Factor Calculation for the 2013 Assessment

For assessed and excluded students in a given subject, the raking adjustment factor $STU\_RAKE_k$ was computed as follows:

First, the weight for student k was initialized as follows:

$$STUSAWT_k^{adj(0)} = STU\_BWT_k \times SCH\_TRIM_k \times SCH\_NRAF_k \times STU\_NRAF_k \times SUBJFAC_k$$

where

- $STU\_BWT_k$ 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,
- $STU\_NRAF_k$ is the student-level nonresponse adjustment factor for student k, and
- $SUBJFAC_k$ is the subject factor for student k.

Then, the sequence of weights for the first iteration was calculated as follows for student k in category c of dimension d:

For dimension 1:

$$STUSAWT_k^{adj(1)} = \frac{TOTAL_{c(1)}}{\sum_{R_{c(1)} \cup E_{c(1)}} STUSAWT_k^{adj(0)}} \times STUSAWT_k^{adj(0)}$$

For dimension 2:

$$STUSAWT_k^{adj(2)} = \frac{TOTAL_{c(2)}}{\sum_{R_{c(2)} \cup E_{c(2)}} STUSAWT_k^{adj(1)}} \times STUSAWT_k^{adj(1)}$$

For dimension 3:

$$STUSAWT_k^{adj(3)} = \frac{TOTAL_{c(3)}}{\sum_{R_{c(3)} \cup E_{c(3)}} STUSAWT_k^{adj(2)}} \times STUSAWT_k^{adj(2)}$$

For dimension 4:

$$STUSAWT_k^{adj(4)} = \frac{TOTAL_{c(4)}}{\sum_{R_{c(4)} \cup E_{c(4)}} STUSAWT_k^{adj(3)}} \times STUSAWT_k^{adj(3)}$$

where

- $R_{c(d)}$ is the set of all assessed students in category c of dimension d,
- $E_{c(d)}$ is the set of all excluded students in category c of dimension d, and
- $Total_{c(d)}$ is the control total for category c of dimension d.

The process is said to converge if the maximum difference between the sum of adjusted weights and the control totals is 1.0 for each category in each dimension. If after the sequence of adjustments the maximum difference was greater than 1.0, the process continues to the next iteration, cycling back to the first dimension with the initial weight for student k equaling $STUSAWT_k^{adj(4)}$ from the previous iteration. The process continued until convergence was reached.

Once the process converged, the adjustment factor was computed as follows:

$$STU\_RAKE_k = \frac{STUSAWT_k}{STU\_BWT_k \times SCH\_TRIM_k \times SCH\_NRAF_k \times STU\_NRAF_k \times SUBJFAC_k}$$

where $STUSAWT_k$ is the weight for student k after convergence.

The process was done independently for the full sample and for each replicate.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/raking_adjustment_factor_calculation_for_the_2013_assessment.aspx

# NAEP Technical DocumentationComputation of Replicate W eights for the 2013 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, approximately unbiased estimates of sampling variance are obtained. This section gives the specifics for generating the replicate weights for the 2013 assessment samples. The theory that underlies the jackknife variance estimators used in NAEP studies is discussed in the section Replicate Variance Estimation.

In general, the process of creating jackknife replicate weights takes place at both the school and student level. The precise implementation differs between those samples that involve the selection of Primary Sampling Units (PSUs) and those where the school is the first stage of sampling. The procedure for this second kind of sample also differed starting in 2011 from all previous NAEP assessments. The change that was implemented permitted the introduction of a finite population correction factor at the school sampling stage, developed by Rizzo and Rust (2011). In assessments prior to 2011, this adjustment factor has always been implicitly assumed equal to 1.0, resulting in some overestimation of the sampling variance.

For each sample, the calculation of replicate weighting factors at the school level was conducted in a series of steps. First, each school was assigned to one of 62 variance estimation strata. Then, a random subset of schools in each variance estimation stratum was assigned a replicate factor of between 0 and 1. Next, the remaining subset of schools in the same variance stratum was assigned a complementary replicate factor greater than 1. All schools in the other variance estimation strata were assigned a replicate factor of exactly 1. This process was repeated for each of the 62 variance estimation strata so that 62 distinct replicate factors were assigned to each school in the sample.

This process was then repeated at the student level. Here, each individual sampled student was assigned to one of 62 variance estimation strata, and 62 replicate factors with values either between 0 and 1, greater than 1, or exactly equal to 1 were assigned to each student.

For example, consider a single hypothetical student. For replicate 37, that student's student replicate factor might be 0.8, while for the school to which the student belongs, for replicate 37, the school replicate factor might be 1.6. Of course, for a given student, for most replicates, either the student replicate factor, the school replicate factor, or (usually) both, is equal to 1.0.

A replicate weight was calculated for each student, for each of the 62 replicates, using weighting procedures similar to those used for the full-sample weight. Each replicate weight contains the school and student replicate factors described above. By repeating the various weighting procedures on each set of replicates, 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 in stratum j can be expressed as follows:

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

where

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

Specific school and student nonresponse and student-level raking 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 and raking 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/2013/computation_of_replicate_weights_for_the_2013_assessment.aspx

# NAEP Technical Documentation Defining Variance Strata and Forming Replicates for the 2013 Assessment

In the NAEP 2013 assessment, replicates were formed separately for each sample indicated by grade (4, 8, 12), school type (public, private), and assessment subject (mathematics, reading). To reflect the school-level finite population corrections in the variance estimators for the two-stage samples used for the mathematics and reading assessments, replication was carried out at both the school and student levels.

The first step in forming replicates was to create preliminary variance strata in each primary stratum. This was done by sorting the appropriate sampling unit (school or student) in the order of its selection within the primary stratum and then pair off adjacent sampling units into preliminary variance strata. Sorting sample units by their order of sample selection reflects the implicit stratification and systematic sampling features of the sample design. Within each primary stratum with an even number of sampling units, all of the preliminary variance strata consisted of pairs of sampling units. However, within primary strata with an odd number of sampling units, all but one variance strata consisted of pairs of sampling units, while the last one consisted of three sampling units.

The next step is to form the final variance strata by combining preliminary strata if appropriate. If there were more than 62 preliminary variance strata within a primary stratum, the preliminary variance strata were grouped to form 62 final variance strata. This grouping effectively maximized the distance in the sort order between grouped preliminary variance strata. The first 62 preliminary variance strata, for example, were assigned to 62 different final variance strata in order (1 through 62), with the next 62 preliminary variance strata assigned to final variance strata 1 through 62, so that, for example, preliminary variance stratum 1, preliminary variance stratum 63, preliminary variance stratum 125 (if in fact there were that many), etc., were all assigned to the first final variance stratum.

If, on the other hand, there were fewer than 62 preliminary variance strata within a primary stratum, then the number of final variance strata was set equal to the number of preliminary variance strata. For example, consider a primary stratum with 111 sampled units sorted in their order of selection. The first two units were in the first preliminary variance stratum; the next two units were in the second preliminary variance stratum, and so on, resulting in 54 preliminary variance strata with two sample units each (doublets). The last three sample units were in the 55th preliminary variance stratum (triplet). Since there are no more than 62 preliminary variance strata, these were also the final variance strata.

Within each preliminary variance stratum containing a pair of sampling units, one sampling unit was randomly assigned as the first variance unit and the other as the second variance unit. Within each preliminary variance stratum containing three sampling units, the three first-stage units were randomly assigned variance units 1 through 3.

Reading and Mathematics Assessments

At the school-level for these samples, formation of preliminary variance strata did not pertain to certainty schools, since they are not subject to sampling variability, but only to noncertainty schools. The primary stratum for noncertainty schools was the highest school-level sampling stratum variable listed below, and the order of selection was defined by sort order on the school sampling frame.

- Trial Urban District Assessment (TUDA) districts, remainder of states (for states with TUDAs), or entire states for the public school samples at grades 4, 8, and 12; and
- Private school affiliation (Catholic, non-Catholic) for the private school samples at grades 4, 8, and 12.

At the student-level, all students were assigned to variance strata. The primary stratum was school, and the order of selection was defined by session number and position on the administration schedule.

Within each pair of preliminary variance strata, one first-stage unit, designated at random, was assigned as the first variance unit and the other first-stage unit as the second variance unit. Within each triplet preliminary variance stratum, the three schools were randomly assigned variance units 1 through 3.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/defining_variance_strata_and_forming_replicates_for_the_2013_assessment.aspx

# NAEP Technical DocumentationComputing School-Level Replicate Factors for the 2013 Assessment

The replicate variance estimation approach for the mathematics and reading assessments involved finite population corrections at the school level. The calculation of school-level replicate factors for these assessments depended upon whether or not a school was selected with certainty. For certainty schools, the school-level replicate factors for all replicates are set to unity – this is true regardless of whether or not the variance replication method uses finite population corrections – since certainty schools are not subject to sampling variability. Alternatively, one can view the finite population correction factor for such schools as being equal to zero. Thus, for each certainty school in a given assessment, the school-level replicate factor for each of the 62 replicates (r = 1, ..., 62) was assigned as follows:

$$SCH\_REPFAC_{js}(r) = 1$$

where $SCH\_REPFAC_{js}(r)$ is the school-level replicate factor for school s in primary stratum j for the r-th replicate.

For noncertainty schools, where preliminary variance strata were formed by grouping schools into pairs or triplets, school-level replicate factors were calculated for each of the 62 replicates based on this grouping. For schools in variance strata comprising pairs of schools, the school-level replicate factors,$SCH\_REPFAC_{js}(r)$,r = 1,..., 62, were calculated as follows:

$$SCH\_REPFAC_{js}(r) = \begin{cases} 1+\sqrt{\left(1-min(\pi_{j1},\pi_{j2})\right)}, & \text{for } js \in R_{jr}, U_{js}=1 \\ 1-\sqrt{\left(1-min(\pi_{j1},\pi_{j2})\right)}, & \text{for } js \in R_{jr}, U_{js}=2 \\ 1, & \text{for } js \notin R_{jr} \end{cases}$$

where

- $\min(\pi_{j1}, \pi_{j2})$ is the smallest school probability between the two schools comprising $R_{jr}$,
- $R_{jr}$ is the set of schools within the r-th variance stratum for primary stratum j, and
- $U_{js}$ is the variance unit (1 or 2) for school s in primary stratum j.

For noncertainty schools in preliminary variance strata comprising three schools, the school-level replicate factors $SCH\_REPFAC_{js}(r)$, r = 1,..., 62 were calculated as follows:

For school s from primary stratum j, variance stratum r,

$$SCH\_REPFAC_{js}(r) = \begin{cases} 1+\dfrac{\sqrt{\left(1-min(\pi_{j1},\pi_{j2},\pi_{j3}\right)}}{2}, & for\ js \in R_{jr},\ U_{js}=1 \\ 1+\dfrac{\sqrt{\left(1-min(\pi_{j1},\pi_{j2},\pi_{j3}\right)}}{2}, & for\ js \in R_{jr},\ U_{js}=2 \\ 1-\sqrt{\left(1-min(\pi_{j1},\pi_{j2},\pi_{j3}\right)}, & for\ js \in R_{jr},\ U_{js}=3 \end{cases}$$

while for r' = r + 31 (mod 62):

$$SCH\_REPFAC_{js}(r') = \begin{cases} 1+\dfrac{\sqrt{\left(1-min(\pi_{j1},\pi_{j2},\pi_{j3}\right)}}{2}, & \text{for } js \in R_{jr},\ U_{js}=1 \\ 1-\sqrt{\left(1-min(\pi_{j1},\pi_{j2},\pi_{j3}\right)}, & \text{for } js \in R_{jr},\ U_{js}=2 \\ 1+\dfrac{\sqrt{\left(1-min(\pi_{j1},\pi_{j2},\pi_{j3}\right)}}{2}, & \text{for } js \in R_{jr},\ U_{js}=3 \end{cases}$$

and for all other r* other than r and r':

$$SCH\_REPFAC_{js}(r^*) = 1;$$

where

- $\min(\pi_{j1}, \pi_{j2}, \pi_{j3})$ is the smallest school probability among the three schools comprising $R_{jr}$,
- $R_{jr}$ is the set of schools within the r-th variance stratum for primary stratum j, and
- $U_{js}$ is the variance unit (1, 2, or 3) for school s in primary stratum j.

In primary strata with fewer than 62 variance strata, the replicate weights for the "unused" variance strata (the remaining ones up to 62) for these schools were set equal to the school base weight (so that those replicates contribute nothing to the variance estimate).

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/computing_school_level_replicate_factors_for_the_2013_assessment_.aspx

---

# NAEP Technical DocumentationComputing Student-Level Replicate Factors for the 2013 Assessment

For the mathematics and reading assessments, which involved school-level finite population corrections, the student-level replication factors were calculated the same way regardless of whether or not the student was in a certainty school.

For students in student-level variance strata comprising pairs of students, the student-level replicate factors, $STU\_REPFAC_{jsk}(r)$, r = 1,..., 62, were calculated as follows:

$$STU\_REPFAC_{jsk}(r) = \begin{cases} 1+\sqrt{\pi_s}, & \text{for } jsk \in R_{jsr}, U_{jsk} = 1 \\ 1-\sqrt{\pi_s}, & \text{for } jsk \in R_{jsr}, U_{jsk} = 2 \\ 1, & \text{for } jsk \notin R_{jsr} \end{cases}$$

where

- $\pi_s$ is the probability of selection for school s,
- $R_{jsr}$ is the set of students within the r-th variance stratum for school s in primary stratum j, and
- $U_{jsk}$ is the variance unit (1 or 2) for student k in school s in stratum j.

For students in variance strata comprising three students, the student-level replicate factors $STU\_REPFAC_{jsk}(r)$, r = 1,..., 62, were calculated as follows:

$$STU\_REPFAC_{jsk}(r) = \begin{cases} 1+\frac{\sqrt{\pi_s}}{2}, & \text{for } jsk \in R_{jsr}, U_{jsk} = 1 \\ 1+\frac{\sqrt{\pi_s}}{2}, & \text{for } jsk \in R_{jsr}, U_{jsk} = 2 \\ 1-\sqrt{\pi_s}, & \text{for } jsk \in R_{jsr}, U_{jsk} = 3 \end{cases}$$

while for r' = r + 31 (mod 62):

$$STU\_REPFAC_{jsk}(r') = \begin{cases} 1+\frac{\sqrt{\pi_s}}{2}, & \text{for } jsk \in R_{jsr}, U_{jsk} = 1 \\ 1-\sqrt{\pi_s}, & \text{for } jsk \in R_{jsr}, U_{jsk} = 2 \\ 1+\frac{\sqrt{\pi_s}}{2}, & \text{for } jsk \in R_{jsr}, U_{jsk} = 3 \end{cases}$$

and for all other r* other than r and r':

$$STU\_REPFAC_{jsk}(r^*) = 1$$

where

- $\pi_s$ is the probability of selection for school s,
- $R_{jsr}$ is the set of students within the r-th replicate stratum for school s in stratum j, and
- $U_{jsk}$ is the variance unit (1, 2, or 3) for student k in school s in stratum j.

Note, for students in certainty schools, where $\pi_s = 1$, the student replicate factors are 2 and 0 in the case of pairs, and 1.5, 1.5, and 0 in the case of triples.

