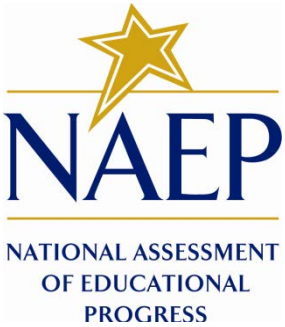
NATIONAL CENTER FOR EDUCATION STATISTICS
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

National Assessment of Educational Progress (NAEP) 2018 and 2019

Appendix A-C

Appendix A: External Advisory Committees
Appendix B: NAEP 2011 Weighting Procedures
Appendix C: 2018 Sampling Memo

OMB# 1850-0928 v.5



April 2017

NATIONAL CENTER FOR EDUCATION STATISTICS
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

National Assessment of Educational Progress (NAEP) 2018 and 2019

Appendix A

External Advisory Committees

OMB# 1850-0928 v.5

March 2017

Table of Contents

Appendix A-1: NAEP Design and Analysis Committee....4
Appendix A-2: NAEP Validity Studies Panel....4
Appendix A-3: NAEP Quality Assurance Technical Panel....5
Appendix A-4: NAEP National Indian Education Study Technical Review Panel....5
Appendix A-5: NAEP Civics Standing Committee....6
Appendix A-6: NAEP Economics Standing Committee....6
Appendix A-7: Geography Standing Committee....6
Appendix A-8: NAEP Mathematics Standing Committee....7
Appendix A-9: NAEP Reading Standing Committee....7
Appendix A-10: NAEP Science Standing Committee....8
Appendix A-11: NAEP Survey Questionnaires Standing Committee....8
Appendix A-12: NAEP Technology and Engineering Literacy Standing Committee....9
Appendix A-13: NAEP U.S. History Standing Committee....9
Appendix A-14: NAEP Writing Standing Committee....10
Appendix A-15: NAEP Principals' Panel Standing Committee....10
Appendix A-16: NAEP Mathematics Translation Review Committee....11
Appendix A-17: NAEP Science Translation Review Committee....11
Appendix A-18: NAEP Grade 8 Social Sciences Translation Review Committee....11

Appendix A-1: NAEP Design and Analysis Committee

Name	Affiliation
Betsy Becker	Florida State University
Peter Behuniak	University of Connecticut
Lloyd Bond	University of North Carolina, Greensboro (Emeritus)/ Carnegie Foundation (retired)
Steve Elliott	University of Colorado
Derek Briggs	University of Colorado
Steve Elliott	Arizona State University
Ben Hansen	University of Michigan
Matthew Johnson	Columbia University
Brian Junker	Carnegie Mellon University
David Kaplan	University of Wisconsin-Madison
Kenneth Koedinger	Carnegie Mellon University
Yan Li	University of Maryland
Sophia Rabe-Hesketh	University of California, Berkeley
Michael Rodriguez	University of Minnesota
S. Lynne Stokes	Southern Methodist University
Chun Wang	University of Minnesota

Appendix A-2: NAEP Validity Studies Panel

Name	Affiliation
Peter Behuniak	University of Connecticut
George Bohrnstedt	American Institutes for Research, Washington, DC
Jim Chromy	RTI International (Emeritus Fellow), Raleigh, NC
Phil Daro	Strategic Education Research (SERP) Institute, Berkeley, CA University of
Richard Duran	California
David Grissmer	University of Virginia
Larry Hedges	Northwestern University
Gerunda Hughes	Howard University
Ina Mullis	Boston College
Scott Norton	Council of Chief State School Officers, Washington, DC
Jim Pellegrino	University of Illinois at Chicago/Learning Sciences Research Institute
Gary Phillips	American Institutes for Research, Washington, DC
Lorrie Shepard	University of Colorado at Boulder
David Thissen	The University of North Carolina at Chapel Hill
Gerald Tindal	University of Oregon
Sheila Valencia	Universty of Washington

Appendix A-3: NAEP Quality Assurance Technical Panel

Name	Affiliation
Jamal Abedi	University of California, Davis
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
Jim Pellegrino	University of Illinois at Chicago/Learning Sciences Research Institute
Mark Reckase	Michigan State University
Michael (Mike) Russell	Boston College
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
Susan C. Faircloth	University of North Carolina Wilmington
Jeremy MacDonald	Rocky Boy Elementary, Box, Elder, MT
Rebecca Izzo-Manymules	Southwest Indian Polytechnic Institute, Albuquerque, NM
Jeannette Muskett Miller	Tohatchi High School, Tohatchi, NM
Debora Norris	Salt River Pima-Maicopa Indian Community
Martin Reinhardt	Northern Michigan University
Tarajean Yazzie-Mintz	Wakanyeja ECE Initative/American Indian College Fund, Denver, CO

Appendix A-5: NAEP Civics Standing Committee

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

Appendix A-6: NAEP Economics Standing Committee

Name	Affiliation
Kris Bertelsen	Little Rock Branch-Federal Reserve Bank of St. Louis, Little Rock, AR
Stephen Buckles	Vanderbilt University
Steven L. Cobb	University of North Texas
Jaime Festa-Daigle	Lake Havasu High School, Lake Havasu City, AZ
Julie Heath	University of Memphis
Richard MacDonald	St. Cloud State University
Andrea Morgan	Oregon Department of Education, Salem, OR
Kevin Smith	Renaissance High School, Detroit, MI
William Walstad	University of Nebraska–Lincoln

Appendix A-7: Geography Standing Committee

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

Appendix A-8: NAEP Mathematics Standing Committee

Name	Affiliation
Jennifer Alvarez	Sultana Elementary School, Ontario, CA
Daniel Chazan	University of Maryland, College Park
Carl Cowen	Indiana University–Purdue University
Julie Guthrie	Texas Education Agency
Kathleen Heid	Pennsylvania State University
Mark Howell	Gonzaga College High School, Washington, DC
Russ Keglovits	Nevada Department of Education, Carson City, NV
Carolyn Maher	Rutgers University
Michele Mailhot	Maine Department of Education, Augusta, ME
Brian Nelson	Curtis Corner Middle School, Wakefield, RI
Matthew Owens	Spring Valley High School, Columbia, SC
Carole Philip	Alice Deal Middle School, Washington, DC
Melisa M. Ramos Trinidad	Educación Bilingüe Luis Muñoz Iglesias, Cidra, PR
Ann Trescott	Stella Maris Academy, La Jolla, CA