# NAEP Technical Documentation Replicate Variance Estimation for the 2013 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, and differences between these for different subgroups, as well as for more complex statistics such as linear or logistic regression coefficients.

In general, the paired jackknife replicate variance procedure involves initially 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 assigning, to one unit at random from the first variance stratum, a replicate weighting factor of less than 1.0, while assigning the remaining unit a complementary replicate factor greater than 1.0, and assigning all other units from the other (H - 1) strata a replicate factor of 1.0. This procedure is carried out for each variance stratum resulting in H replicates, each of which provides an estimate of the population total.

In general, this process is repeated for subsequent levels of sampling. In practice, this is not practicable for a design with three or more stages of sampling, and the marginal improvement in precision of the variance estimates would be negligible in all such cases in the NAEP setting. Thus in NAEP, when a two-stage design is used – sampling schools and then students – beginning in 2011 replication is carried out at both stages. (See Rizzo and Rust (2011) for a description of the methodology.) When a three-stage design is used, involving the selection of geographic Primary Sampling Units (PSUs), then schools, and then students, the replication procedure is only carried out at the first stage of sampling (the PSU stage for noncertainty PSUs, and the school stage within certainty PSUs). In this situation, the school and student variance components are correctly estimated, and the overstatement of the between-PSU variance component is relatively very small.

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

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

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 $v(\hat{\bar{Y}})$ 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

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

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

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,

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

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

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

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 (mod 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 $r$, where $\bar{y}_r(J)$ 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 replacing $\bar{y}_r$ with $\bar{y}_{r'}(J)$ for replicate $r'$, where $\bar{y}_{r'}(J)$ 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:

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

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/2013/replicate_variance_estimation_for_the_2013_assessment.aspx

---

NAEP Technical Documentation Website

# NAEP Technical Documentation Quality Control on Weighting Procedures for the 2013 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 participation, exclusion, and accommodation rates;
- checking demographic data of individual schools;
- comparisons with 2011 demographic data; and
- nonresponse bias analyses.

Final Participation, Exclusion, and Accommodation Rates

Nonresponse Bias Analyses

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 minority enrollment, median income (based on the school ZIP code area), and urban-centric locale. At the student level, the tabulations included race/ethnicity, gender, relative age, 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/2013/quality_control_on_weighting_procedures_for_the_2013_assessment.aspx

NAEP Technical Documentation Website

# NAEP Technical Documentation Final Participation, Exclusion, and Accommodation Rates for the 2013 Assessment

Final participation, exclusion, and accommodation rates are presented in quality control tables for each grade and subject by geographic domain and school type. School-level participation rates have been calculated according to National Center for Education Statistics (NCES) standards as they have been for previous assessments.

School-level participation rates were below 85 percent for private schools at all three grades (4, 8, and 12). Student-level participation rates were also below 85 percent for grade 12 public school student sample overall and in specific states: Connecticut, Florida, Illinois, Iowa, Massachusetts, New Hampshire, New Jersey, and West Virginia. As required by NCES standards, nonresponse bias analyses were conducted on each reporting group falling below the 85 percent participation threshold.

Grade 4 Mathematics
Grade 4 Reading

Grade 8 Mathematics
Grade 8 Reading

Grade 12 Mathematics
Grade 12 Reading

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/final_participation_exclusion_and_accommodation_rates_for_the_2013_assessment.aspx

# NAEP Technical Documentation Participation, Exclusion, and Accommodation Rates for Grade 4 Mathematics for the 2013 Assessment

The following table displays the school- and student-level response, exclusion, and accommodation rates for the grade 4 mathematics assessment by school type and jurisdiction. 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 sample schools only and do not reflect any effect of substitution. The rates weighted by the base weight and enrollment show the approximate proportion of the student population in the jurisdiction 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, grade 4 mathematics assessment, by school type and jurisdiction: 2013

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| All | 8,760 | 97.30 | 90.45 | 214,900 | 1.40 | 94.57 | 13.55 |
| National all[1] | 8,590 | 97.27 | 90.32 | 209,800 | 1.41 | 94.57 | 13.44 |
| Northeast all | 1,480 | 95.63 | 85.22 | 34,500 | 1.29 | 93.85 | 15.68 |
| Midwest all | 2,190 | 97.27 | 88.80 | 47,300 | 1.32 | 94.84 | 12.87 |
| South all | 2,740 | 98.20 | 93.44 | 73,600 | 1.37 | 94.71 | 14.38 |
| West all | 2,120 | 96.86 | 91.04 | 51,800 | 1.62 | 94.57 | 10.98 |
| National public | 8,060 | 99.69 | 99.54 | 202,700 | 1.52 | 94.49 | 14.22 |
| Alabama | 120 | 100.00 | 100.00 | 3,200 | 1.10 | 94.82 | 5.15 |
| Alaska | 200 | 99.48 | 96.56 | 3,100 | 1.14 | 93.18 | 21.85 |
| Arizona | 120 | 100.00 | 100.00 | 3,400 | 1.20 | 95.07 | 12.97 |
| Arkansas | 120 | 100.00 | 100.00 | 3,400 | 1.24 | 94.66 | 15.16 |
| California | 300 | 99.17 | 98.75 | 9,000 | 1.93 | 94.79 | 8.78 |
| Colorado | 120 | 100.00 | 100.00 | 3,400 | 1.15 | 92.34 | 12.11 |
| Connecticut | 120 | 97.22 | 97.25 | 3,200 | 1.36 | 93.85 | 15.52 |
| Delaware | 100 | 100.00 | 100.00 | 3,400 | 2.10 | 94.36 | 13.58 |
| District of Columbia | 140 | 100.00 | 100.00 | 2,300 | 1.37 | 95.09 | 17.59 |
| Florida | 240 | 100.00 | 100.00 | 6,900 | 1.84 | 94.11 | 20.24 |
| Georgia | 170 | 100.00 | 100.00 | 5,300 | 1.43 | 94.18 | 11.22 |
| Hawaii | 120 | 100.00 | 100.00 | 3,500 | 1.25 | 94.70 | 10.64 |
| Idaho | 130 | 100.00 | 100.00 | 3,500 | 1.29 | 95.24 | 9.58 |
| Illinois | 200 | 97.98 | 98.40 | 5,100 | 1.00 | 94.40 | 15.44 |
| Indiana | 120 | 100.00 | 100.00 | 3,300 | 1.52 | 95.18 | 17.03 |
| Iowa | 140 | 100.00 | 100.00 | 3,100 | 0.70 | 95.16 | 14.50 |
| Kansas | 150 | 100.00 | 100.00 | 3,400 | 1.62 | 94.79 | 15.16 |
| Kentucky | 160 | 100.00 | 100.00 | 4,700 | 1.45 | 94.67 | 11.30 |
| Louisiana | 130 | 100.00 | 100.00 | 3,300 | 1.08 | 94.49 | 18.38 |
| Maine | 160 | 100.00 | 100.00 | 3,400 | 2.11 | 93.95 | 17.44 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Mathematics Assessment.

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| Maryland | 170 | 100.00 | 100.00 | 4,700 | 0.99 | 94.22 | 17.30 |
| Massachusetts | 190 | 100.00 | 100.00 | 5,200 | 2.03 | 93.74 | 17.18 |
| Michigan | 190 | 100.00 | 100.00 | 4,600 | 1.96 | 94.14 | 11.02 |
| Minnesota | 130 | 100.00 | 100.00 | 3,500 | 1.37 | 94.85 | 10.62 |
| Mississippi | 120 | 100.00 | 100.00 | 3,300 | 0.76 | 95.44 | 6.73 |
| Missouri | 130 | 100.00 | 100.00 | 3,600 | 1.41 | 95.42 | 11.20 |
| Montana | 200 | 99.85 | 98.28 | 3,400 | 1.68 | 93.92 | 8.56 |
| Nebraska | 170 | 100.00 | 100.00 | 3,500 | 1.72 | 95.37 | 14.37 |
| Nevada | 120 | 100.00 | 100.00 | 3,500 | 1.41 | 95.75 | 22.90 |
| New Hampshire | 130 | 100.00 | 100.00 | 3,400 | 1.22 | 93.74 | 14.78 |
| New Jersey | 120 | 100.00 | 100.00 | 3,300 | 1.17 | 94.85 | 16.62 |
| New Mexico | 150 | 99.69 | 99.48 | 4,200 | 1.22 | 95.06 | 16.90 |
| New York | 160 | 98.84 | 96.79 | 4,500 | 1.23 | 92.27 | 20.02 |
| North Carolina | 160 | 100.00 | 100.00 | 4,800 | 1.24 | 94.19 | 14.17 |
| North Dakota | 270 | 99.86 | 99.19 | 3,700 | 2.56 | 95.57 | 9.78 |
| Ohio | 210 | 100.00 | 100.00 | 4,700 | 1.33 | 94.29 | 13.52 |
| Oklahoma | 140 | 100.00 | 100.00 | 3,600 | 1.85 | 94.35 | 13.95 |
| Oregon | 130 | 100.00 | 100.00 | 3,500 | 2.12 | 94.18 | 15.23 |
| Pennsylvania | 170 | 100.00 | 100.00 | 4,500 | 1.64 | 94.30 | 12.95 |
| Rhode Island | 120 | 100.00 | 100.00 | 3,400 | 1.12 | 94.98 | 15.17 |
| South Carolina | 120 | 100.00 | 100.00 | 3,200 | 1.08 | 96.08 | 11.87 |
| South Dakota | 190 | 100.00 | 100.00 | 3,400 | 1.42 | 95.36 | 10.56 |
| Tennessee | 120 | 100.00 | 100.00 | 3,400 | 1.34 | 94.21 | 13.54 |
| Texas | 310 | 100.00 | 100.00 | 9,200 | 1.65 | 95.36 | 17.92 |
| Utah | 120 | 99.08 | 99.32 | 3,600 | 1.25 | 94.79 | 12.66 |
| Vermont | 220 | 100.00 | 100.00 | 3,000 | 1.37 | 95.04 | 15.72 |
| Virginia | 110 | 100.00 | 100.00 | 3,300 | 1.51 | 94.35 | 13.07 |
| Washington | 120 | 99.09 | 99.35 | 3,600 | 2.17 | 93.50 | 14.12 |
| West Virginia | 150 | 100.00 | 100.00 | 3,200 | 1.71 | 94.77 | 10.03 |
| Wisconsin | 190 | 100.00 | 100.00 | 4,400 | 1.79 | 95.42 | 16.21 |
| Wyoming | 200 | 100.00 | 100.00 | 3,500 | 1.01 | 94.65 | 12.76 |
| DoDEA[2] | 120 | 99.23 | 98.08 | 3,700 | 1.66 | 95.05 | 12.20 |
| Trial Urban (TUDA) Districts and Other Jurisdictions | | | | | | | |
| Albuquerque | 50 | 100.00 | 100.00 | 1,700 | 1.15 | 94.71 | 20.47 |
| Atlanta | 60 | 100.00 | 100.00 | 2,000 | 0.98 | 95.42 | 9.76 |
| Austin | 60 | 100.00 | 100.00 | 1,700 | 2.04 | 93.69 | 30.80 |
| Baltimore City | 70 | 100.00 | 100.00 | 1,600 | 1.59 | 94.32 | 19.27 |
| Boston | 80 | 100.00 | 100.00 | 2,000 | 3.69 | 93.72 | 19.59 |
| Charlotte | 50 | 100.00 | 100.00 | 1,700 | 1.19 | 94.18 | 12.81 |
| Chicago | 100 | 100.00 | 100.00 | 2,500 | 1.07 | 94.85 | 19.30 |
| Cleveland | 90 | 100.00 | 100.00 | 1,500 | 4.26 | 93.62 | 22.29 |
| Dallas | 50 | 100.00 | 100.00 | 1,700 | 2.33 | 95.79 | 35.42 |
| Detroit | 70 | 100.00 | 100.00 | 1,300 | 4.88 | 90.92 | 14.80 |
| Fresno | 50 | 100.00 | 100.00 | 1,800 | 0.90 | 93.58 | 7.51 |
| Hillsborough | 60 | 100.00 | 100.00 | 1,700 | 1.17 | 95.74 | 23.30 |
| Houston | 80 | 100.00 | 100.00 | 2,600 | 1.88 | 96.62 | 27.25 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Mathematics Assessment.

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| Jefferson County, KY | 50 | 100.00 | 100.00 | 1,700 | 1.74 | 94.66 | 11.61 |
| Los Angeles | 80 | 100.00 | 100.00 | 2,500 | 1.96 | 95.80 | 9.83 |
| Miami | 90 | 100.00 | 100.00 | 2,300 | 2.35 | 95.07 | 28.05 |
| Milwaukee | 70 | 100.00 | 100.00 | 1,500 | 3.40 | 94.68 | 26.55 |
| New York City | 80 | 100.00 | 100.00 | 2,500 | 1.33 | 91.74 | 27.56 |
| Philadelphia | 60 | 100.00 | 100.00 | 1,600 | 3.45 | 94.71 | 15.82 |
| San Diego | 50 | 100.00 | 100.00 | 1,500 | 1.48 | 95.18 | 11.80 |
| District of Columbia (TUDA) | 90 | 100.00 | 100.00 | 1,500 | 1.97 | 95.52 | 18.06 |
| National private | 410 | 71.19 | 64.52 | 3,300 | 0.08 | 95.61 | 4.38 |
| Catholic | 130 | 88.65 | 89.70 | 1,700 | 0.06 | 95.60 | 4.95 |
| Non-Catholic private | 280 | 56.94 | 52.97 | 1,600 | 0.11 | 95.62 | 3.92 |
| Puerto Rico | 170 | 100.00 | 100.00 | 5,100 | 0.24 | 94.47 | 27.19 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Mathematics Assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/participation_exclusion_and_accommodation_rates_for_grade_4_mathematics_for_the_2013_assessment.aspx

# NAEP Technical Documentation Participation, Exclusion, and Accommodation Rates for Grade 4 Reading for the 2013 Assessment

The following table displays the school- and student-level response, exclusion, and accommodation rates for the grade 4 reading assessment by school type and jurisdiction. 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 sample schools only and do not reflect any effect of substitution. The rates weighted by the base weight and enrollment show the approximate proportion of the student population in the jurisdiction 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, grade 4 r eading assessment, by school type and jurisdiction: 2013

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| All | 8,590 | 97.27 | 90.32 | 216,400 | 2.52 | 94.78 | 12.17 |
| National all[1] | 8,590 | 97.27 | 90.32 | 216,400 | 2.52 | 94.78 | 12.17 |
| Northeast all | 1,480 | 95.63 | 85.22 | 35,600 | 1.72 | 93.97 | 15.30 |
| Midwest all | 2,190 | 97.27 | 88.80 | 48,700 | 2.01 | 95.04 | 12.22 |
| South all | 2,740 | 98.20 | 93.44 | 76,000 | 3.39 | 95.00 | 12.25 |
| West all | 2,120 | 96.86 | 91.04 | 53,500 | 2.13 | 94.71 | 9.92 |
| National public | 8,060 | 99.69 | 99.54 | 209,100 | 2.69 | 94.70 | 12.87 |
| Alabama | 120 | 100.00 | 100.00 | 3,400 | 1.14 | 95.49 | 5.39 |
| Alaska | 200 | 99.48 | 96.56 | 3,300 | 1.45 | 93.65 | 20.65 |
| Arizona | 120 | 100.00 | 100.00 | 3,500 | 1.08 | 95.46 | 13.24 |
| Arkansas | 120 | 100.00 | 100.00 | 3,600 | 1.11 | 95.16 | 15.34 |
| California | 300 | 99.17 | 98.75 | 9,300 | 2.50 | 94.88 | 7.73 |
| Colorado | 120 | 100.00 | 100.00 | 3,500 | 1.52 | 93.66 | 12.61 |
| Connecticut | 120 | 97.22 | 97.25 | 3,400 | 1.58 | 94.29 | 15.33 |
| Delaware | 100 | 100.00 | 100.00 | 3,500 | 4.70 | 94.34 | 10.38 |
| District of Columbia | 140 | 100.00 | 100.00 | 2,400 | 1.65 | 94.46 | 17.41 |
| Florida | 240 | 100.00 | 100.00 | 7,100 | 2.96 | 93.98 | 19.02 |
| Georgia | 170 | 100.00 | 100.00 | 5,400 | 4.90 | 95.34 | 8.13 |
| Hawaii | 120 | 100.00 | 100.00 | 3,600 | 1.80 | 93.97 | 10.48 |
| Idaho | 130 | 100.00 | 100.00 | 3,600 | 1.49 | 94.99 | 9.32 |
| Illinois | 200 | 97.98 | 98.40 | 5,200 | 1.24 | 95.13 | 14.76 |
| Indiana | 120 | 100.00 | 100.00 | 3,500 | 2.43 | 94.40 | 16.31 |
| Iowa | 140 | 100.00 | 100.00 | 3,200 | 1.08 | 95.11 | 14.42 |
| Kansas | 150 | 100.00 | 100.00 | 3,500 | 1.82 | 95.07 | 13.41 |
| Kentucky | 160 | 100.00 | 100.00 | 4,800 | 2.99 | 94.97 | 9.74 |
| Louisiana | 130 | 100.00 | 100.00 | 3,400 | 1.16 | 94.73 | 18.61 |
| Maine | 160 | 100.00 | 100.00 | 3,500 | 1.69 | 93.65 | 17.87 |
| Maryland | 170 | 100.00 | 100.00 | 4,900 | 12.86 | 94.40 | 5.70 |
| Massachusetts | 190 | 100.00 | 100.00 | 5,300 | 2.66 | 93.77 | 15.53 |
| Michigan | 190 | 100.00 | 100.00 | 4,800 | 3.81 | 94.64 | 9.66 |
| Minnesota | 130 | 100.00 | 100.00 | 3,600 | 2.71 | 94.93 | 9.61 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Reading Assessment.

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| Mississippi | 120 | 100.00 | 100.00 | 3,400 | 0.53 | 94.99 | 6.85 |
| Missouri | 130 | 100.00 | 100.00 | 3,700 | 1.23 | 95.26 | 11.16 |
| Montana | 200 | 99.85 | 98.28 | 3,500 | 2.86 | 94.40 | 7.33 |
| Nebraska | 170 | 100.00 | 100.00 | 3,600 | 3.57 | 95.83 | 14.26 |
| Nevada | 120 | 100.00 | 100.00 | 3,700 | 1.50 | 95.10 | 22.73 |
| New Hampshire | 130 | 100.00 | 100.00 | 3,500 | 2.56 | 93.45 | 13.48 |
| New Jersey | 120 | 100.00 | 100.00 | 3,400 | 1.72 | 94.87 | 15.27 |
| New Mexico | 150 | 99.69 | 99.48 | 4,300 | 1.02 | 94.55 | 15.04 |
| New York | 160 | 98.84 | 96.79 | 4,600 | 1.35 | 93.06 | 20.15 |
| North Carolina | 160 | 100.00 | 100.00 | 5,000 | 1.80 | 94.88 | 13.06 |
| North Dakota | 270 | 99.86 | 99.19 | 3,800 | 4.06 | 96.28 | 8.73 |
| Ohio | 210 | 100.00 | 100.00 | 4,800 | 2.61 | 94.58 | 12.80 |
| Oklahoma | 140 | 100.00 | 100.00 | 3,700 | 1.72 | 94.58 | 14.35 |
| Oregon | 130 | 100.00 | 100.00 | 3,700 | 2.49 | 93.98 | 12.20 |
| Pennsylvania | 170 | 100.00 | 100.00 | 4,600 | 2.29 | 94.42 | 12.53 |
| Rhode Island | 120 | 100.00 | 100.00 | 3,500 | 1.34 | 94.78 | 14.43 |
| South Carolina | 120 | 100.00 | 100.00 | 3,300 | 1.73 | 94.64 | 9.74 |
| South Dakota | 190 | 100.00 | 100.00 | 3,500 | 2.22 | 95.69 | 9.26 |
| Tennessee | 120 | 100.00 | 100.00 | 3,500 | 3.10 | 95.34 | 12.29 |
| Texas | 310 | 100.00 | 100.00 | 9,500 | 4.90 | 95.50 | 14.40 |
| Utah | 120 | 99.08 | 99.32 | 3,700 | 3.05 | 93.71 | 10.29 |
| Vermont | 220 | 100.00 | 100.00 | 3,100 | 1.17 | 95.05 | 15.65 |
| Virginia | 110 | 100.00 | 100.00 | 3,400 | 1.54 | 94.93 | 12.21 |
| Washington | 120 | 99.09 | 99.35 | 3,700 | 2.81 | 93.71 | 12.45 |
| West Virginia | 150 | 100.00 | 100.00 | 3,300 | 1.78 | 93.62 | 8.89 |
| Wisconsin | 190 | 100.00 | 100.00 | 4,500 | 1.61 | 94.97 | 16.63 |
| Wyoming | 200 | 100.00 | 100.00 | 3,600 | 1.25 | 94.38 | 13.00 |
| DoDEA[2] | 120 | 99.23 | 98.08 | 3,800 | 5.95 | 95.48 | 7.39 |
| Trial Urban (TUDA) Districts and Other Jurisdictions | | | | | | | |
| Albuquerque | 50 | 100.00 | 100.00 | 1,800 | 0.74 | 93.43 | 17.51 |
| Atlanta | 60 | 100.00 | 100.00 | 2,000 | 1.12 | 95.96 | 9.39 |
| Austin | 60 | 100.00 | 100.00 | 1,700 | 3.90 | 94.12 | 27.06 |
| Baltimore City | 70 | 100.00 | 100.00 | 1,700 | 15.85 | 93.62 | 4.33 |
| Boston | 80 | 100.00 | 100.00 | 2,000 | 4.33 | 94.03 | 17.64 |
| Charlotte | 50 | 100.00 | 100.00 | 1,700 | 0.90 | 94.49 | 11.72 |
| Chicago | 100 | 100.00 | 100.00 | 2,600 | 1.45 | 94.58 | 18.56 |
| Cleveland | 90 | 100.00 | 100.00 | 1,500 | 4.70 | 94.08 | 22.22 |
| Dallas | 50 | 100.00 | 100.00 | 1,700 | 17.11 | 96.08 | 24.30 |
| Detroit | 70 | 100.00 | 100.00 | 1,300 | 5.51 | 92.09 | 13.44 |
| Fresno | 50 | 100.00 | 100.00 | 1,800 | 2.36 | 94.94 | 6.04 |
| Hillsborough | 60 | 100.00 | 100.00 | 1,800 | 1.07 | 94.92 | 23.00 |
| Houston | 80 | 100.00 | 100.00 | 2,700 | 6.41 | 96.63 | 23.90 |
| Jefferson County, KY | 50 | 100.00 | 100.00 | 1,800 | 5.28 | 95.03 | 7.56 |
| Los Angeles | 80 | 100.00 | 100.00 | 2,500 | 2.10 | 94.63 | 10.75 |
| Miami | 90 | 100.00 | 100.00 | 2,400 | 4.51 | 95.37 | 26.36 |
| Milwaukee | 70 | 100.00 | 100.00 | 1,500 | 4.08 | 93.65 | 25.71 |
| New York City | 80 | 100.00 | 100.00 | 2,500 | 1.62 | 92.44 | 27.13 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Reading Assessment.

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| Philadelphia | 60 | 100.00 | 100.00 | 1,600 | 3.83 | 94.61 | 15.31 |
| San Diego | 50 | 100.00 | 100.00 | 1,600 | 2.32 | 94.74 | 10.45 |
| District of Columbia (TUDA) | 90 | 100.00 | 100.00 | 1,600 | 2.26 | 94.50 | 17.21 |
| National private | 410 | 71.19 | 64.52 | 3,400 | 0.53 | 95.85 | 4.05 |
| Catholic | 130 | 88.65 | 89.70 | 1,700 | 0.23 | 95.75 | 3.84 |
| Non-Catholic private | 280 | 56.94 | 52.97 | 1,600 | 0.79 | 95.96 | 4.22 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Reading Assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/participation_exclusion_and_accommodation_rates_for_grade_4_reading_for_the_2013_assessment.aspx

# NAEP Technical Documentation Participation, Exclusion, and Accommodation Rates for Grade 8 Mathematics for the 2013 Assessment

The following table displays the school- and student-level response, exclusion, and accommodation rates for the grade 8 mathematics assessment by school type and jurisdiction. 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 sample schools only and do not reflect any effect of substitution. The rates weighted by the base weight and enrollment show the approximate proportion of the student population in the jurisdiction 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, grade 8 mathematics assessment, by school type and jurisdiction: 2013

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| All | 7,370 | 96.97 | 84.74 | 201,500 | 1.47 | 93.14 | 11.88 |
| National all[1] | 7,240 | 96.94 | 84.59 | 195,600 | 1.48 | 93.15 | 11.79 |
| Northeast all | 1,160 | 93.53 | 75.06 | 32,700 | 1.60 | 92.00 | 15.85 |
| Midwest all | 1,920 | 97.62 | 85.21 | 44,100 | 1.42 | 93.69 | 11.78 |
| South all | 2,380 | 97.75 | 86.70 | 68,800 | 1.51 | 93.24 | 11.59 |
| West all | 1,720 | 97.42 | 89.08 | 48,000 | 1.41 | 93.28 | 9.25 |
| National public | 6,760 | 99.48 | 99.61 | 189,400 | 1.59 | 93.02 | 12.25 |
| Alabama | 110 | 100.00 | 100.00 | 3,000 | 1.04 | 94.23 | 5.14 |
| Alaska | 150 | 99.91 | 98.79 | 3,000 | 1.08 | 91.72 | 18.75 |
| Arizona | 120 | 99.03 | 99.16 | 3,200 | 1.30 | 93.42 | 10.71 |
| Arkansas | 110 | 100.00 | 100.00 | 3,200 | 1.93 | 95.00 | 13.92 |
| California | 260 | 100.00 | 100.00 | 8,400 | 1.49 | 93.59 | 7.91 |
| Colorado | 120 | 100.00 | 100.00 | 3,100 | 1.12 | 93.47 | 11.50 |
| Connecticut | 110 | 98.00 | 97.87 | 3,100 | 2.05 | 92.44 | 13.92 |
| Delaware | 70 | 100.00 | 100.00 | 3,200 | 1.31 | 90.65 | 14.90 |
| District of Columbia | 90 | 100.00 | 100.00 | 2,100 | 0.96 | 91.26 | 20.71 |
| Florida | 230 | 100.00 | 100.00 | 6,400 | 1.70 | 91.06 | 15.32 |
| Georgia | 130 | 100.00 | 100.00 | 4,800 | 1.55 | 93.38 | 9.82 |
| Hawaii | 60 | 100.00 | 100.00 | 3,200 | 1.67 | 90.26 | 12.28 |
| Idaho | 100 | 100.00 | 100.00 | 3,100 | 1.06 | 94.15 | 8.42 |
| Illinois | 190 | 100.00 | 100.00 | 4,800 | 1.01 | 94.48 | 13.83 |
| Indiana | 110 | 97.06 | 96.65 | 3,000 | 1.64 | 92.49 | 13.95 |
| Iowa | 120 | 100.00 | 100.00 | 3,100 | 0.77 | 93.74 | 13.28 |
| Kansas | 130 | 100.00 | 100.00 | 3,300 | 1.67 | 93.94 | 11.23 |
| Kentucky | 140 | 99.04 | 99.21 | 4,300 | 2.08 | 94.54 | 10.09 |
| Louisiana | 150 | 100.00 | 100.00 | 3,200 | 1.06 | 94.14 | 14.26 |
| Maine | 120 | 100.00 | 100.00 | 2,900 | 1.33 | 92.79 | 15.99 |
| Maryland | 160 | 100.00 | 100.00 | 4,400 | 1.74 | 92.08 | 13.33 |
| Massachusetts | 140 | 100.00 | 100.00 | 4,800 | 2.01 | 91.98 | 16.11 |
| Michigan | 170 | 100.00 | 100.00 | 4,200 | 2.46 | 92.93 | 10.55 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Mathematics Assessment.