Appendix A-9: NAEP Reading Standing Committee

Name	Affiliation
Marilyn Adams	Brown University
Peter Afflerbach	University of Maryland
Patricia Alexander	University of Maryland
Margretta Browne	Richard Montgomery High School, Silver Spring, MD
Julie Coiro	University of Rhode Island
Bridget Dalton	University of Colorado Boulder
Valerie Harrison	Claflin University
Karen Malone	Fort Wingate High School, Fort Wingate, NM
Pamela Mason	Harvard Graduate School of Education
Margaret McKeown	University of Pittsburgh
P. David Pearson	University of California, Berkeley
Jenny Thomson	University of Sheffield, Sheffield, UK
Monica Verra-Tirado	Florida Department of Education, Tallahassee, FL
Victoria Young	Texas Education Agency, Austin, TX
Zynia Zepeda	Crane Elementary School District, Yuma, AZ

Appendix A-10: NAEP Science Standing Committee

Name	Affiliation
Alicia Cristina Alonzo	Michigan State University
George Deboer	American Association for the Advancement of Science, Washington, DC
Alex Decaria	Millersville University
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
Hunter Gehlbach	Harvard University
Camille Farrington	University of Chicago, Chicago, IL
Gerunda Hughes	Howard University
David Kaplan	University of Wisconsin-Madison
Henry Levin	Teachers College, Columbia University
Stanley Presser	University of Maryland
Augustina Reyes	University of Houston, Houston, TX
Leslie Rutkowski	Indiana University Bloomington
Jonathon Stout	Lock Haven University
Roger Tourangeau	Westat, Rockville, MD
Akane Zusho	Fordham University

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

Name	Affiliation
Keith Barton	Indiana University Bloomington
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
Chris Dede	Harvard Graduate School of Education
Richard Duran	University of California, Santa Barbara
Maurice Frazier	Oscar Smith High School, Chesapeake, VA
Camilla Gagliolo	Arlington Public Schools, Arlington, VA
Christopher Hoadley	New York University
Eric Klopfer	Massachusetts Institute of Technology
Beth McGrath	Stevens Institute of Technology
Greg Pearson	National Academy of Engineering, Washington, DC
John Poggio	University of Kansas
Erin Reilly	University of Southern California
Troy Sadler	University of Missouri Science Education Center, Columbia, MO
Kimberly Scott	Arizona State University
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
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 Writing Standing Committee

Name	Affiliation
Margretta Browne	Montgomery County Public Schools, Silver Spring, MD
Robert Crongeyer	Robla School, Sacramento, CA
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
Drew Sterner	Tamanend Middle School, Warrington, PA
Sherry Swain	National Writing Project, Berkeley, CA
Victoria Young	Texas Education Agency, Austin, TX

Appendix A-15: 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
Jason Mix	Howard Lake–Waverly–Winsted High School, Howard Lake, MN

Appendix A-16: NAEP Mathematics Translation Review Committee

Name	Affiliation
Gilberto Cuevas	Texas State University, San Marcos, TX
Néstor Díaz	Coral Gables Senior High School, Coral Gables, FL
David Feliciano	P.S./M.S 29, The Melrose School, Bronx, NY
Yvonne Fuentes	Author and Spanish Linguist, Carrollton, GA
Flor Yanira Gurrola Valenzuela	Washington Middle School, Albuquerque, NM
Luz N. Rosario Cristóbal	Puerto Rico Department of Education, Hato Rey, PR
Melisa M. Ramos Trinidad	Educación Bilingüe Luis Muñoz Iglesias, Cidra, PR
Sonia Suazo	Escuela Salvador Brau Elemental, Cayey, PR
Enid Valle	Kalamazoo College, Kalamazoo, MI

Appendix A-17: NAEP Science Translation Review Committee

Name	Affiliation
Néstor Díaz	Coral Gables Senior High School, Coral Gables, FL
Yvonne Fuentes	Author and Spanish Linguist, Carrollton, GA
Myrna Rasmussen	Austin Independent School District, Austin, TX
Enid Valle	Kalamazoo College, Kalamazoo, MI

Appendix A-18: 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, NY
Dagoberto Eli Ramierz	Bilingual Education Expert, Palmhurst, TX
Enid Valle	Kalamazoo College, Kalamazoo, MI

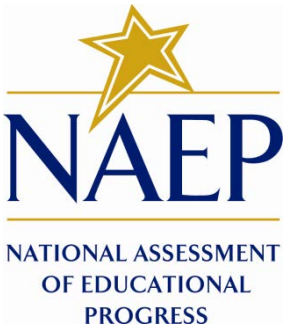
NATIONAL CENTER FOR EDUCATION STATISTICS
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

National Assessment of Educational Progress (NAEP) 2018 and 2019

Appendix B

NAEP 2011 Weighting Procedures

OMB# 1850-0928 v.5



March 2017

NAEP Technical Documentation Website

Weighting Procedures for the 2011 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 2011, weights were developed for students sampled at grades 4, 8, and 12 for assessments in mathematics, reading, science, and a writing computer-based assessment (WCBA). 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, which 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.

Computation of Full-Sample Weights

- Computation of Base Weights
- School and Student Nonresponse Adjustments
- School and Student Weight Trimming Adjustments
- Student Raking Adjustment

Computation of Replicate Weights for Variance Estimation

Quality Control on Weighting Procedures

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 2011 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.

Starting in 2009, an additional weighting adjustment was implemented in the state samples so that estimates for key student-level characteristics were in agreement across assessments in reading, mathematics, and science. This procedure 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. In 2011, a finite population correction (fpc) factor was used in computing variance estimates for the reading, mathematics, and science 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.

Computation of Full-Sample Weights for the 2011 Assessment

Computation of Base Weights

School and Student Nonresponse Weight Adjustments

School and Student Weight Trimming Adjustments

Student Raking Adjustment

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, mathematics, science, and writing [WCBA]). 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 varies by school type and grade. See the links below for descriptions of the school strata for the various assessments.

- Reading, mathematics, and science at grades 4 and 8 for public schools
- Reading, mathematics, and science at grades 4 and 8 for private schools
- WCBA at grades 8 and 12 for public schools
- WCBA at grades 8 and 12 for private schools

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_comp_full_samp_weights.aspx

Computation of Base Weights for the 2011 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);
- sampling frame (new school frame or not); and
- assessment subject (writing [WCBA] samples involved the selection of geographic units known as primary sampling units [PSUs], but reading, mathematics, and science did 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, mathematics, science, or writing (WCBA) assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_base.aspx

---

School Base Weights for the 2011 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); and
- assessment subject (writing [WCBA] samples involved the selection of geographic units known as primary sampling units [PSUs], but reading, mathematics, and science did not).

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

The overall probability of selection of an originally selected school in the WCBA sample reflects two components:

- the probability of selection of the PSU; and
- the probability of selection of the school within the selected PSU from the NAEP public/private school frame.