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| Minnesota | 130 | 98.99 | 99.67 | 2,900 | 1.70 | 91.58 | 9.16 |
| Mississippi | 110 | 100.00 | 100.00 | 3,200 | 0.80 | 93.80 | 6.51 |
| Missouri | 130 | 100.00 | 100.00 | 3,100 | 1.28 | 94.25 | 10.57 |
| Montana | 150 | 99.80 | 98.82 | 3,200 | 1.44 | 92.28 | 9.20 |
| Nebraska | 130 | 100.00 | 100.00 | 3,100 | 1.85 | 93.41 | 12.02 |
| Nevada | 90 | 100.00 | 100.00 | 3,300 | 1.04 | 92.80 | 11.91 |
| New Hampshire | 90 | 100.00 | 100.00 | 3,200 | 1.06 | 91.60 | 15.99 |
| New Jersey | 110 | 100.00 | 100.00 | 3,100 | 1.64 | 92.26 | 16.38 |
| New Mexico | 120 | 99.68 | 99.02 | 4,000 | 1.57 | 93.07 | 12.00 |
| New York | 160 | 93.08 | 95.81 | 4,300 | 1.90 | 91.15 | 19.38 |
| North Carolina | 140 | 100.00 | 100.00 | 4,500 | 1.29 | 92.95 | 13.74 |
| North Dakota | 190 | 99.92 | 99.44 | 3,700 | 2.93 | 94.98 | 11.44 |
| Ohio | 200 | 100.00 | 100.00 | 4,500 | 1.51 | 93.07 | 13.54 |
| Oklahoma | 130 | 100.00 | 100.00 | 3,100 | 1.63 | 92.97 | 14.09 |
| Oregon | 130 | 100.00 | 100.00 | 3,100 | 1.47 | 92.91 | 10.88 |
| Pennsylvania | 160 | 100.00 | 100.00 | 4,300 | 1.70 | 92.17 | 14.66 |
| Rhode Island | 60 | 100.00 | 100.00 | 3,200 | 1.11 | 93.93 | 15.92 |
| South Carolina | 110 | 100.00 | 100.00 | 3,200 | 1.33 | 94.19 | 9.86 |
| South Dakota | 150 | 100.00 | 100.00 | 3,200 | 1.30 | 94.44 | 8.66 |
| Tennessee | 110 | 100.00 | 100.00 | 3,200 | 1.77 | 92.81 | 9.81 |
| Texas | 230 | 100.00 | 100.00 | 8,800 | 1.92 | 93.82 | 12.13 |
| Utah | 120 | 100.00 | 100.00 | 3,300 | 1.53 | 92.07 | 10.15 |
| Vermont | 120 | 100.00 | 100.00 | 3,000 | 0.83 | 93.91 | 15.36 |
| Virginia | 110 | 100.00 | 100.00 | 3,200 | 1.05 | 93.39 | 12.18 |
| Washington | 120 | 100.00 | 100.00 | 3,100 | 2.03 | 90.87 | 11.47 |
| West Virginia | 110 | 100.00 | 100.00 | 3,200 | 1.69 | 92.62 | 9.02 |
| Wisconsin | 170 | 100.00 | 100.00 | 4,300 | 1.51 | 94.25 | 14.73 |
| Wyoming | 100 | 100.00 | 100.00 | 3,300 | 1.50 | 93.66 | 12.51 |
| DoDEA[2] | 70 | 99.40 | 96.83 | 2,600 | 1.15 | 94.47 | 9.23 |
| Trial Urban (TUDA) Districts and Other Jurisdictions | | | | | | | |
| Albuquerque | 30 | 100.00 | 100.00 | 1,400 | 1.53 | 90.76 | 14.44 |
| Atlanta | 30 | 100.00 | 100.00 | 1,600 | 0.72 | 91.57 | 11.10 |
| Austin | 30 | 100.00 | 100.00 | 1,600 | 1.88 | 90.97 | 20.60 |
| Baltimore City | 60 | 100.00 | 100.00 | 1,300 | 1.70 | 89.54 | 19.73 |
| Boston | 40 | 100.00 | 100.00 | 1,800 | 2.55 | 91.61 | 20.88 |
| Charlotte | 40 | 100.00 | 100.00 | 1,500 | 1.29 | 90.94 | 10.11 |
| Chicago | 100 | 100.00 | 100.00 | 2,300 | 1.28 | 94.80 | 17.19 |
| Cleveland | 90 | 100.00 | 100.00 | 1,500 | 2.62 | 91.57 | 28.48 |
| Dallas | 40 | 100.00 | 100.00 | 1,600 | 2.44 | 93.81 | 18.35 |
| Detroit | 50 | 100.00 | 100.00 | 1,100 | 4.29 | 91.58 | 15.07 |
| Fresno | 20 | 100.00 | 100.00 | 1,400 | 1.74 | 92.52 | 7.06 |
| Hillsborough | 50 | 100.00 | 100.00 | 1,600 | 1.35 | 93.78 | 20.46 |
| Houston | 50 | 100.00 | 100.00 | 2,400 | 2.21 | 92.37 | 14.67 |
| Jefferson County, KY | 40 | 100.00 | 100.00 | 1,600 | 1.65 | 93.37 | 12.72 |
| Los Angeles | 70 | 100.00 | 100.00 | 2,200 | 1.54 | 94.39 | 10.83 |
| Miami | 80 | 100.00 | 100.00 | 2,300 | 2.25 | 92.63 | 18.78 |
| Milwaukee | 60 | 100.00 | 100.00 | 1,500 | 4.10 | 91.60 | 25.55 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Mathematics Assessment.

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| New York City | 90 | 99.00 | 97.58 | 2,400 | 1.72 | 91.78 | 26.10 |
| Philadelphia | 50 | 100.00 | 100.00 | 1,400 | 3.74 | 92.67 | 20.69 |
| San Diego | 30 | 100.00 | 100.00 | 1,300 | 2.32 | 92.60 | 11.81 |
| District of Columbia (TUDA) | 40 | 100.00 | 100.00 | 1,100 | 1.69 | 90.15 | 22.20 |
| National private | 400 | 69.63 | 60.45 | 3,400 | 0.26 | 94.74 | 6.54 |
| Catholic | 130 | 87.18 | 84.76 | 1,800 | 0.26 | 95.73 | 5.50 |
| Non-Catholic private | 270 | 53.51 | 48.11 | 1,600 | 0.26 | 93.50 | 7.51 |
| Puerto Rico | 130 | 100.00 | 100.00 | 5,900 | 0.03 | 92.75 | 23.05 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Mathematics Assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/participation_exclusion_and_accommodation_rates_for_grade_8_mathematics_for_the_2013_assessment.aspx

# NAEP Technical Documentation Participation, Exclusion, and Accommodation Rates for Grade 8 Reading for the 2013 Assessment

The following table displays the school- and student-level response, exclusion, and accommodation rates for the grade 8 reading assessment by school type and jurisdiction. 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 sample schools only and do not reflect any effect of substitution. The rates weighted by the base weight and enrollment show the approximate proportion of the student population in the jurisdiction 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, grade 8 r eading assessment, by school type and jurisdiction: 2013

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| All | 7,240 | 96.94 | 84.59 | 199,100 | 2.15 | 93.11 | 10.76 |
| National all[1] | 7,240 | 96.94 | 84.59 | 199,100 | 2.15 | 93.11 | 10.76 |
| Northeast all | 1,160 | 93.53 | 75.06 | 33,300 | 1.55 | 91.80 | 15.53 |
| Midwest all | 1,920 | 97.62 | 85.21 | 45,100 | 1.93 | 93.48 | 11.08 |
| South all | 2,380 | 97.75 | 86.70 | 69,900 | 2.60 | 93.39 | 9.99 |
| West all | 1,720 | 97.42 | 89.08 | 48,900 | 2.08 | 93.21 | 8.32 |
| National public | 6,760 | 99.48 | 99.61 | 192,900 | 2.32 | 92.93 | 11.16 |
| Alabama | 110 | 100.00 | 100.00 | 3,100 | 1.14 | 94.26 | 4.83 |
| Alaska | 150 | 99.91 | 98.79 | 3,100 | 1.40 | 91.91 | 18.39 |
| Arizona | 120 | 99.03 | 99.16 | 3,300 | 1.47 | 93.67 | 9.67 |
| Arkansas | 110 | 100.00 | 100.00 | 3,200 | 1.96 | 93.21 | 13.36 |
| California | 260 | 100.00 | 100.00 | 8,500 | 2.52 | 93.42 | 6.74 |
| Colorado | 120 | 100.00 | 100.00 | 3,200 | 1.15 | 93.46 | 10.89 |
| Connecticut | 110 | 98.00 | 97.87 | 3,100 | 2.13 | 91.38 | 13.88 |
| Delaware | 70 | 100.00 | 100.00 | 3,200 | 3.49 | 91.59 | 12.23 |
| District of Columbia | 90 | 100.00 | 100.00 | 2,100 | 1.82 | 91.33 | 19.57 |
| Florida | 230 | 100.00 | 100.00 | 6,500 | 1.86 | 91.72 | 15.15 |
| Georgia | 130 | 100.00 | 100.00 | 4,900 | 3.80 | 93.67 | 8.18 |
| Hawaii | 60 | 100.00 | 100.00 | 3,300 | 1.93 | 90.58 | 12.33 |
| Idaho | 100 | 100.00 | 100.00 | 3,200 | 1.61 | 93.64 | 7.76 |
| Illinois | 190 | 100.00 | 100.00 | 4,900 | 1.44 | 93.76 | 12.94 |
| Indiana | 110 | 97.06 | 96.65 | 3,100 | 1.90 | 93.12 | 13.75 |
| Iowa | 120 | 100.00 | 100.00 | 3,100 | 1.27 | 93.44 | 12.16 |
| Kansas | 130 | 100.00 | 100.00 | 3,300 | 1.72 | 93.42 | 11.72 |
| Kentucky | 140 | 99.04 | 99.21 | 4,300 | 3.28 | 93.93 | 8.47 |
| Louisiana | 150 | 100.00 | 100.00 | 3,300 | 1.24 | 93.78 | 14.15 |
| Maine | 120 | 100.00 | 100.00 | 3,000 | 1.55 | 92.34 | 15.16 |
| Maryland | 160 | 100.00 | 100.00 | 4,400 | 9.41 | 93.77 | 5.45 |
| Massachusetts | 140 | 100.00 | 100.00 | 4,900 | 2.15 | 91.82 | 15.04 |
| Michigan | 170 | 100.00 | 100.00 | 4,300 | 3.53 | 93.66 | 9.68 |
| Minnesota | 130 | 98.99 | 99.67 | 3,000 | 2.33 | 91.30 | 8.43 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Reading Assessment.

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| Mississippi | 110 | 100.00 | 100.00 | 3,200 | 0.70 | 93.72 | 6.55 |
| Missouri | 130 | 100.00 | 100.00 | 3,100 | 1.02 | 92.55 | 10.62 |
| Montana | 150 | 99.80 | 98.82 | 3,200 | 2.29 | 91.61 | 7.51 |
| Nebraska | 130 | 100.00 | 100.00 | 3,200 | 2.99 | 92.32 | 10.14 |
| Nevada | 90 | 100.00 | 100.00 | 3,400 | 1.00 | 92.19 | 10.91 |
| New Hampshire | 90 | 100.00 | 100.00 | 3,200 | 2.93 | 91.46 | 14.28 |
| New Jersey | 110 | 100.00 | 100.00 | 3,200 | 2.64 | 92.01 | 14.78 |
| New Mexico | 120 | 99.68 | 99.02 | 4,000 | 1.70 | 93.39 | 10.00 |
| New York | 160 | 93.08 | 95.81 | 4,400 | 0.96 | 90.46 | 20.03 |
| North Carolina | 140 | 100.00 | 100.00 | 4,600 | 1.72 | 92.51 | 12.29 |
| North Dakota | 190 | 99.92 | 99.44 | 3,800 | 4.30 | 94.07 | 9.52 |
| Ohio | 200 | 100.00 | 100.00 | 4,600 | 2.22 | 93.08 | 13.08 |
| Oklahoma | 130 | 100.00 | 100.00 | 3,200 | 1.39 | 93.43 | 12.42 |
| Oregon | 130 | 100.00 | 100.00 | 3,200 | 1.45 | 92.62 | 11.30 |
| Pennsylvania | 160 | 100.00 | 100.00 | 4,300 | 1.78 | 91.94 | 14.51 |
| Rhode Island | 60 | 100.00 | 100.00 | 3,300 | 1.37 | 92.96 | 15.18 |
| South Carolina | 110 | 100.00 | 100.00 | 3,200 | 1.88 | 94.03 | 7.48 |
| South Dakota | 150 | 100.00 | 100.00 | 3,300 | 2.95 | 95.01 | 6.02 |
| Tennessee | 110 | 100.00 | 100.00 | 3,200 | 3.13 | 93.54 | 7.75 |
| Texas | 230 | 100.00 | 100.00 | 8,900 | 3.51 | 93.78 | 10.05 |
| Utah | 120 | 100.00 | 100.00 | 3,400 | 3.05 | 93.00 | 8.36 |
| Vermont | 120 | 100.00 | 100.00 | 3,100 | 0.92 | 92.93 | 15.08 |
| Virginia | 110 | 100.00 | 100.00 | 3,300 | 1.40 | 92.97 | 10.56 |
| Washington | 120 | 100.00 | 100.00 | 3,200 | 2.46 | 91.22 | 9.78 |
| West Virginia | 110 | 100.00 | 100.00 | 3,200 | 1.82 | 93.10 | 7.60 |
| Wisconsin | 170 | 100.00 | 100.00 | 4,400 | 1.61 | 94.11 | 14.45 |
| Wyoming | 100 | 100.00 | 100.00 | 3,400 | 1.14 | 93.15 | 12.27 |
| DoDEA[2] | 70 | 99.40 | 96.83 | 2,600 | 3.84 | 94.13 | 7.11 |
| Trial Urban (TUDA) Districts and Other Jurisdictions | | | | | | | |
| Albuquerque | 30 | 100.00 | 100.00 | 1,400 | 2.04 | 93.46 | 11.79 |
| Atlanta | 30 | 100.00 | 100.00 | 1,700 | 1.02 | 92.20 | 10.98 |
| Austin | 30 | 100.00 | 100.00 | 1,600 | 3.35 | 88.54 | 18.36 |
| Baltimore City | 60 | 100.00 | 100.00 | 1,300 | 16.39 | 89.73 | 5.14 |
| Boston | 40 | 100.00 | 100.00 | 1,800 | 3.41 | 93.05 | 18.94 |
| Charlotte | 40 | 100.00 | 100.00 | 1,500 | 1.68 | 92.20 | 9.90 |
| Chicago | 100 | 100.00 | 100.00 | 2,300 | 1.60 | 94.72 | 16.76 |
| Cleveland | 90 | 100.00 | 100.00 | 1,500 | 3.52 | 91.90 | 27.75 |
| Dallas | 40 | 100.00 | 100.00 | 1,600 | 3.51 | 93.98 | 15.20 |
| Detroit | 50 | 100.00 | 100.00 | 1,100 | 5.74 | 91.37 | 12.53 |
| Fresno | 20 | 100.00 | 100.00 | 1,500 | 3.10 | 93.27 | 5.86 |
| Hillsborough | 50 | 100.00 | 100.00 | 1,600 | 1.94 | 91.85 | 19.74 |
| Houston | 50 | 100.00 | 100.00 | 2,400 | 3.80 | 93.58 | 12.29 |
| Jefferson County, KY | 40 | 100.00 | 100.00 | 1,600 | 4.30 | 94.71 | 9.49 |
| Los Angeles | 70 | 100.00 | 100.00 | 2,300 | 2.70 | 94.30 | 9.97 |
| Miami | 80 | 100.00 | 100.00 | 2,400 | 2.88 | 94.21 | 18.45 |
| Milwaukee | 60 | 100.00 | 100.00 | 1,500 | 4.06 | 93.15 | 25.08 |
| New York City | 90 | 99.00 | 97.58 | 2,400 | 1.46 | 91.17 | 26.00 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Reading Assessment.

| School type and jurisdiction | Number of schools in original sample, rounded | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled, rounded | Weighted percent of students excluded | Weighted student participation rates (percent) after makeups | Weighted percent of students accommodated |
|---|---|---|---|---|---|---|---|
| Philadelphia | 50 | 100.00 | 100.00 | 1,400 | 3.79 | 91.35 | 20.91 |
| San Diego | 30 | 100.00 | 100.00 | 1,300 | 2.58 | 93.78 | 10.58 |
| District of Columbia (TUDA) | 40 | 100.00 | 100.00 | 1,100 | 2.53 | 90.18 | 22.13 |
| National private | 400 | 69.63 | 60.45 | 3,500 | 0.30 | 95.45 | 6.32 |
| Catholic | 130 | 87.18 | 84.76 | 1,900 | 0.21 | 96.07 | 4.96 |
| Non-Catholic private | 270 | 53.51 | 48.11 | 1,600 | 0.39 | 94.67 | 7.56 |

[1] Includes national public, national private, and Bureau of Indian Education schools located in the United States and all Department of Defense Education Activity schools, but not schools in Puerto Rico.

[2] Department of Defense Education Activity schools.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Reading Assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/participation_exclusion_and_accommodation_rates_for_grade_8_reading_for_the_2013_assessment.aspx

# NAEP Technical Documentation Participation, Exclusion, and Accommodation Rates for Grade 12 Mathematics for the 2013 Assessment

The following table displays the school- and student-level response, exclusion, and accommodation rates for the grade 12 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 sample schools only and do not reflect any effect of substitution. The rates weighted by the base weight and enrollment show the approximate proportion of the student population in the jurisdiction 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, grade 12 mathematics assessment, by school type and geographic r egion: 2013

| School type and geographic region | Number of schools in original sample | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled | Weighted percentage of students excluded | Weighted student participation rates (percent) after makeups | Weighted percentage of students accommodated |
|---|---|---|---|---|---|---|---|
| All | 2,200 | 89.51 | 82.66 | 62,200 | 2.16 | 84.33 | 8.65 |
| National all[1] | 2,200 | 89.51 | 82.66 | 62,200 | 2.16 | 84.33 | 8.65 |
| Northeast all | 510 | 89.05 | 81.63 | 16,200 | 2.29 | 81.79 | 11.95 |
| Midwest all | 650 | 87.14 | 83.20 | 16,600 | 1.65 | 83.87 | 8.61 |
| South all | 710 | 89.42 | 85.99 | 20,300 | 2.31 | 86.52 | 7.98 |
| West all | 330 | 92.21 | 77.24 | 9,100 | 2.32 | 83.37 | 7.15 |
| National public | 2,030 | 92.95 | 93.31 | 60,400 | 2.31 | 84.17 | 8.77 |
| Arkansas | 100 | 100.00 | 100.00 | 2,900 | 2.78 | 92.09 | 8.61 |
| Connecticut | 110 | 98.93 | 99.45 | 3,200 | 1.76 | 81.22 | 8.71 |
| Florida | 120 | 99.05 | 99.30 | 3,300 | 3.21 | 77.25 | 12.67 |
| Idaho | 100 | 100.00 | 100.00 | 3,000 | 1.65 | 89.17 | 6.72 |
| Illinois | 130 | 90.38 | 93.98 | 3,300 | 1.85 | 85.16 | 9.79 |
| Iowa | 120 | 100.00 | 100.00 | 3,300 | 1.13 | 83.05 | 10.78 |
| Massachusetts | 110 | 99.04 | 99.45 | 3,200 | 2.21 | 81.71 | 11.13 |
| Michigan | 140 | 100.00 | 100.00 | 4,000 | 1.90 | 86.94 | 8.78 |
| New Hampshire | 80 | 100.00 | 100.00 | 4,100 | 1.61 | 76.64 | 11.22 |
| New Jersey | 110 | 98.14 | 98.57 | 3,300 | 1.89 | 84.10 | 14.28 |
| South Dakota | 140 | 99.74 | 99.07 | 3,100 | 1.51 | 87.48 | 5.78 |
| Tennessee | 130 | 100.00 | 100.00 | 4,100 | 2.51 | 88.15 | 7.84 |
| West Virginia | 90 | 100.00 | 100.00 | 3,300 | 2.00 | 83.68 | 7.01 |
| Remaining jurisdictions[2] | 570 | 91.16 | 90.91 | 16,200 | 2.26 | 84.41 | 10.55 |
| National private | 160 | 53.34 | 55.43 | 1,800 | 0.63 | 86.51 | 7.32 |
| Catholic | 40 | 68.06 | 79.95 | 1,000 | 0.83 | 85.53 | 5.46 |
| Non-Catholic private | 120 | 38.52 | 50.25 | 800 | 0.42 | 87.96 | 9.28 |

[1] Includes national public, national private, Bureau of Indian Education, and Department of Defense Education Activity schools located in the United States.

[2] Includes national public schools not part of the state assessment.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Mathematics Assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/participation_exclusion_and_accommodation_rates_for_grade_12_mathematics_for_the_2013_assessment.aspx

# NAEP Technical DocumentationParticipation, Exclusion, and Accommodation Rates for Grade 12 Reading for the 2013 Assessment

The following table displays the school- and student-level response, exclusion, and accommodation rates for the grade 12 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 sample schools only and do not reflect any effect of substitution. The rates weighted by the base weight and enrollment show the approximate proportion of the student population in the jurisdiction 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, grade 12 r eading assessment, by school type and geographic r egion: 2013

| School type and geographic region | Number of schools in original sample | School participation rates (percent) before substitution (weighted by base weight and enrollment) | School participation rates (percent) before substitution (weighted by base weight only) | Number of students sampled | Weighted percentage of students excluded | Weighted student participation rates (percent) after makeups | Weighted percentage of students accommodated |
|---|---|---|---|---|---|---|---|
| All | 2,200 | 89.51 | 82.66 | 62,300 | 2.41 | 83.89 | 8.55 |
| National all[1] | 2,200 | 89.51 | 82.66 | 62,300 | 2.41 | 83.89 | 8.55 |
| Northeast all | 510 | 89.05 | 81.63 | 16,500 | 2.16 | 80.91 | 12.89 |
| Midwest all | 650 | 87.14 | 83.20 | 16,700 | 2.05 | 84.05 | 8.75 |
| South all | 710 | 89.42 | 85.99 | 20,000 | 2.87 | 85.51 | 7.18 |
| West all | 330 | 92.21 | 77.24 | 9,000 | 2.24 | 83.58 | 7.14 |
| National public | 2,030 | 92.95 | 93.31 | 60,400 | 2.56 | 83.77 | 8.73 |
| Arkansas | 100 | 100.00 | 100.00 | 3,000 | 2.56 | 90.21 | 8.24 |
| Connecticut | 110 | 98.93 | 99.45 | 3,400 | 2.34 | 79.77 | 8.70 |
| Florida | 120 | 99.05 | 99.30 | 3,300 | 3.55 | 77.34 | 12.14 |
| Idaho | 100 | 100.00 | 100.00 | 3,200 | 1.66 | 88.68 | 6.42 |
| Illinois | 130 | 90.38 | 93.98 | 3,400 | 2.29 | 83.72 | 9.92 |
| Iowa | 120 | 100.00 | 100.00 | 3,500 | 1.51 | 84.26 | 10.62 |
| Massachusetts | 110 | 99.04 | 99.45 | 3,200 | 1.87 | 79.84 | 11.31 |
| Michigan | 140 | 100.00 | 100.00 | 3,900 | 4.01 | 87.21 | 6.17 |
| New Hampshire | 80 | 100.00 | 100.00 | 4,300 | 2.55 | 76.91 | 10.25 |
| New Jersey | 110 | 98.14 | 98.57 | 3,300 | 1.80 | 84.67 | 14.78 |
| South Dakota | 140 | 99.74 | 99.07 | 3,300 | 1.60 | 86.17 | 5.16 |
| Tennessee | 130 | 100.00 | 100.00 | 3,900 | 2.88 | 88.82 | 7.13 |
| West Virginia | 90 | 100.00 | 100.00 | 3,400 | 2.37 | 84.28 | 6.89 |
| Remaining jurisdictions[2] | 570 | 91.16 | 90.91 | 15,200 | 2.77 | 83.98 | 10.05 |
| National private | 160 | 53.34 | 55.43 | 1,900 | 0.84 | 85.52 | 6.67 |
| Catholic | 40 | 68.06 | 79.95 | 1,100 | 0.92 | 84.67 | 4.01 |
| Non-Catholic | 120 | 38.52 | 50.25 | 800 | 0.75 | 86.75 | 9.41 |

[1] Includes national public, national private, Bureau of Indian Education, and Department of Defense Education Activity schools located in the United States.

[2] Includes national public schools not part of the state assessment.

NOTE: Numbers of schools are rounded to nearest ten, and numbers of students are rounded to nearest hundred. 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), 2013 Reading Assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/participation_exclusion_and_accommodation_rates_for_grade_12_reading_for_the_2013_assessment.aspx

# NAEP Technical Documentation Nonresponse Bias Analyses for the 2013 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 2013 assessment indicated a need for school nonresponse bias analyses for private school samples in grades 4, 8, and 12 (operational subjects). Student nonresponse bias analyses were necessary for the grade 12 public school student sample overall and in specific states, for both reading and mathematics: Connecticut, Florida, Iowa, Massachusetts, New Hampshire, and West Virginia. Additionally, a student nonresponse bias analysis was required for the grade 12 public school student sample in Illinois based on the weighted response rate for reading, while such an analysis was required for grade 12 public school student sample in New Jersey based on the weighted response rate for mathematics. Thus, three separate school-level analyses and nine separate student-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 2013 NRBA report (657.56 KB).

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), and race/ethnicity percentages (mean).

Based on the school characteristics available, for the private school samples at grade 4, 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 grade 8 and 12 private school samples. For grade 8, this result is evidently related to the fact that, among non-Catholic schools, larger schools were less likely to respond. Thus, when making adjustments to address the underrepresentation of non-Catholic schools among the respondents, the result is to over represent smaller schools at the expense of larger ones. The limited school sample sizes involved means that it is not possible to make adjustments that account fully for all school characteristics. For grade 12, the analyses suggested potential bias for percentage Asian and percentage Two or more races. Please see the full report for more details.

Each student-level analysis was conducted in two parts. The first part of the analysis examined the potential for nonresponse bias that was introduced through student nonresponse. The second part of the analysis examined the potential for bias after accounting for the effects of nonresponse weight adjustments. The characteristics examined were gender, race/ethnicity, relative age, National School Lunch Program eligibility, student disability (SD) status, and English language learner (ELL) status.

Based on the student characteristics available, there does not appear to be evidence of substantial potential bias resulting from student nonresponse. Please see the full report for more details.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2013/nonresponse_bias_analyses_for_the_2013_assessment.aspx

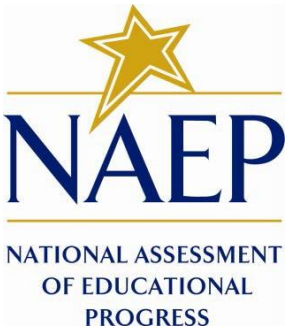
# *NATIONAL CENTER FOR EDUCATION STATISTICS NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS*

## *National Assessment of Education Progress (NAEP) 2019 and 2020*

### *Appendix B2*

### *NAEP 2012 Long Term Trend (LTT) Weighting Procedures Design*

*OMB# 1850-0928 v.14*




February 2019

## 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

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

---

## 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.

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

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

Trimming of School Base Weights

Trimming of Student Weights

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- $\delta$), 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 $\delta$ 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

**Main QC Findings of Interest**

Participation, Exclusion, and Accommodation Rates

Nonresponse Bias Analysis

- 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.

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.

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.

Age 9 Mathematics
Age 9 Reading

Age 13 Mathematics
Age 13 Reading

Age 17 Mathematics
Age 17 Reading

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.

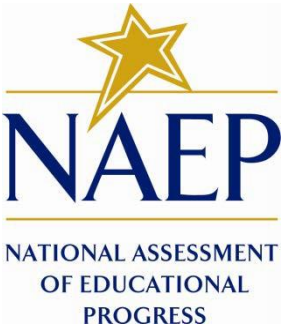
http://nces.ed.gov/nationsreportcard/tdw/weighting/2012/2012_weighting_nonresponse_bias_analysis.asp x

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

*National Assessment of Educational Progress (NAEP)*
*2019 and 2020*

# *Appendix C1*
# *2019 Sampling Memo*

*OMB# 1850-0928 v.14*




September 2018
No changes since v.10

**Date:** February 28, 2018

**Memo:** 2019-1.1A/1.1B/1.1D/1.1E

**To:**

| | |
|---|---|
| William Ward, NCES | Chris Averett |
| Ed Kulick, ETS | Kavemuii Murangi |
| David Freund, ETS | Erin Wiley |
| Amy Dresher, ETS | Dwight Brock |
| Cathy White, Pearson | David Hubble |
| Saira Brenner, Fulcrum | Yiting Dai |
| Dianne Walsh | Jing Kang |
| Lauren Byrne | Sabrina Zhang |
| Lisa Rodriguez | Leslie Wallace |
| Rick Rogers | Natalia Weil |
| Rob Dymowski | Greg Binzer |
| William Wall | |

**From:** Amy Lin, John Burke, and Lloyd Hicks

**Reviewer:** Keith Rust

**Subject:** Sample Design for 2019 NAEP - DRAFT

# I. Introduction

For 2019, the NAEP assessment involves the following components:

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

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

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

D. Pilot tests in reading, mathematics, and vocabulary at grades 4 and 8.

Below is a summary list of the features of the 2019 sample design.

1. The alpha samples for grades 4 and 8 public, and the delta samples for private schools at grades 4 and 8, will be used for the operational assessments in reading and mathematics.

2. The beta public school samples and the epsilon private school samples at grades 4 and 8 will be used for the national science assessments and the various pilot tests. The beta and epsilon samples at grade 12 will be used for the operational reading, mathematics, and science assessments.

3. As in recent NAEP studies, each Trial Urban District Assessment (TUDA) sample will form part of the corresponding state sample, and each state sample will form part of the national sample. There are twenty-seven Trial Urban District Assessment (TUDA) participants. These are the same districts that participated in 2017.

4. Schools in the alpha and delta samples will be assessed using DBA with tablets. Schools in the beta and epsilon samples will receive a mixture of DBA assessments, using tablets, and pencil and paper (PBA) assessments.

5. All BIE schools and students will be included in the operational samples at grades 4 and 8. This is because, after a hiatus in 2017, the National Indian Education Study (NIES) is resuming. Having all BIE students in sample is designed to provide detailed national results for American Indian and Alaskan Native (AIAN) students in reading and mathematics, as part of the NIES.

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

7. As in 2017, the Department of Defense Schools are expected to be reported as a single jurisdiction (DoDEA).

8. At grade 12, there will be no state-level samples.

9. Oversampling of private schools at grades 4 and 8 will be done at the same level as 2017. Response rates permitting, this will allow separate reporting for reading and mathematics for Catholic and non-Catholic schools at grades 4 and 8, but no further breakdowns by private school type.

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

11. In the beta samples, there will be moderate oversampling of schools with moderate to high proportions of Black, Hispanic, and American Indian and Alaska Native students.