Substitute schools for the 2011 private school national assessment

Substitute public schools for the 2011 computer-based assessment

Substitute private schools for the 2011 computer-based assessment

The overall selection probability of a school from the new school frame in a reading, mathematics, science, or WCBA 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.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_base_school.aspx

Student Base Weights for the 2011 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 that the particular session type was assigned to the school, that the student was assigned to the session type;
- $SUBJFAC_{jsk}$ 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 mathematics, reading, science, or writing (WCBA) booklets to the students. The subject factor varies by grade, subject, school type (public/private), jurisdiction (states participating in the science assessment/states declining to participate in the science assessment), and it is equal to the inverse of the booklet proportions (mathematics, reading, science, or WCBA) 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 $1/(1\text{-}OFF_{js}/100)$.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_base_stud.aspx

School and Student Nonresponse Weight Adjustments for the 2011 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/2011/weighting_2011_nonresp.aspx

School Nonresponse Weight Adjustment for the 2011 Assessment

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

- grade (4, 8, 12);
- school type (public, private); and
- assessment subject (reading, mathematics, science, writing [WCBA]).

Development of Initial School Nonresponse Cells

Development of Final School Nonresponse Cells

School Nonresponse Adjustment Factor Calculation

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_nonresp_schl.aspx

Development of Final School Nonresponse Cells for the 2011 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 Reading, Mathematics, and Science at Grades 4 and 8

For the 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 jurisdiction or TUDA district.

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

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 affiliation.

Public School Samples for WCBA at Grades 8 and 12

For the public school samples, cells with similar high percentage Black/Hispanic status within a given census region and urbanicity stratum were collapsed first. If further collapsing was required, cells with the most similar urbanicity strata within a given census region were combined next. No further collapsing occurred after all levels of urbanicity strata were collapsed. That is, collapsing never occurred across census region.

Private School Samples for WCBA at Grades 8 and 12

For the private school samples, if collapsing was necessary, all census region cells within a given affiliation were collapsed. However, collapsing never occurred across affiliation.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_nonresp_schl_final.aspx

Development of Initial School Nonresponse Cells for the 2011 Assessment

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 2011, the reading, mathematics, and science assessments were spiraled together, but writing (WCBA) was not spiraled with any other subject.

The initial nonresponse cells for the various NAEP 2011 samples are described below.

Public School Samples for Reading, Mathematics, and Science 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, or achievement level, or median income, or grade enrollment.

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

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

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

- affiliation;
- census division stratum;
- urbanicity stratum; and
- race/ethnicity classification.

Public School Samples for WCBA at Grades 8 and 12

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

- census region;
- urban-centric locale; and
- indicator of high percentage of Black/Hispanic students within school.

Private School Samples for WCBA at Grades 8 and 12

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

- affiliation; and
- census region.

School Nonresponse Adjustment Factor Calculation for the 2011 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 grade enrollment corresponding to the original sampled school.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_nonresp_schl_factor.aspx

---

Student Nonresponse Weight Adjustment for the 2011 Assessment

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 (reading, mathematics, science, writing [WCBA]).

Development of Initial Student Nonresponse Cells

Development of Final Student Nonresponse Cells

Student Nonresponse Adjustment Factor Calculation

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_nonresp_stud.aspx

Development of Final Student Nonresponse Cells for the 2011 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 group, 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/ELL for any sample.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_nonresp_stud_final.aspx

Development of Initial Student Nonresponse Cells for the 2011 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 2011, the reading, mathematics, and science assessments were spiraled together, but writing (WCBA) was not spiraled with any other subject. The initial student nonresponse cells for the various NAEP 2011 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, Mathematics, and Science 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[1] (classified into "older" 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.

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

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

- SD/ELL;
- school nonresponse cell;
- age[1] (classified into "older" 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.

Public School Samples for WCBA at Grades 8 and 12

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

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

Private School Samples for WCBA at Grades 8 and 12

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

- SD/ELL;

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

[1] Older students are those born before October 1, 2000, for grade 4; October 1, 1996, for grade 8; and October 1, 1992, for grade 12.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_nonresp_stud_initial.aspx

---

Student Nonresponse Adjustment Factor Calculation for the 2011 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/2011/weighting_2011_nonresp_stud_factor.aspx

---

School and Student Weight Trimming Adjustments for the 2011 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/2011/weighting_2011_trimming_adj.aspx

Trimming of School Base Weights for the 2011 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.

Twenty-one (21) new schools had their weights trimmed: seven at grade 4, and fourteen at grade 8.

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/2011/weighting_2011_base_schtrim.aspx

Trimming of Student Weights for the 2011 Assessment

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 for reading and mathematics at grades 4 and 8, and science at grade 8; by the nation for the public school and private school samples for writing [WCBA] at grades 8 and 12; and by affiliation (Catholic, Conservative Christian, Lutheran, and Other private) for the private school samples for reading, mathematics, and science at grades 4 and 8.

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:

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

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 2011 assessment, relatively few students had weights considered excessively large. Out of the approximately 987,000 students included in the combined 2011 assessment samples, 710 students had their weights trimmed.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_studtrim.aspx

Student Weight Raking Adjustment for the 2011 Assessment

Development of Final Raking Dimensions

Raking Adjustment Control Totals

Raking Adjustment Factor Calculation

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. The new 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. The new weighting step 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 2011, 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 science public school samples at grade 8. 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, mathematics, and science samples combined.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_stud_raking.aspx

Development of Final Raking Dimensions for the 2011 Assessment

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 TUDA 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 Two or More Races; Asian with the Pacific Islander; 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/2011/weighting_2011_stud_raking_final_dim.aspx

Raking Adjustment Control Totals for the 2011 Assessment

The control totals used in the raking procedure for NAEP 2011 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/2011/weighting_2011_stud_raking_ctrl_tots.aspx

---

Raking Adjustment Factor Calculation for the 2011 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 intialized 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

- 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 equalling $STUSAWT_k^{adj(4)}$ from the previous iteration. The process continued until convergence was reached.

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

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

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/2011/weighting_2011_stud_raking_factor_cal.aspx

---

Computation of Replicate Weights for the 2011 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 2011 assessment samples. The theory that underlies the jackknife variance estimators used in NAEP studies is discussed in the section Replicate Variance Estimation.

Defining Variance Strata and Forming Replicates

Computing School-Level Replicate Factors

Computing Student-Level Replicate Factors

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 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 past assessments 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 stratum 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.

In the case of PSU-based samples, the replication procedure was only carried out at the school level. Conceptually, one can include this process under the framework of replication at both school and student levels but where the replicate factors at the student level are equal to 1.0 for every replicate for every student.

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/2011/weighting_2011_repwts.aspx

---

Computing School-Level Replicate Factors for the 2011 Assessment

The calculation of school-level replicate factors for an assessment depended on the variance replication procedure, with or without finite population corrections for the first stage of sampling.

Mathematics, Reading, and Science Assessments

The replicate variance estimation approach for the mathematics, reading, and science 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 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 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,

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

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

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

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.

Computer-Based Writing Assessments

The replicate variance estimation approach for the computer-based writing assessments did not involve school-level finite population corrections. As described in the section Defining Variance Strata and Forming Replicates, variance strata were defined by grouping first-stage units (schools or geographic PSUs) into pairs or triplets. The calculation of the school-level replicate factors for each of the 62 replicates was based on this grouping.

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

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

where

- 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 schools in variance strata comprising three first-stage units, school-level replicate factors for school were calculated as follows:

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

$$SCH\_REPFAC_{js}(r) = \begin{cases} 1.5, & \text{for } js \in R_{jr}, U_{js} = 1 \\ 1.5, & \text{for } js \in R_{jr}, U_{js} = 2 \\ 0, & \text{for } js \in R_{jr}, U_{js} = 3 \end{cases}$$

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

$$SCH\_REPFAC_{js}(r^1) = \begin{cases} 1.5, & \text{for } js \in R_{jr}, U_{js} = 1 \\ 1.0, & \text{for } js \in R_{jr}, U_{js} = 2 \\ 1.5, & \text{for } js \in R_{jr}, U_{js} = 3 \end{cases}$$

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

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

where

- 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/2011/weighting_2011_repwts_schl.aspx

---

Computing Student-Level Replicate Factors for the 2011 Assessment

The calculation of student-level replicate factors for an assessment depended on whether the variance replication procedure incorporated finite population corrections at the first stage of sampling.

Mathematics, Reading, and Science Assessments

For the mathematics, reading, and science 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

- π_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 $r' = r + 31 \pmod{62}$:

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

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

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

where

- π_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.

Computer-Based Writing Assessments

For the computer-based writing assessments, which did not involve first-stage finite population corrections and there were no certainty schools, replication at the student level was not relevant. As a consequence, the replicate factors for all replicates for every student in the computer-based writing assessments were set to unity. That is, the student-level replicate factors, $STU\_REPFAC_{jsk}(r)$, $r = 1,..., 62$, were calculated as follows:

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

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_repwts_stud.aspx

Defining Variance Strata and Forming Replicates for the 2011 Assessment

In the NAEP 2011 assessment, replicates were formed separately for each sample indicated by grade (4, 8, 12), school type (public, private), and assessment subject (mathematics, reading, science, computer-based writing). To reflect the school-level finite population corrections in the variance estimators for the two-stage samples used for mathematics, reading, and science assessments, variance strata were formed at both the school and student levels for these samples. For the computer-based writing assessments, which did not use school-level finite population corrections in the variance estimators, variance strata were formed only at the first-stage level.

The first step in forming replicates was to create variance strata in each primary stratum. In 2011, the mathematics, reading, and science assessments required formation of variance strata at both the school and student levels. The computer-based writing assessments required formation of variance strata only at the first-stage level.

The next step was to sort the appropriate sampling unit (school or student) in the order of its selection within the primary stratum and then pair off 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.

If there were more than 62 preliminary variance strata within a primary stratum, the preliminary variance strata were grouped to form 62 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, Mathematics, and Science Assessments

As described above, the mathematics, reading, and science assessments required variance strata at both the school and student level.

At the school-level for these samples, formation of 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 and 8; and
- Private school affiliation (Catholic, Lutheran, Conservative Christian, and Other private) for the private school samples at grades 4 and 8.

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.

Computer-Based Writing Assessments

As described above, variance strata for the computer-based writing assessments were formed at the first-stage sampling level and so differed by certainty and noncertainty PSUs. For noncertainty PSUs, the first-stage sampling 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 sampling 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, variance 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 demographic and socioeconomic characteristics such as percentage of Black youth and percentage of children whose family income is below the poverty threshold. The formation of these 19 variance 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 variance strata must reflect the sampling of schools within the certainty PSUs. Variance 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, minority status, 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 variance strata. Within each primary stratum with an even number of first-stage units, all of the preliminary variance strata were pairs, and within primary strata with an odd number of first-stage units, one of the variance strata was a triplet (the last one), and all others were pairs.

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

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/2011/weighting_2011_repwts_strata.aspx

---

Replicate Variance Estimation for the 2011 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:

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

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:

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

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

$$E(\bar{y}_r(J) - \bar{y}_r)^2 = \frac{1}{(n_r/2)}\frac{n_r - n_r/2}{n_r}\frac{1}{n_r - 1}\sum_{i=1}^{n_r}\left(y_{r_2} - \bar{y}_r\right)^2 =$$

$$= \frac{1}{n_r(n_r - 1)}\sum_{i=1}^{n_r}\left(y_{r_2} - \bar{y}_r\right)^2 = \frac{s_r^2}{n_r}$$

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,

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

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

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

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

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

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, 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):

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

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

$$E(\bar{y}_r(J) - \bar{y}_r)^2 = \frac{1}{(2n_r/3)} \frac{n_r - (2n_r/3)}{n_r} \frac{1}{n_r - 1} \sum_{i=1}^{n_r} (y_{r_2} - \bar{y}_r)^2$$
$$= \frac{1}{2n_r(n_r - 1)} \sum_{i=1}^{n_r} (y_{r_2} - \bar{y}_r)^2 = \frac{s_r^2}{2n_r}$$

$$E(\bar{y}_{r'}(J)-\bar{y}_r)^2 = \frac{1}{(2n_r/3)}\frac{n_r-(2n_r/3)}{n_r}\frac{1}{n_r-1}\sum_{i=1}^{n_r}(y_{r_2}-\bar{y}_r)^2$$

$$= \frac{1}{2n_r(n_r-1)}\sum_{i=1}^{n_r}(y_{r_2}-\bar{y}_r)^2 = \frac{s_r^2}{2n_r}$$

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/2011/weighting_2011_repwts_appdx.aspx

---

Quality Control on Weighting Procedures for the 2011 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 2009 demographic data; and
- nonresponse bias analyses.