Table 1. Target sample sizes of assessed students, and expected number of participating schools, for 2019 NAEP

| | Spiral | Jurisdictions | | Students | | |
|---|---|---|---|---|---|---|
| | Spiral Indic. | States (incl. DC, DoDEA) | Urban districts | Public school students | Private school students | Total |
| Grade 4 | | | | | | |
| Nat'l/state reading (DBA) | DS | 52 | 27 | 176,000 | 3,700 | 179,000 |
| Nat'l/state math (DBA) | DS | 52 | 27 | 144,000 | 3,000 | 147,000 |
| Puerto Rico (DBA) | DP | 1 | | 3,000 | 0 | 3,000 |
| **Total - alpha** | | | | **323,000** | | **323,000** |
| **Total- delta** | | | | | **6,700** | **6,700** |
| Typical max. no. students/school | | | | 50 | 50 | |
| Average assessed students/school | | | | 40 | 25 | |
| **Total schools - alpha, delta** | | | | **8,075** | **268** | **8,343** |
| | | | | | | |
| Science (DBA) | DA | | | 17,100 | 1,900 | 19,000 |
| Science (PBA) | PA | | | 8,100 | 900 | 9,000 |
| Math Pilot | DA | | | 10,350 | 1,150 | 11,500 |
| Reading Pilot | DA | | | 4,050 | 450 | 4,500 |
| Vocabulary initial-Pilot | DA | | | 1,980 | 220 | 2,200 |
| **Total - beta** | | | | **41,580** | | **41,580** |
| **Total - epsilon** | | | | | **4,620** | **4,620** |
| Typical max. no. students/school | | | | 62 | 62 | |
| Average assessed students/school | | | | 50 | 25 | |
| **Total schools - beta, epsilon** | | | | **832** | **185** | **1,017** |
| | | | | | | |
| **Total number of students grade 4** | | | | **364,580** | **11,320** | **375,900** |
| **Total number of schools grade 4** | | | | **8,907** | **453** | **9,360** |

Table 1. Target sample sizes of assessed students, and expected number of participating schools, for 2019 NAEP (Continued)

| | Spiral | Jurisdictions | | Students | | |
|---|---|---|---|---|---|---|
| | Spiral Indic. | States (incl. DC, DoDEA) | Urban districts | Public school students | Private school students | Total |
| **Grade 8** | | | | | | |
| Nat'l/state reading (DBA) | DS, DT | 52 | 27 | 176,000 | 3,700 | 179,000 |
| Nat'l/state math (DBA) | DS, DT | 52 | 27 | 144,000 | 3,000 | 147,000 |
| Puerto Rico (DBA) | DP | 1 | | 3,000 | 0 | 3,000 |
| **Total - alpha** | | | | **323,000** | | **323,000** |
| **Total- delta** | | | | | **6,700** | **6,700** |
| Typical max. no. students/school | | | | 50 | 50 | |
| Average assessed students/school | | | | 47 | 25 | |
| **Total schools - alpha, delta** | | | | **6,870** | **268** | **7,138** |
| | | | | | | |
| Science (DBA) | DA | | | 17,100 | 1,900 | 19,000 |
| Science (PBA) | PA | | | 9,000 | 1,000 | 10,000 |
| Math Pilot | DA | | | 10,350 | 1,150 | 11,500 |
| Reading pilot | DA | | | 4,050 | 450 | 4,500 |
| Vocabulary initial-Pilot | DA | | | 1,980 | 220 | 2,200 |
| **Total – beta** | | | | **42,480** | | **42,480** |
| **Total – epsilon** | | | | | **4,720** | **4,720** |
| Typical max. no. students/school | | | | **63** | **63** | |
| Average assessed students/school | | | | 52 | 25 | |
| **Total schools - beta, epsilon** | | | | **817** | **189** | **1,006** |
| | | | | | | |
| **Total number of students grade 8** | | | | **365,480** | **11,420** | **376,900** |
| **Total number of schools grade 8** | | | | **7,687** | **457** | **8,144** |

Table 1. Target sample sizes of assessed students, and expected number of participating schools, for 2019 NAEP (Continued)

| | Spiral | Jurisdictions | | Students | | |
|---|---|---|---|---|---|---|
| | Spiral Indic. | States (incl. DC, DoDEA) | Urban districts | Public School students | Private school students | Total |
| **Grade 12** | | | | | | |
| Reading (DBA) | DA | | | 13,500 | 1,500 | 15,000 |
| Reading (PBA) | PA | | | 11,700 | 1,300 | 13,000 |
| Math (DBA) | DA | | | 12,600 | 1,400 | 14,000 |
| Math (PBA) | PA | | | 12,600 | 1,400 | 14,000 |
| Science (DBA) | DA | | | 17,100 | 1,900 | 19,000 |
| Science (PBA) | PA | | | 9,900 | 1,100 | 11,000 |
| **Total - beta** | | | | **77,400** | | **77,400** |
| **Total- epsilon** | | | | | **8,600** | **8,600** |
| Typical max. no. students/school | | | | 68 | 68 | |
| Average assessed students/school | | | | 50 | 40 | |
| **Total schools – beta, epsilon** | | | | **1,548** | **215** | **1,763** |
| | | | | | | |
| **Total number of students grade 12** | | | | **77,400** | **8,600** | **86,000** |
| **Total number of schools grade 12** | | | | **1,548** | **215** | **1,763** |
| | | | | | | |
| | | | | | | |
| GRAND TOTAL STUDENTS | | | | 807,460 | 31,340 | 838,800 |
| GRAND TOTAL SCHOOLS | | | | **18,142** | **1,125** | **19,267** |

# II. Assessment Types

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

Table 2. NAEP 2019 assessment types and IDs

| ID | Type | Subjects | Grades | Schools | Comments |
|---|---|---|---|---|---|
| DS | Operational DBA | Reading, math (22:27) | 4, 8 | Public, Private | All schools in the alpha (except Puerto Rico) and delta samples. |
| DA | Operational, and pilot DBA | Science, reading, math, vocabulary (190:45:115:22) | 4, 8 | Public, Private | All schools in the beta and epsilon samples. |
| PA | Operational | Science | 4, 8 | Public, Private | All schools in the beta and epsilon samples. |
| DA | Operational | Reading, math, science, (15:14:19) | 12 | Public, Private | All schools in the beta and epsilon samples. |
| PA | Operational | Reading, math, science (13:14:11) | 12 | Public, Private | All schools in the beta and epsilon samples. |
| DP | Operational | Mathematics | 4, 8 | Public | Puerto Rico alpha samples |

# III. Sample Types and Sizes

In similar fashion to past years (but somewhat different), we will identify four different types of school samples: Alpha, Beta, Delta, and Epsilon. These distinguish sets of schools that will be conducting distinct portions of the assessment.

## 1. Alpha Samples at Grades 4 and 8

These are public school samples for grades 4 and 8. They will be used for the operational state-by-state assessments in reading and mathematics, and contribute to the national samples for these subjects as well. There will be alpha samples for each state, DC, DoDEA, BIE, and Puerto Rico.

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

A. At each grade, the target student sample size is 5,700 per state. The goal in each state (before considering the contribution of TUDA districts) is to roughly assess 2,700 student for math and 2,200 students for reading. The DS session type will be used.

B. There will be samples for twenty-seven TUDA districts. For the six large TUDA districts (New York, Los Angeles, Chicago, Miami-Dade, Clark Co., and Houston) the assessed student target sample sizes are three-quarters the size of a state sample (3,675). The target sample size after considering attrition is 4,275.

C. For the remaining 21 TUDA districts, the assessed student target sample sizes are half the size of a state sample (2,450). The target sample size after inflation to account for attrition is 2,850.

D. Note that, above, there is a conflict between sample size requirements at the state level, and the TUDA district level. This will be resolved as in previous years: the districts will have the target samples indicated in B and C, and reflected in Table 3. For the states that contain one or more of these districts, the target sample size indicated in A (and shown in Table 3) will be used to determine a school sampling rate for the state, which will be applied to the balance of the state outside the TUDA district(s). Thus the target student sample sizes, shown in Table 3, for states that contain a TUDA district, are only 'design targets', and are smaller than the final total sample size for the state, but larger than the sample for the balance of the state, exclusive of its TUDA districts. In the case of the District of Columbia, the state sample size requirement is that all schools and students be included. This renders moot any requirements for the DC TUDA sample, which by default consists of all schools operated by the DCPS district (but excludes charter schools in DC, even though those are all included in the state sample, as these are not operated by DCPS).

E. In Puerto Rico, the target sample size is 4,000 per grade (grades 4 and 8), with the goal of assessing 3,000 students. Under normal circumstances this target would be set at 3,500, but because of the rapid and substantial shifts in the school population in Puerto Rico, this has been increased to provide some insurance against attrition due to closed schools and declining enrollments.

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

As mentioned above, the NAEP 2019 design includes an oversample of high proportion American Indian schools in certain states (as part of the NIES design). These schools will be sampled at higher rates than the other schools. The NIES oversample will take place in Arizona, Minnesota, North Carolina, Oregon, Utah, Washington, and Wisconsin. Schools with relatively large percentages of American Indian students will be separately stratified, as explained below, and oversampled by factors ranging from 3 to 6 based on state and grade. Table 3 below shows the thresholds used to define the NIES oversampling strata along with their corresponding oversampling factors.

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

| State | Grade 4 | | Grade 8 | |
|---|---|---|---|---|
| | Percent American Indian thresholds | Oversampling factor | Percent American Indian thresholds | Oversampling factor |
| Arizona | 50 | 4 | 50 | 3 |
| Minnesota | 10 | 5 | 10 | 5 |
| North Carolina | 10 | 6 | 10 | 6 |
| Oregon | 10 | 6 | 10 | 6 |
| Utah | 5 | 6 | 5 | 6 |
| Washington | 10 | 6 | 10 | 6 |
| Wisconsin | 10 | 6 | 10 | 6 |

Table 4 shows the target student sample sizes, and the approximate counts of schools to be selected in the alpha samples, along with the school and student frame counts, by state and TUDA districts for grades 4 and 8. The table also identifies the jurisdictions where we take all schools and where we take all students. Note that the additional sample that will result from NIES oversampling is not included in this table.

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

Table 4. Grade 4 and 8 school and student frame counts, expected school sample sizes, and initial target student sample sizes for the 2019 state-by-state and TUDA district assessments (Alpha samples)

| Jurisdiction | Grade 4 | | | | | Grade 8 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Schools in frame | Schools in sample | Students in frame | Overall target student sample size | | Schools in frame | Schools in sample | Students in frame | Overall target student sample size | |
| Alabama | 709 | 120 | 57,548 | 5,700 | | 456 | 118 | 55,820 | 5,700 | |
| Alaska | 352 | 185 | 9,361 | 5,700 | | 270 | 131 | 9,019 | 5,700 | |
| Arizona | 1,193 | 123 | 86,472 | 5,700 | | 793 | 122 | 83,469 | 5,700 | |
| Arkansas | 480 | 121 | 36,937 | 5,700 | | 303 | 114 | 36,503 | 5,700 | |
| Bureau of Indian Education | 137 | 137 | 3,357 | 3,357 | ** | 113 | 113 | 2,936 | 2,936 | ** |
| California | 5,979 | 119 | 471,633 | 5,700 | | 2,933 | 120 | 455,487 | 5,700 | |
| Colorado | 1,054 | 123 | 67,814 | 5,700 | | 567 | 121 | 65,088 | 5,700 | |
| Connecticut | 602 | 121 | 39,544 | 5,700 | | 339 | 118 | 40,679 | 5,700 | |
| Delaware | 119 | 99 | 10,393 | 5,700 | | 61 | 61 | 10,105 | 5,700 | * |
| District of Columbia | 119 | 119 | 5,536 | 5,536 | ** | 69 | 69 | 4,520 | 4,520 | ** |
| DoDEA Schools | 110 | 95 | 7,547 | 5,700 | | 65 | 65 | 5,629 | 5,629 | ** |
| Florida | 2,225 | 118 | 212,520 | 5,700 | | 1,219 | 119 | 202,235 | 5,700 | |
| Georgia | 1,248 | 115 | 133,243 | 5,700 | | 562 | 115 | 129,475 | 5,700 | |
| Hawaii | 205 | 118 | 15,494 | 5,700 | | 83 | 62 | 13,314 | 5,700 | |
| Idaho | 381 | 128 | 22,864 | 5,700 | | 209 | 100 | 22,319 | 5,700 | |
| Illinois | 2,205 | 124 | 149,235 | 5,700 | | 1,561 | 123 | 151,830 | 5,700 | |
| Indiana | 1,050 | 119 | 78,837 | 5,700 | | 489 | 116 | 79,653 | 5,700 | |
| Iowa | 638 | 128 | 37,147 | 5,700 | | 368 | 118 | 35,691 | 5,700 | |
| Kansas | 704 | 132 | 37,202 | 5,700 | | 393 | 125 | 36,033 | 5,700 | |
| Kentucky | 721 | 120 | 52,221 | 5,700 | | 417 | 121 | 50,755 | 5,700 | |
| Louisiana | 760 | 121 | 55,735 | 5,700 | | 488 | 120 | 51,981 | 5,700 | |
| Maine | 320 | 147 | 13,444 | 5,700 | | 202 | 112 | 13,473 | 5,700 | |
| Maryland | 903 | 119 | 67,399 | 5,700 | | 373 | 117 | 61,983 | 5,700 | |
| Massachusetts | 958 | 120 | 70,968 | 5,700 | | 485 | 116 | 71,662 | 5,700 | |
| Michigan | 1,711 | 123 | 111,240 | 5,700 | | 1,083 | 123 | 114,211 | 5,700 | |
| Minnesota | 956 | 126 | 65,262 | 5,700 | | 712 | 128 | 63,732 | 5,700 | |
| Mississippi | 423 | 118 | 38,316 | 5,700 | | 287 | 112 | 36,486 | 5,700 | |
| Missouri | 1,166 | 129 | 69,574 | 5,700 | | 709 | 127 | 67,833 | 5,700 | |

Table 4. Grade 4 and 8 school and student frame counts, expected school sample sizes, and initial target student sample sizes for the 2019 state-by-state and TUDA district assessments (Alpha samples) (Continued)

| Jurisdiction | Grade 4 | | | | | Grade 8 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Schools in frame | Schools in sample | Students in frame | Overall target student sample size | | Schools in frame | Schools in sample | Students in frame | Overall target student sample size | |
| Montana | 392 | 174 | 11,534 | 5,700 | | 271 | 136 | 10,811 | 5,700 | |
| Nebraska | 532 | 146 | 23,315 | 5,700 | | 294 | 114 | 22,561 | 5,700 | |
| Nevada | 394 | 119 | 35,875 | 5,700 | | 171 | 91 | 34,346 | 5,700 | |
| New Hampshire | 270 | 135 | 13,734 | 5,700 | | 142 | 89 | 14,078 | 5,700 | |
| New Jersey | 1,371 | 120 | 99,697 | 5,700 | | 765 | 118 | 99,117 | 5,700 | |
| New Mexico | 444 | 128 | 26,208 | 5,700 | | 232 | 110 | 25,079 | 5,700 | |
| New York | 2,471 | 118 | 201,226 | 5,700 | | 1,498 | 117 | 196,197 | 5,700 | |
| North Carolina | 1,457 | 119 | 118,118 | 5,700 | | 728 | 117 | 117,176 | 5,700 | |
| North Dakota | 261 | 166 | 8,471 | 5,700 | | 184 | 142 | 7,789 | 5,700 | |
| Ohio | 1,740 | 121 | 129,087 | 5,700 | | 1,093 | 119 | 131,562 | 5,700 | |
| Oklahoma | 869 | 132 | 50,988 | 5,700 | | 583 | 127 | 48,784 | 5,700 | |
| Oregon | 746 | 128 | 43,589 | 5,700 | | 428 | 124 | 42,824 | 5,700 | |
| Pennsylvania | 1,607 | 118 | 130,442 | 5,700 | | 888 | 116 | 131,525 | 5,700 | |
| Puerto Rico | 931 | 169 | 31,308 | 4,000 | | 398 | 161 | 30,211 | 4,000 | |
| Rhode Island | 164 | 111 | 10,777 | 5,700 | | 60 | 60 | 10,720 | 5,700 | * |
| South Carolina | 643 | 118 | 57,878 | 5,700 | | 306 | 115 | 54,617 | 5,700 | |
| South Dakota | 312 | 163 | 10,517 | 5,700 | | 246 | 135 | 9,657 | 5,700 | |
| Tennessee | 995 | 120 | 77,202 | 5,700 | | 584 | 119 | 73,441 | 5,700 | |
| Texas | 4,431 | 118 | 399,283 | 5,700 | | 2,251 | 119 | 383,849 | 5,700 | |
| Utah | 621 | 118 | 50,010 | 5,700 | | 256 | 113 | 47,320 | 5,700 | |
| Vermont | 216 | 216 | 6,204 | 6,204 | ** | 121 | 121 | 5,999 | 5,999 | ** |
| Virginia | 1,109 | 117 | 97,550 | 5,700 | | 379 | 114 | 95,187 | 5,700 | |
| Washington | 1,231 | 122 | 81,904 | 5,700 | | 609 | 122 | 79,084 | 5,700 | |
| West Virginia | 417 | 138 | 20,578 | 5,700 | | 190 | 110 | 20,464 | 5,700 | |
| Wisconsin | 1,099 | 128 | 61,686 | 5,700 | | 649 | 123 | 61,152 | 5,700 | |
| Wyoming | 192 | 137 | 7,639 | 5,700 | | 89 | 89 | 7,042 | 5,700 | * |