Main QC Findings of Interest

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/2011/weighting_2011_qa.aspx

Main Quality Control Findings of Interest for the 2011 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), for public schools of the Bureau of Indian Education (BIE) for grades 4 and 8, and for Colorado for grade 8. Student-level participation rates were also below 85 percent for students in Detroit for grade 8 public schools. As required by NCES standards, nonresponse bias analyses were conducted on each reporting group falling below the 85 percent participation threshold.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/weighting_2011_qa_findings.aspx

Final Participation, Exclusion, and Accommodation Rates

Grade 4 Mathematics
Grade 4 Reading

Grade 8 Mathematics
Grade 8 Reading
Grade 8 Science
Grade 8 Writing (WCBA)

Grade 12 Writing (WCBA)

Participation, Exclusion, and Accommodation Rates for Grade 8 Mathematics for the 2011 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: 2011

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	**8,700**	**97.57**	**88.42**	**209,000**	**2.45**	**92.72**	**9.74**
National all[1]	8,600	97.54	88.31	204,000	2.46	92.71	9.67
Northeast all	1,400	95.45	79.60	33,200	2.25	91.82	14.40
Midwest all	2,400	98.76	91.49	45,900	2.62	93.33	9.84
South all	2,700	97.65	89.96	73,300	3.22	93.13	8.35
West all	2,100	97.80	89.74	50,200	1.30	92.19	8.11
National public	7,500	99.79	99.76	192,000	2.64	92.53	10.14
Alabama	125	100.00	100.00	3,200	1.21	93.72	3.59
Alaska	167	99.90	97.93	2,900	3.15	89.48	14.39
Arizona	136	99.02	99.08	3,200	1.14	92.97	8.87
Arkansas	126	100.00	100.00	3,200	1.36	92.53	11.63
California	257	100.00	100.00	8,400	1.08	91.81	7.47
Colorado	130	99.87	97.31	3,200	0.85	92.71	9.96
Connecticut	119	100.00	100.00	3,100	1.31	92.90	12.26
Delaware	68	100.00	100.00	3,200	3.08	93.15	10.86
District of Columbia	102	100.00	100.00	3,000	4.37	89.99	15.28
Florida	233	100.00	100.00	7,100	1.83	92.56	16.13
Georgia	132	100.00	100.00	4,800	2.73	92.94	7.41
Hawaii	81	100.00	100.00	3,300	1.87	91.67	10.69
Idaho	113	100.00	100.00	3,400	1.28	94.32	7.23
Illinois	223	100.00	100.00	4,700	2.32	93.35	11.56
Indiana	113	100.00	100.00	3,200	2.56	92.93	12.15

[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.

NOTE: 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), 2011 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
Iowa	138	100.00	100.00	3,000	1.44	93.13	13.93
Kansas	148	100.00	100.00	3,200	1.33	93.02	9.08
Kentucky	156	100.00	100.00	4,600	3.33	93.50	7.96
Louisiana	163	100.00	100.00	3,100	1.44	92.63	12.76
Maine	143	100.00	100.00	3,100	1.55	91.83	14.44
Maryland	176	99.05	98.82	4,300	6.32	92.30	6.55
Massachusetts	154	99.46	98.47	4,500	4.03	91.65	14.99
Michigan	180	100.00	100.00	4,900	3.60	92.84	7.81
Minnesota	168	100.00	100.00	3,500	2.11	93.04	8.65
Mississippi	121	100.00	100.00	3,100	1.07	93.56	6.19
Missouri	136	100.00	100.00	3,000	1.35	93.71	10.32
Montana	200	99.86	98.41	3,000	1.58	89.95	9.34
Nebraska	169	100.00	100.00	3,000	3.59	93.50	9.13
Nevada	100	99.70	97.35	3,300	3.07	93.73	8.62
New Hampshire	96	100.00	100.00	3,100	1.75	90.87	13.94
New Jersey	116	100.00	100.00	3,000	4.21	92.20	13.88
New Mexico	136	99.09	99.40	4,000	1.96	91.30	9.83
New York	174	99.08	99.67	4,700	1.38	91.04	18.35
North Carolina	159	100.00	100.00	5,000	1.84	91.83	12.36
North Dakota	209	99.99	99.47	2,700	4.26	94.66	8.97
Ohio	194	100.00	100.00	4,300	5.02	92.54	9.68
Oklahoma	149	100.00	100.00	3,000	9.82	92.29	4.19
Oregon	144	99.10	99.26	3,400	1.43	93.12	10.54
Pennsylvania	168	100.00	100.00	4,500	2.43	91.60	13.12
Rhode Island	61	100.00	100.00	3,100	1.26	92.06	13.48
South Carolina	116	100.00	100.00	3,200	3.76	93.57	7.84
South Dakota	261	100.00	100.00	3,500	1.72	94.31	7.29
Tennessee	123	100.00	100.00	3,300	3.77	91.46	7.85
Texas	236	99.09	99.63	9,100	5.16	93.67	5.11
Utah	125	100.00	100.00	3,500	2.72	91.14	8.43
Vermont	124	100.00	100.00	2,400	1.12	93.96	14.82
Virginia	108	100.00	100.00	3,200	2.87	93.47	9.27
Washington	140	100.00	100.00	3,700	1.66	91.82	10.17
West Virginia	117	100.00	100.00	3,200	1.51	93.34	9.24
Wisconsin	179	100.00	100.00	4,100	2.00	92.92	14.18

[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.

NOTE: 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), 2011 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
Wyoming	108	100.00	100.00	2,500	1.25	92.42	10.81
BIE	116	83.16	84.68	1,100	1.85	91.35	11.45
DoDEA [2]	72	98.56	95.31	2,100	2.68	95.40	7.98
Trial Urban (TUDA) Districts							
Albuquerque	34	100.00	100.00	1,400	3.42	89.29	12.40
Atlanta	26	100.00	100.00	1,600	2.50	92.80	8.38
Austin	24	100.00	100.00	1,800	4.64	91.43	9.07
Baltimore City	75	100.00	100.00	1,400	12.50	87.44	7.50
Boston	45	100.00	100.00	1,400	5.79	92.02	18.94
Charlotte	38	100.00	100.00	1,700	1.32	92.39	11.29
Chicago	117	100.00	100.00	2,200	3.33	95.57	15.84
Cleveland	71	100.00	100.00	1,300	5.59	91.35	24.17
Dallas	41	100.00	100.00	1,700	4.89	93.88	6.24
Detroit	63	100.00	100.00	1,900	8.07	84.39	8.23
Fresno	23	100.00	100.00	1,500	1.19	91.56	6.64
Hillsborough	50	100.00	100.00	1,600	1.84	93.23	20.81
Houston	50	100.00	100.00	2,500	5.61	92.79	5.47
Jefferson County, KY	46	100.00	100.00	1,700	3.16	91.83	8.08
Los Angeles	71	100.00	100.00	2,300	1.30	92.39	9.21
Miami	86	100.00	100.00	2,900	1.87	92.99	17.82
Milwaukee	59	100.00	100.00	1,400	4.90	91.92	25.48
New York City	91	100.00	100.00	2,500	1.04	91.35	24.15
Philadelphia	60	100.00	100.00	1,500	6.71	90.86	17.82
San Diego	31	100.00	100.00	1,300	2.83	94.84	7.94
District of Columbia (TUDA)	50	100.00	100.00	1,700	6.58	88.34	18.13
National private	930	74.40	69.89	8,800	0.50	94.65	4.54
Catholic	332	93.23	92.09	4,600	0.53	95.07	3.32
Non-Catholic private	598	57.54	58.75	4,200	0.47	94.26	5.70
Lutheran	141	92.73	91.45	1,000	0.33	96.31	2.13
Conservative Christian	149	72.51	74.78	1,200	0.47	94.46	3.67
Other private	308	45.71	45.85	2,000	0.49	93.90	7.06
Puerto Rico	110	100.00	100.00	5,100	1.03	93.14	17.11

[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.