Table 4. Grade 4 and 8 school and student frame counts, expected school sample sizes, and initial target student sample sizes for the 2019 state-by-state and TUDA district assessments (Alpha samples) (Continued)

| Jurisdiction | Grade 4 | | | | | Grade 8 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Schools in frame | Schools in sample | Students in frame | Overall target student sample size | | Schools in frame | Schools in sample | Students in frame | Overall target student sample size | |
| Albuquerque | 95 | 57 | 7,412 | 2,850 | | 40 | 40 | 6,691 | 2,850 | * |
| Atlanta | 55 | 55 | 4,285 | 2,850 | * | 23 | 23 | 3,554 | 3,554 | ** |
| Austin | 80 | 56 | 6,867 | 2,850 | | 22 | 22 | 5,427 | 2,850 | * |
| Baltimore City | 128 | 64 | 6,716 | 2,850 | | 96 | 62 | 5,504 | 2,850 | |
| Boston | 72 | 57 | 4,086 | 2,850 | | 43 | 43 | 3,667 | 3,667 | ** |
| Charlotte | 105 | 57 | 11,696 | 2,850 | | 46 | 35 | 11,007 | 2,850 | |
| Chicago | 433 | 93 | 27,360 | 4,275 | | 434 | 93 | 27,895 | 4,275 | |
| Clark County, NV | 226 | 87 | 25,311 | 4,275 | | 80 | 58 | 24,676 | 4,275 | |
| Cleveland | 71 | 71 | 2,754 | 2,754 | ** | 70 | 70 | 2,685 | 2,685 | ** |
| Dallas | 151 | 58 | 13,325 | 2,850 | | 41 | 41 | 10,873 | 2,850 | * |
| Denver | 102 | 59 | 7,108 | 2,850 | | 60 | 47 | 6,060 | 2,850 | |
| Detroit | 65 | 55 | 3,889 | 2,850 | | 49 | 49 | 2,963 | 2,963 | ** |
| Duval County, FL | 119 | 58 | 10,313 | 2,850 | | 50 | 35 | 8,873 | 2,850 | |
| Fresno | 68 | 55 | 5,788 | 2,850 | | 19 | 19 | 5,147 | 2,850 | * |
| Fort Worth | 85 | 57 | 7,073 | 2,850 | | 32 | 32 | 5,977 | 2,850 | * |
| Guilford County, NC | 74 | 56 | 5,492 | 2,850 | | 29 | 29 | 5,339 | 2,850 | * |
| Hillsborough County, FL | 176 | 58 | 16,522 | 2,850 | | 87 | 50 | 15,096 | 2,850 | |
| Houston | 174 | 86 | 17,729 | 4,275 | | 61 | 49 | 13,063 | 4,275 | |
| Jefferson County, KY | 100 | 59 | 7,718 | 2,850 | | 43 | 29 | 7,306 | 2,850 | |
| Los Angeles | 496 | 87 | 45,361 | 4,275 | | 122 | 75 | 36,142 | 4,275 | |
| Miami | 285 | 88 | 26,690 | 4,275 | | 177 | 82 | 26,957 | 4,275 | |
| Milwaukee | 111 | 65 | 5,668 | 2,850 | | 83 | 56 | 4,977 | 2,850 | |
| New York City | 788 | 88 | 73,248 | 4,275 | | 524 | 88 | 66,513 | 4,275 | |
| Philadelphia | 148 | 58 | 11,227 | 2,850 | | 112 | 54 | 8,849 | 2,850 | |
| San Diego | 120 | 59 | 9,125 | 2,850 | | 38 | 38 | 7,433 | 2,850 | * |
| Shelby County, TN | 120 | 59 | 9,250 | 2,850 | | 61 | 44 | 8,277 | 2,850 | |
| District of Columbia PS | 76 | 76 | 3,584 | 3,584 | ** | 32 | 32 | 2,394 | 2,394 | ** |

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

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

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

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

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

| Jurisdiction | Grade 4 | | | | | Grade 8 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Schools in frame | Schools in sample | Students in frame | Overall target student sample size | | Schools in frame | Schools in sample | Students in frame | Overall target student sample size | |
| Alabama | 709 | 120 | 57,548 | 5,700 | | 456 | 117 | 55,820 | 5,700 | |
| Alaska | 352 | 184 | 9,361 | 5,700 | | 270 | 131 | 9,019 | 5,700 | |
| Arizona | 1,193 | 123 | 86,472 | 5,700 | | 793 | 123 | 83,469 | 5,700 | |
| Arkansas | 480 | 121 | 36,937 | 5,700 | | 303 | 114 | 36,503 | 5,700 | |
| Bureau Of Indian Education | 137 | 137 | 3,357 | 3,357 | ** | 113 | 113 | 2,936 | 2,936 | ** |
| California | 5,979 | 305 | 471,633 | 14,945 | | 2,933 | 240 | 455,487 | 15,064 | |
| Colorado | 1,054 | 169 | 67,814 | 7,950 | | 567 | 157 | 65,088 | 8,018 | |
| Connecticut | 602 | 121 | 39,544 | 5,700 | | 339 | 118 | 40,679 | 5,700 | |
| Delaware | 119 | 99 | 10,393 | 5,700 | | 61 | 61 | 10,105 | 5,700 | * |
| District Of Columbia | 119 | 119 | 5,536 | 5,536 | ** | 69 | 69 | 4,520 | 4,520 | ** |
| DoDEA Schools | 110 | 95 | 7,547 | 5,700 | | 65 | 65 | 5,629 | 5,629 | ** |
| Florida | 2,225 | 293 | 212,520 | 14,238 | | 1,219 | 256 | 202,235 | 14,238 | |
| Georgia | 1,248 | 166 | 133,243 | 8,367 | | 562 | 135 | 129,475 | 9,098 | |
| Hawaii | 205 | 118 | 15,494 | 5,700 | | 83 | 61 | 13,314 | 5,700 | |
| Idaho | 381 | 128 | 22,864 | 5,700 | | 209 | 100 | 22,319 | 5,700 | |
| Illinois | 2,205 | 194 | 149,235 | 8,927 | | 1,561 | 194 | 151,830 | 8,924 | |
| Indiana | 1,050 | 119 | 78,837 | 5,700 | | 489 | 116 | 79,653 | 5,700 | |
| Iowa | 638 | 128 | 37,147 | 5,700 | | 368 | 118 | 35,691 | 5,700 | |
| Kansas | 704 | 132 | 37,202 | 5,700 | | 393 | 125 | 36,033 | 5,700 | |
| Kentucky | 721 | 162 | 52,221 | 7,709 | | 417 | 133 | 50,755 | 7,730 | |
| Louisiana | 760 | 121 | 55,735 | 5,700 | | 488 | 120 | 51,981 | 5,700 | |
| Maine | 320 | 147 | 13,444 | 5,700 | | 202 | 112 | 13,473 | 5,700 | |
| Maryland | 903 | 169 | 67,399 | 7,983 | | 373 | 167 | 61,983 | 8,044 | |
| Massachusetts | 958 | 170 | 70,968 | 8,222 | | 485 | 153 | 71,662 | 9,076 | |
| Michigan | 1,711 | 174 | 111,240 | 8,350 | | 1,083 | 168 | 114,211 | 8,515 | |
| Minnesota | 956 | 126 | 65,262 | 5,700 | | 712 | 128 | 63,732 | 5,700 | |
| Mississippi | 423 | 118 | 38,316 | 5,700 | | 287 | 112 | 36,486 | 5,700 | |
| Missouri | 1,166 | 129 | 69,574 | 5,700 | | 709 | 127 | 67,833 | 5,700 | |
| Montana | 392 | 174 | 11,534 | 5,700 | | 271 | 136 | 10,811 | 5,700 | |

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

| Jurisdiction | Grade 4 | | | | | Grade 8 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Schools in frame | Schools in sample | Students in frame | Overall target student sample size | | Schools in frame | Schools in sample | Students in frame | Overall target student sample size | |
| Nebraska | 532 | 145 | 23,315 | 5,700 | | 294 | 114 | 22,561 | 5,700 | |
| Nevada | 394 | 124 | 35,875 | 5,945 | | 171 | 91 | 34,346 | 5,874 | |
| New Hampshire | 270 | 135 | 13,734 | 5,700 | | 142 | 89 | 14,078 | 5,700 | |
| New Jersey | 1,371 | 120 | 99,697 | 5,700 | | 765 | 118 | 99,117 | 5,700 | |
| New Mexico | 444 | 152 | 26,208 | 6,923 | | 232 | 123 | 25,079 | 7,021 | |
| New York | 2,471 | 164 | 201,226 | 7,899 | | 1,498 | 165 | 196,197 | 8,042 | |
| North Carolina | 1,457 | 215 | 118,118 | 10,570 | | 728 | 165 | 117,176 | 10,604 | |
| North Dakota | 261 | 166 | 8,471 | 5,700 | | 184 | 142 | 7,789 | 5,700 | |
| Ohio | 1,740 | 189 | 129,087 | 8,332 | | 1,093 | 186 | 131,562 | 8,269 | |
| Oklahoma | 869 | 132 | 50,988 | 5,700 | | 583 | 127 | 48,784 | 5,700 | |
| Oregon | 746 | 128 | 43,589 | 5,700 | | 428 | 124 | 42,824 | 5,700 | |
| Pennsylvania | 1,607 | 166 | 130,442 | 8,059 | | 888 | 162 | 131,525 | 8,167 | |
| Puerto Rico | 931 | 169 | 31,308 | 4,000 | | 398 | 161 | 30,211 | 4,000 | |
| Rhode Island | 164 | 111 | 10,777 | 5,700 | | 60 | 60 | 10,720 | 5,700 | * |
| South Carolina | 643 | 118 | 57,878 | 5,700 | | 306 | 115 | 54,617 | 5,700 | |
| South Dakota | 312 | 163 | 10,517 | 5,700 | | 246 | 135 | 9,657 | 5,700 | |
| Tennessee | 995 | 165 | 77,202 | 7,866 | | 584 | 149 | 73,441 | 7,907 | |
| Texas | 4,431 | 361 | 399,283 | 17,881 | | 2,251 | 251 | 383,849 | 17,999 | |
| Utah | 621 | 118 | 50,010 | 5,700 | | 256 | 113 | 47,320 | 5,700 | |
| Vermont | 216 | 216 | 6,204 | 6,204 | ** | 121 | 121 | 5,999 | 5,999 | ** |
| Virginia | 1,109 | 117 | 97,550 | 5,700 | | 379 | 114 | 95,187 | 5,700 | |
| Washington | 1,231 | 122 | 81,904 | 5,700 | | 609 | 122 | 79,084 | 5,700 | |
| West Virginia | 417 | 138 | 20,578 | 5,700 | | 190 | 110 | 20,464 | 5,700 | |
| Wisconsin | 1,099 | 181 | 61,686 | 8,024 | | 649 | 168 | 61,152 | 8,085 | |
| Wyoming | 192 | 137 | 7,639 | 5,700 | | 89 | 89 | 7,042 | 5,700 | * |
| Total | 52,343 | 8,314 | 3,831,663 | 369,705 | | 29,024 | 7,082 | 3,732,513 | 370,482 | |

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

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

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

## Stratification

Each state and grade will be stratified separately, but using a common approach in all cases. TUDA districts will be separated from the balance of their state, and each part stratified separately. The first level of stratification will be based on urban-centered type of location. This variable has 12 levels (some of which may not be present in a given state or TUDA district), and these will be collapsed so that each of the resulting location categories contains at least 9 percent of the student population (12 percent for large TUDA districts and 18 percent for small TUDA districts). In those states with school oversampling for NIES, the schools to be oversampled will be placed in a separate stratum, apart from the location strata used for other schools.

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

Within the resulting location and minority group classes (of which there are likely to be from three to fifteen, depending upon the jurisdiction), and the NIES oversampling stratum in states where this is applicable, schools will be sorted by a measure derived from school level results from the most recent available state achievement tests at the relevant grade. In general, mathematics test results will be used, but where these are not available, reading results will be used. In the few states that do not have math or reading tests at grades 4 and 8 (or where we are unable to match the results to the NAEP school frame), instead of achievement data, schools will be sorted using a measure of socio-economic status. This is the median household income of the 5-digit ZIP Code area where the school is located, based on the 2016 ACS (5-year) data. For BIE and DoDEA schools neither achievement data nor income data are available, and so grade enrollment is used in these cases.

Once the schools are sorted by location class, minority enrollment class, and achievement data (or household income), a systematic sample of schools will be selected using a random start. Schools will be sampled with probability proportional to size. The exact details of this process are described in the individual sampling specification memos.

## 2. Beta Sample

The beta sample comprises the national public school samples at grades 4, 8, and 12. At grades 4 and 8 the beta samples will be used for the national science assessments (PBA and DBA) and for pilot tests of reading, math, and vocabulary (DBA-only). At grade 12 the beta sample will be used for the operational reading, mathematics, and science assessments (PBA and DBA). Each of these samples will be nationally representative, selected to have minimal overlap with the alpha sample schools at the same grade. The number of students targeted per school will be 62 at grade 4, 63 at grade 8, and 68 at grade 12.

In order to increase the likelihood that the results for American Indian/Alaskan Native (AIAN) students can be reported for the operational samples, we will oversample high-AIAN public schools. That is, a public school with more than 5 AIAN students and greater than 5 percent AIAN enrollment will be given four times the chance of selection of a public school of the same size with a lower AIAN percentage. For all other schools, whenever there are more than 10 Black or Hispanic students enrolled and the combined Black and Hispanic enrollment exceeds 15 percent, the school will be given twice the chance of selection of a public school of the same size with a lower percentage of these two groups. This approach is effective in increasing the sample sizes of AIAN, Black, and Hispanic students without inducing undesirably large design effects on the sample, either overall, or for particular subgroups.

### Stratification

The Beta samples will have an implicit stratification, using a hierarchy of stratifiers and a serpentine sort. The highest level of the hierarchy is Census division (9 implicit strata). The next stratifier in the hierarchy is type of location, which has twelve categories. Many of the type of location strata nested within Census divisions will be collapsed with neighboring type of location cells (this will occur if the expected school sample size within the cell is less than 4.0). These geographic strata will be subdivided into three substrata: 1) schools being oversampled for AIAN, 2) schools being oversampled for Blacks and Hispanics, and 3) low-minority schools not being oversampled. If the expected sample size in an oversampled substratum is less than 8.0, it will be left as is. If the expected sample size is greater than 8.0, then it will be subdivided into up to four substrata (two for expected sample size up to 12.0, three for expected sample size up to 16.0, and four for expected sample size greater than 16.0). For the oversampling strata, the subdivision will be by percentage AIAN or percentage Black and Hispanic, as appropriate. For the low-minority sampling strata, the subdivision will be by state or groups of contiguous states. Within these substrata, the schools are to

be sorted by school type (public, BIE, DoDEA) and median household income from the 2016 5-year ACS (using a serpentine sort within the school type substrata).