NOTE: 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), 2011 Assessment.

Participation, Exclusion, and Accommodation Rates for Grade 8 Writing (WCBA) for the 2011 Assessment

The following table displays the school- and student-level response, exclusion, and accommodation rates for the grade 8 writing [WCBA] 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 writing (WCBA) assessment, by school type and jurisdiction: 2011

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	**1,000**	**97.27**	**87.29**	**27,400**	**1.72**	**93.99**	**8.05**
National all[1]	1,000	97.27	87.29	27,400	1.72	93.99	8.05
Northeast all	178	95.36	80.39	4,400	2.08	93.20	11.42
Midwest all	204	98.83	93.61	5,300	1.78	94.19	10.14
South all	400	97.15	86.74	10,700	1.71	94.62	6.48
West all	265	97.43	86.34	7,000	1.38	93.43	5.80
National public	890	99.73	99.86	25,200	1.84	93.99	8.49
National private	157	71.21	68.66	2,200	0.28	93.97	3.01
Catholic	50	95.53	95.23	1,100	0.53	94.72	3.12
Non-Catholic private	107	52.06	56.42	1,100	0.08	93.35	2.92

[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.
NOTE: 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), 2011 Assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/resp_excl_accomm_rates_g8writing_2011.aspx

Participation, Exclusion, and Accommodation Rates for Grade 12 Writing (WCBA) for the 2011 Assessment

The following table displays the school- and student-level response, exclusion, and accommodation rates for the grade 12 writing [WCBA] 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 12 writing (WCBA) assessment, by school type and jurisdiction: 2011

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	**1,400**	**93.52**	**89.26**	**36,500**	**2.33**	**86.97**	**6.65**
National all[1]	1,400	93.52	89.26	36,500	2.33	86.97	6.65
Northeast all	245	91.91	79.32	6,400	2.09	84.18	8.87
Midwest all	266	96.93	95.76	7,300	2.20	86.40	8.79
South all	495	94.66	90.91	13,200	2.72	88.36	4.79
West all	367	89.70	85.00	9,700	2.02	87.73	5.70
National public	1,200	96.04	97.12	33,400	2.52	86.96	6.84
National private	177	67.23	67.45	3,100	0.27	87.16	4.68
Catholic	55	76.60	75.42	1,500	0.11	85.96	3.31
Non-Catholic private	122	58.35	65.42	1,600	0.42	88.34	6.02

[1] Includes national public, national private, Bureau of Indian Education, and Department of Defense Education Activity schools located in the United States.
NOTE: 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), 2011 Writing Assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/resp_excl_accomm_rates_gr12writing_2011.aspx

Participation, Exclusion, and Accommodation Rates for Grade 4 Mathematics for the 2011 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: 2011

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	**9,400**	**97.44**	**91.77**	**242,000**	**2.05**	**94.53**	**11.62**
National all[1]	9,200	97.41	91.65	237,000	2.06	94.53	11.51
Northeast all	1,600	95.19	84.70	37,800	1.78	94.25	15.43
Midwest all	2,400	97.76	91.40	51,900	1.82	94.37	11.48
South all	2,900	97.64	93.31	84,200	2.65	94.73	10.67
West all	2,300	98.27	94.56	60,400	1.54	94.58	10.16
National public	8,200	99.83	99.86	225,000	2.22	94.44	12.12
Alabama	117	98.95	99.76	3,300	1.16	95.00	4.17
Alaska	202	100.00	100.00	3,100	2.81	92.61	17.86
Arizona	127	99.03	99.25	4,200	1.02	94.29	15.32
Arkansas	123	100.00	100.00	3,900	1.00	94.87	14.07
California	292	100.00	100.00	10,100	1.56	95.27	7.07
Colorado	124	100.00	100.00	4,000	1.14	92.20	14.32
Connecticut	116	100.00	100.00	3,400	1.27	93.37	15.55
Delaware	109	100.00	100.00	3,900	3.59	94.12	11.92
District of Columbia	153	100.00	100.00	2,400	5.24	94.54	14.27
Florida	235	100.00	100.00	8,000	1.59	94.51	18.87
Georgia	174	100.00	100.00	6,000	1.65	94.48	10.35
Hawaii	118	100.00	100.00	3,800	1.76	93.35	11.27
Idaho	137	100.00	100.00	4,000	1.21	95.34	8.51
Illinois	193	100.00	100.00	5,600	2.26	93.33	12.84
Indiana	117	100.00	100.00	3,900	2.15	94.69	14.06

[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.

NOTE: 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), 2011 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
Iowa	141	100.00	100.00	3,700	1.41	94.94	14.81
Kansas	148	99.18	99.30	3,500	1.65	94.25	12.54
Kentucky	158	100.00	100.00	5,500	3.08	94.48	8.54
Louisiana	134	100.00	100.00	3,900	1.74	93.65	17.54
Maine	166	100.00	100.00	3,500	1.59	94.48	15.07
Maryland	173	100.00	100.00	5,200	5.61	94.66	11.17
Massachusetts	202	100.00	100.00	5,600	3.17	94.16	15.25
Michigan	172	100.00	100.00	4,700	2.15	94.11	8.75
Minnesota	148	100.00	100.00	4,100	1.50	94.02	12.52
Mississippi	117	100.00	100.00	3,400	0.83	94.99	5.87
Missouri	131	100.00	100.00	3,900	1.65	93.58	10.16
Montana	206	100.00	100.00	3,500	1.52	94.24	7.68
Nebraska	181	100.00	100.00	3,400	1.50	95.57	14.36
Nevada	118	100.00	100.00	4,400	2.29	94.93	22.47
New Hampshire	133	100.00	100.00	3,600	1.74	93.95	15.33
New Jersey	118	99.17	98.83	3,600	3.30	94.52	14.19
New Mexico	155	100.00	100.00	4,700	2.61	93.89	14.86
New York	158	100.00	100.00	5,200	1.33	94.13	20.59
North Carolina	173	100.00	100.00	5,900	1.78	94.01	12.44
North Dakota	272	99.97	99.61	3,400	3.59	95.26	9.31
Ohio	205	100.00	100.00	4,900	2.32	94.03	13.21
Oklahoma	140	100.00	100.00	3,500	8.27	95.42	6.72
Oregon	149	99.08	99.36	4,100	2.67	93.47	14.60
Pennsylvania	167	100.00	100.00	5,200	1.39	94.13	12.58
Rhode Island	125	100.00	100.00	3,500	0.93	94.41	13.31
South Carolina	114	100.00	100.00	3,700	1.27	94.15	10.28
South Dakota	209	100.00	100.00	3,600	1.78	95.31	9.00
Tennessee	119	100.00	100.00	3,900	3.40	93.79	10.37
Texas	306	99.08	99.23	11,000	4.15	95.24	8.31
Utah	130	100.00	100.00	4,500	2.02	93.69	10.20
Vermont	226	100.00	100.00	3,000	1.58	94.12	14.70
Virginia	115	100.00	100.00	4,100	2.08	94.76	11.71
Washington	141	100.00	100.00	4,400	1.90	94.31	13.75
West Virginia	152	100.00	100.00	3,400	1.50	94.58	8.78
Wisconsin	190	100.00	100.00	4,900	1.66	94.98	15.56

[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.

NOTE: 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), 2011 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
Wyoming	202	100.00	100.00	3,200	1.61	93.72	11.97
BIE	135	83.26	84.85	1,200	1.38	91.91	16.51
DoDEA [2]	120	98.91	97.32	3,800	2.74	94.06	10.38
Trial Urban (TUDA) Districts							
Albuquerque	52	100.00	100.00	2,000	2.74	93.15	19.19
Atlanta	66	100.00	100.00	2,100	0.99	96.13	8.38
Austin	55	100.00	100.00	2,000	3.96	94.12	16.52
Baltimore City	70	100.00	100.00	1,700	11.23	93.05	8.11
Boston	86	100.00	100.00	1,900	4.64	93.74	16.87
Charlotte	57	100.00	100.00	1,900	1.15	94.44	11.73
Chicago	95	100.00	100.00	2,700	2.35	94.44	20.21
Cleveland	86	100.00	100.00	1,600	5.59	94.44	20.97
Dallas	55	100.00	100.00	2,000	2.94	96.63	7.70
Detroit	58	100.00	100.00	1,500	5.71	88.79	6.41
Fresno	53	100.00	100.00	2,200	1.27	94.15	6.77
Hillsborough	56	100.00	100.00	1,800	1.68	95.05	25.94
Houston	86	100.00	100.00	3,100	4.12	95.17	14.46
Jefferson County, KY	56	100.00	100.00	2,200	4.91	95.22	9.50
Los Angeles	76	100.00	100.00	2,600	1.76	94.87	9.35
Miami	87	100.00	100.00	2,900	2.85	96.27	22.60
Milwaukee	68	100.00	100.00	1,500	2.75	94.32	27.73
New York City	82	100.00	100.00	2,700	1.63	94.24	27.11
Philadelphia	57	100.00	100.00	1,700	3.88	94.52	15.93
San Diego	54	100.00	100.00	1,900	2.61	94.87	8.20
District of Columbia (TUDA)	106	100.00	100.00	1,600	6.48	93.98	15.15
National private	748	73.51	68.47	6,300	0.30	95.60	4.59
Catholic	264	96.27	95.93	3,400	0.25	95.86	3.75
Non-Catholic private	484	55.34	56.27	2,900	0.35	95.37	5.33
Lutheran	107	94.87	92.38	722	0.36	96.62	3.31
Conservative Christian	123	73.13	70.86	925	0.30	94.16	2.79
Other private	254	42.23	44.57	1,300	0.37	95.74	6.73
Puerto Rico	139	100.00	100.00	4,800	0.47	94.52	22.94

[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.

NOTE: 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), 2011 Assessment.

Participation, Exclusion, and Accommodation Rates for Grade 4 Reading for the 2011 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 reading assessment, by school type and jurisdiction: 2011

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	**9,200**	**97.41**	**91.65**	**245,000**	**3.65**	**94.57**	**9.64**
National all[1]	9,200	97.41	91.65	245,000	3.65	94.57	9.64
Northeast all	1,600	95.19	84.70	39,100	3.75	94.18	13.51
Midwest all	2,400	97.76	91.40	53,700	2.45	94.61	10.44
South all	2,900	97.64	93.31	87,100	5.19	94.61	8.07
West all	2,300	98.27	94.56	62,600	2.19	94.76	8.75
National public	8,200	99.83	99.86	233,000	3.92	94.55	10.12
Alabama	117	98.95	99.76	3,400	2.27	95.31	3.52
Alaska	202	100.00	100.00	3,200	2.03	92.56	19.65
Arizona	127	99.03	99.25	4,400	1.42	94.43	12.84
Arkansas	123	100.00	100.00	4,100	1.26	94.70	13.45
California	292	100.00	100.00	10,500	2.21	95.22	6.10
Colorado	124	100.00	100.00	4,100	1.43	92.67	13.44
Connecticut	116	100.00	100.00	3,500	2.20	93.99	14.80
Delaware	109	100.00	100.00	4,000	6.98	95.06	7.62
District of Columbia	153	100.00	100.00	2,400	3.26	94.66	16.24
Florida	235	100.00	100.00	8,300	2.17	94.55	18.02
Georgia	174	100.00	100.00	6,300	6.31	94.42	5.64
Hawaii	118	100.00	100.00	3,900	2.27	93.39	11.39
Idaho	137	100.00	100.00	4,100	1.81	95.46	7.40
Illinois	193	100.00	100.00	5,800	1.63	93.82	13.30
Indiana	117	100.00	100.00	4,000	1.20	95.23	14.49

[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.

[2] Department of Defense Education Activity.