## 3. Delta Samples

These are the private school samples at grades 4 and 8 for conducting the operational assessments in reading and mathematics. The sample sizes are large enough to report results by Catholic and non-Catholic at grades 4 and 8. Approximately half the sample at each grade will be from Catholic schools. The number of students targeted per school will be 50 at each grade.

### Stratification

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

## 4. Epsilon Sample

With regard to subjects and grades assessed, this sample is analogous to the beta sample, but for private schools. However, in contrast to the beta sample, there will be no oversampling of high minority schools. The same stratification variables will be used as for the delta samples. The epsilon sample schools will have minimum overlap with the delta sample schools which, given the respective sample sizes, means that no schools will be selected for both the delta and epsilon samples at the same grade. The number of students targeted per school will be 62 at grade 4, 63 at grade 8, and 68 at grade 12.

## IV. New Schools

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

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

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

## V. Substitute Samples

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

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

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

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

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

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

# VI. Contingency Samples

The districts that are taking part in the TUDA program are volunteers. Thus it is possible that at some point over the next few months, a given district might choose to opt out of the TUDA program for 2019. However, it is not acceptable for all schools in such a district to decline NAEP, as then the state estimates will be adversely affected. Thus to deal with this possibility, in each TUDA district, subsamples of the alpha sample schools will be identified as contingency samples. In the event that the district withdraws from the TUDA program prior to the selection of the student sample, all alpha sampled schools from that district will be dropped from the sample, with the exception of those selected in the contingency sample. The contingency sample will provide a proportional representation of the district, within the aggregate state sample. Student sampling in those schools will then proceed in the same way as for the other schools within the same state.

# VII. Student Sampling

Students within the sampled schools will be selected with equal probability. The student sampling parameters vary by sample type (Alpha, Beta, Delta, and Epsilon) and grade, as described below.

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

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

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

### Alpha Sample, Puerto Rico Grades 4 and 8

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

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

### Delta Samples, Grades 4 and 8

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

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

### Beta and Epsilon Samples, Grades 4, 8, and 12

A. At grade 4 all students will be selected, up to 70. If the school has more than 70 students, 62 will be selected. Of these students, 50 will be assigned to DBA and the rest to PBA. In schools with fewer than 21 students, all will be assigned to DBA or all to PBA. In schools with 32 to 37 students, 25 will be assigned to DBA and the rest to PBA. In all other schools, 25/31 of the students will be assigned to DBA with the rest to PBA.

B. At grade 8 all students will be selected, up to 70. If the school has more than 70 students, 63 will be selected. Of these students, 50 will be assigned to DBA and the rest to PBA. In schools with fewer than 21 students, all will be assigned to DBA or all to PBA. In schools with 31 to 37 students, 25 will be assigned to DBA and the rest to PBA. In all other schools, 50/63 of the students will be assigned to DBA with the rest to PBA.

C. At grade 12 all students will be selected, up to 75. If the school has more than 75 students, 68 will be selected. Of these students, 38 will be assigned to DBA and the rest to PBA. In schools with fewer than 20 students, all will be assigned to DBA or all to PBA. In schools with 32 to 36 students, 19 will be assigned to DBA and the rest to PBA. In all other schools, 19/34 of the students will be assigned to DBA with the rest to PBA.

## VIII. Weighting Requirements

### The Operational Reading and Mathematics Assessments, Grade 4 and 8

The sample weights will reflect probabilities of selection, school and student nonresponse, any trimming, and the random assignment to the particular subject. There will be separate replication schemes by grade and public/private. Weights will also be derived for the Puerto Rico KaSA assessment at grades 4 and 8.

### The Operational Reading and Mathematics Assessments, Grade 12, and Science Assessment, Grades 4, 8, and 12

The exact weighting requirements for these samples have yet to be determined. One possibility is that three sets of weights will be required – for DBA alone, PBA alone, and DBA/PBA combined. The sample weights will reflect probabilities of selection, school and student nonresponse, any trimming, and the random assignment to the particular subject. There will be a separate replication scheme by grade and public/private.

### Pilot Assessments in Reading, Mathematics, and Vocabulary, at Grades 4 and 8

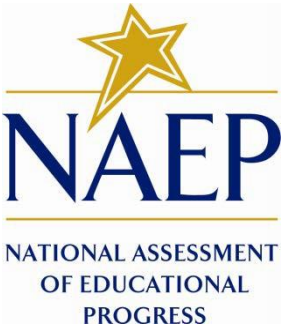
As is standard practice, only preliminary weights will be provided for these assessments. The sample weights will reflect probabilities of selection, and the random assignment to the particular subject (necessary because these assessments are spiraled in with other assessment components).

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

*National Assessment of Educational Progress (NAEP)*
*2019 and 2020*

# *Appendix C2*
# *2020 Sampling Memo*

*OMB# 1850-0928 v.14*




March 2019

| | | | |
|---|---|---|---|
| **Date:** | February 26, 2019 | **Memo:** | 2020-m01v01psu/s |

**To:**

| | |
|---|---|
| William Ward, NCES | William Wall |
| Ed Kulick, ETS | Rob Dymowski |
| David Freund, ETS | Chris Averett |
| Amy Dresher, ETS | Kavemuii Murangi |
| Cathy White, Pearson | John Burke |
| Saira Brenner, Fulcrum | Dwight Brock |
| Greg Binzer | Joel Wakesberg |
| Lauren Byrne | Jing Kang |
| Lisa Rodriguez | Veronique Lieber |
| Rick Rogers | Shaohua Dong |

**From:** Dave Hubble

**Reviewers:** Keith Rust, Leslie Wallace

**Subject:** Sample Design for 2020 NAEP - Overview

## I. Introduction

For 2020, the sample design involves only one component: Long-Term Trend (LTT) Paper-based Assessment (PBA).

1. LTT Reading Operational assessments at ages 9, 13, and 17;

2. LTT Mathematics Operational assessments at ages 9, 13, and 17;

There will be no pilot assessments in 2020 LTT PBA.

The target sample sizes of assessed students for LTT are shown in Table 1 (which also shows the estimated numbers of sampled schools before attrition). Unlike most years, the NAEP 2020 LTT assessment components will take place at various seasons throughout the school year. With that in mind, the last column was added to provide the season in which the assessment will be fielded.

Table 1. 2020 NAEP Sample Sizes (Public and Private) and Season Fielded

| | Session | Public school students | Private school students | Total students | Season Fielded |
|---|---|---|---|---|---|
| **Age 9** | | | | | |
| LTT Math (O) | | 7,200 | 800 | 8,000 | |
| LTT Reading (O) | LT09 | 7,200 | 800 | 8,000 | |
| Subtotal | | 14,400 | 1,600 | 16,000 | |
| **Schools** | | **430** | **220** | **650** | **Winter** |
| **Age 13** | | | | | |
| LTT Math (O) | | 7,200 | 800 | 8,000 | |
| LTT Reading (O) | LT13 | 7,200 | 800 | 8,000 | |
| Subtotal | | 14,400 | 1,600 | 16,000 | |
| **Schools** | | **440** | **190** | **630** | **Fall** |
| **Age 17** | | | | | |
| LTT Math (O) | | 7,200 | 800 | 8,000 | |
| LTT Reading (O) | LT17 | 7,200 | 800 | 8,000 | |
| Subtotal | | 14,400 | 1,600 | 16,000 | |
| **Schools** | | **490** | **130** | **620** | **Spring** |
| **GRAND TOTAL** | | 43,200 | 4,800 | 48,000 | |
| **Schools** | | **1,360** | **540** | **1,900** | |

(O) = Operational

## II. Assessment Types

For 2020 NAEP, there is only one type of assessment. While the detailed target counts of LTT assessed students are provided in Table 1, a summary of major points follows.

The LTT spiral at ages 9, 13, and 17. This paper-based assessment (PBA) will be conducted in LTT PSUs. The spiral includes Math and Reading operational samples. The LTT session type has a target of 16,000 assessed students each at age 9, age 13, and age 17. Note, 10% of the assessed students are allocated to private schools. This roughly represents a proportional sample, as about 10% of the population attends private schools.

## III. Primary Sampling Units Selection and Overlap Control

As the LTT assessments are national, with a total sample size of assessed students of about 48,000, for reasons of operational efficiency in conducting the assessments a sample of Primary Sampling Units (PSUs) was selected, and all sampled schools were drawn from within the sampled PSUs.

The PSUs were created from aggregates of counties. Data on counties were obtained from the 2010 Census, and the definitions of Metropolitan Statistical Areas (MeSAs) used were the December 2009 Office of Management and Budget (OMB) definitions. Each Metropolitan Statistical Area (MeSA) constitutes a PSU, except that MeSAs that cross state boundaries were split into their individual regional components.

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 almost contiguous counties (separated by MeSA counties) otherwise. Each PSU falls within a single state.

This process generated a frame of approximately 1,000 PSUs. The PSUs were stratified, using characteristics aggregated from county-level characteristics, found by analysis to be related to NAEP achievement in past assessments. A sample of 105 PSUs was selected for the LTT samples. Twenty-nine large MeSAs were selected with certainty, and the remaining sample was a stratified probability proportional to size (PPS) sample, where the size measure was a function of the number of children as given in the most recent population estimates prepared by the U.S. Census Bureau.

## IV. Stratification and Oversampling

As in the recent past, the plan is to draw separate public and private school samples. This approach has proven to be useful, in that, selecting the samples separately has three 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 schools and private schools; and 3) it makes it easier to use different sort variables for public schools and private schools. It also allows for the possibility of a late change of mind concerning the sample sizes that differ between public and private schools.

Explicit stratification will take place at the PSU level. For schools within PSUs, stratification gains are 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. 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, such as a race/ethnicity percentage variable. The exact set of variables used in sorting the schools prior to sampling will be specified in the particular sampling specification memos.

In addition, we will implement oversampling of certain public schools. In order to increase the likelihood that the results for American Indian/Alaskan Native (AIAN) students can be reported for the operational samples, we will oversample high-AIAN public schools for LTT for ages 9, 13, and 17. That is, a public school with 5 percent or more AIAN enrollment will be given four times the chance of selection of a public school of the same size with a lower AIAN percentage. Recent research into oversampling schemes that could benefit AIAN students indicates that this approach should be effective in increasing the sample sizes of AIAN students, without inducing undesirably large design effects on the sample, either overall or for particular subgroups. In addition, high minority public schools for LTT that are not oversampled for AIAN enrollment will be oversampled for Black and Hispanic enrollment. That is, a public school with 15 percent or more Black and Hispanic combined enrollment will be given twice the chance of selection of a public school of the same size with a lower percentage of these two groups. This approach is effective in increasing the sample sizes of Black and Hispanic students, without inducing undesirably large design effects on the sample, either overall or for particular subgroups. Beyond this, we will also implement the oversampling of AIAN, Black, and Hispanic students at the student level in schools not being oversampled at the school level.

The preliminary 2017/18 CCD and the updated 2017/18 PSS school files were approved for use by NCES. They serve as the basis for the public and private school frames for the 2020 NAEP.

## V. New Schools

To compensate for the fact that files used to create the NAEP school sampling frames are two years out of date at the time of assessment, we will supplement the samples in the LTT PSUs with a sample of new public schools for each age sample. .

The new school samples will be drawn using a three-stage design. The first stage is the selection of the LTT sample PSUs, as discussed above. At the second stage, a national sample of school districts will be selected from the LTT sample PSUs. The sampled districts will be asked to review lists of their respective schools and identify new schools. Frames of new schools will be constructed from these updates, and, at the third stage, new schools will be drawn with probability proportional to size using the same sampling rates as their corresponding original school samples.

Note that the student and school sample sizes in Table 1 do not reflect these new school samples. However, some schools from the original sample will prove to be closed or otherwise ineligible, and the new school procedure essentially compensates for the sample losses from these sources, as well as ensuring full coverage of the population.

## VI. Within PSU Overlap Control with Other Samples

As LTT is the only NAEP sample in 2020 and there are no other NCES-related operational samples (e.g., PIRLS, PISA, etc.) in 2020 there will be no need for LTT within PSU sampling overlap control. Selection of 2020 Field Trial schools for PIRLS and PISA will avoid NAEP LTT sample schools.

## VII. Substitute Samples

A portion of the eligible 2020 LTT sample schools will choose to not participate in the assessment. In order to maintain sample yields, substitute school samples will be selected for each of the 2020 LTT samples. Within the 2020 LTT samples, the order for selecting substitute schools will be from “oldest” to “youngest”. That is, age 17, 13, and then 9. This ordering of samples by age is necessary since no school can be selected as a substitute more than once and there are fewer schools available to serve as substitutes at the higher ages. This will be done separately for both public and private schools. The general steps for selecting substitutes are to put the substitute frames in their original sampling sort order, and take the 'nearest neighbor' of each original sampled school, excluding schools selected for any of the NAEP 2020 LTT samples, schools already selected to serve as a substitute school, and schools which cross PSU or state boundaries, as potential substitutes.

The nearest neighbor is the school adjacent (immediately preceding or succeeding) the original school in the sorted frame with the closer estimated age enrollment value. If estimated age enrollment of both potential substitute schools differs from the original school by the exact same amount, the selection procedure will randomly choose one of the schools. If neither the preceding or succeeding school is eligible to be a substitute, then the sampled school is not assigned a substitute.

In addition, sampled private schools whose school affiliation is unknown will not get substitutes nor can such private schools not in sample serve as substitute schools. Also, new schools will not get substitute schools nor serve as substitutes.

## VIII. Student Sampling

Students within the sampled schools will be selected with equal probability, except in public schools where oversampling of AIAN, Black and Hispanic students will take place. In addition to this, student sample sizes for LTT within each school are determined as the combined result of several factors:

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

2. We do not wish to have a sample that is too clustered for any one assessment subject.

3. We do not wish to have many physical sessions that contain only a very small number of students, as this is inefficient.

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

The plans for LTT below reflect the design that results from considering each of these factors and balancing them.

LTT Private Schools and Oversampled Public Schools

In all private schools and public schools that are oversampled (as described in Section IV), the target sample size is 50 assessed students for each age. We will select all students of a certain age, up to 50. In schools with more than 50 such students we will select 50. There will be only one session type.

LTT Non-Oversampled Public Schools

In public schools not oversampled at the school level (i.e., under 5% AIAN and under 15% Black and Hispanic students), we will select 50 students plus an oversample of up to 5 additional AIAN, Black, and Hispanic students. The maximum number of sample students will be 55 in these schools.

## IX. Weighting Requirements

The LTT operational samples currently require a single set of weights for each subject (LTT Math and LTT Reading at ages 9, 13, and 17), applied to reflect probabilities of selection, school and student nonresponse, any trimming, and the random assignment to the particular subject. There will be a separate replication scheme by age and public/private. LTT Preliminary weights will be developed as required by the DAR contractor.