NOTE: 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), 2011 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
Iowa	141	100.00	100.00	3,800	1.01	95.68	14.99
Kansas	148	99.18	99.30	3,600	2.21	95.16	11.63
Kentucky	158	100.00	100.00	5,700	8.73	94.41	3.78
Louisiana	134	100.00	100.00	4,000	1.33	93.88	17.26
Maine	166	100.00	100.00	3,700	1.56	93.86	14.74
Maryland	173	100.00	100.00	5,400	10.33	94.44	6.72
Massachusetts	202	100.00	100.00	5,800	5.69	94.48	12.31
Michigan	172	100.00	100.00	4,900	3.52	94.40	7.10
Minnesota	148	100.00	100.00	4,300	1.56	94.46	10.47
Mississippi	117	100.00	100.00	3,500	1.06	93.78	5.60
Missouri	131	100.00	100.00	4,100	1.65	94.56	9.29
Montana	206	100.00	100.00	3,600	4.25	93.94	5.13
Nebraska	181	100.00	100.00	3,600	4.33	95.31	10.97
Nevada	118	100.00	100.00	4,500	1.12	95.59	17.76
New Hampshire	133	100.00	100.00	3,700	2.78	93.93	13.85
New Jersey	118	99.17	98.83	3,800	9.09	94.76	8.85
New Mexico	155	100.00	100.00	4,900	5.71	93.43	10.06
New York	158	100.00	100.00	5,400	2.56	93.75	19.39
North Carolina	173	100.00	100.00	6,100	2.21	93.81	12.12
North Dakota	272	99.97	99.61	3,600	6.49	96.01	6.34
Ohio	205	100.00	100.00	5,000	5.77	94.22	9.43
Oklahoma	140	100.00	100.00	3,600	4.97	95.14	9.30
Oregon	149	99.08	99.36	4,300	2.62	94.58	12.70
Pennsylvania	167	100.00	100.00	5,300	2.91	94.28	11.17
Rhode Island	125	100.00	100.00	3,600	2.06	95.01	12.39
South Carolina	114	100.00	100.00	3,900	2.74	94.30	7.26
South Dakota	209	100.00	100.00	3,700	3.18	95.71	7.61
Tennessee	119	100.00	100.00	4,000	7.05	94.71	6.82
Texas	306	99.08	99.23	11,300	9.93	94.83	3.14
Utah	130	100.00	100.00	4,700	4.14	94.15	7.66
Vermont	226	100.00	100.00	3,100	2.38	93.54	13.64
Virginia	115	100.00	100.00	4,200	2.79	94.73	9.77
Washington	141	100.00	100.00	4,500	2.84	95.44	11.67
West Virginia	152	100.00	100.00	3,500	1.70	95.16	7.94
Wisconsin	190	100.00	100.00	5,100	1.86	94.53	15.95

[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.

[2] Department of Defense Education Activity.

NOTE: 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), 2011 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
Wyoming	202	100.00	100.00	3,400	1.98	94.66	12.48
BIE	135	83.26	84.85	1,300	1.62	90.96	14.88
DoDEA [2]	120	98.91	97.32	4,000	6.74	94.10	6.70
Trial Urban (TUDA) Districts							
Albuquerque	52	100.00	100.00	2,100	5.11	92.87	12.48
Atlanta	66	100.00	100.00	2,200	3.81	96.23	5.57
Austin	55	100.00	100.00	2,100	16.49	94.26	4.50
Baltimore City	70	100.00	100.00	1,700	16.89	92.70	3.44
Boston	86	100.00	100.00	2,000	8.06	94.48	14.36
Charlotte	57	100.00	100.00	2,000	1.63	94.57	10.24
Chicago	95	100.00	100.00	2,700	2.10	95.27	19.13
Cleveland	86	100.00	100.00	1,600	5.41	93.03	20.79
Dallas	55	100.00	100.00	2,000	18.48	95.49	2.80
Detroit	58	100.00	100.00	1,500	7.01	88.99	5.39
Fresno	53	100.00	100.00	2,200	2.30	93.71	6.01
Hillsborough	56	100.00	100.00	1,900	2.55	94.61	24.39
Houston	86	100.00	100.00	3,200	14.45	95.27	3.57
Jefferson County, KY	56	100.00	100.00	2,200	9.61	94.80	5.11
Los Angeles	76	100.00	100.00	2,700	1.89	95.11	8.93
Miami	87	100.00	100.00	3,000	3.84	95.68	21.80
Milwaukee	68	100.00	100.00	1,600	2.67	94.75	28.77
New York City	82	100.00	100.00	2,800	2.49	93.01	26.24
Philadelphia	57	100.00	100.00	1,800	3.41	94.40	16.23
San Diego	54	100.00	100.00	2,000	3.61	95.08	7.29
District of Columbia (TUDA)	106	100.00	100.00	1,700	3.93	94.99	18.04
National private	748	73.51	68.47	6,500	0.51	94.91	4.23
Catholic	264	96.27	95.93	3,500	0.45	95.49	3.52
Non-Catholic private	484	55.34	56.27	3,000	0.56	94.41	4.84
Lutheran	107	94.87	92.38	753	1.17	96.23	2.69
Conservative Christian	123	73.13	70.86	950	0.00	95.40	1.96
Other private	254	42.23	44.57	1,300	0.74	93.71	6.48

[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.

[2] Department of Defense Education Activity.

NOTE: 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), 2011 Assessment.

Participation, Exclusion, and Accommodation Rates for Grade 8 Reading for the 2011 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 reading assessment, by school type and jurisdiction: 2011

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	**8,600**	**97.54**	**88.31**	**197,000**	**3.23**	**93.01**	**8.58**
National all[1]	8,600	97.54	88.31	197,000	3.23	93.01	8.58
Northeast all	1,400	95.45	79.60	32,200	3.65	92.23	12.98
Midwest all	2,400	98.76	91.49	44,400	2.87	93.60	9.34
South all	2,700	97.65	89.96	70,800	4.00	93.09	7.01
West all	2,100	97.80	89.74	48,500	2.07	92.92	7.12
National public	7,500	99.79	99.76	186,000	3.48	92.84	8.97
Alabama	125	100.00	100.00	3,000	2.07	94.17	3.84
Alaska	167	99.90	97.93	2,900	1.82	91.24	15.69
Arizona	136	99.02	99.08	3,100	1.17	93.72	8.48
Arkansas	126	100.00	100.00	3,100	1.50	93.81	11.47
California	257	100.00	100.00	8,100	2.17	93.19	6.31
Colorado	130	99.87	97.31	3,100	1.57	92.12	9.83
Connecticut	119	100.00	100.00	2,900	2.25	92.28	12.02
Delaware	68	100.00	100.00	3,100	5.25	93.01	8.95
District of Columbia	102	100.00	100.00	2,900	2.84	89.51	16.48
Florida	233	100.00	100.00	6,900	2.33	91.62	15.10
Georgia	132	100.00	100.00	4,600	4.40	93.50	6.27
Hawaii	81	100.00	100.00	3,200	2.21	92.40	10.01
Idaho	113	100.00	100.00	3,200	1.77	94.08	6.24
Illinois	223	100.00	100.00	4,600	1.62	93.67	11.91
Indiana	113	100.00	100.00	3,100	2.08	92.91	12.55

[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.

[2] Department of Defense Education Activity.

NOTE: 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), 2011 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
Iowa	138	100.00	100.00	2,900	0.75	92.53	13.68
Kansas	148	100.00	100.00	3,100	1.90	93.46	8.49
Kentucky	156	100.00	100.00	4,400	7.22	94.27	4.21
Louisiana	163	100.00	100.00	3,000	1.00	92.69	13.36
Maine	143	100.00	100.00	3,000	1.73	92.31	14.05
Maryland	176	99.05	98.82	4,200	8.43	91.82	4.03
Massachusetts	154	99.46	98.47	4,400	6.32	92.18	12.42
Michigan	180	100.00	100.00	4,700	4.83	93.15	6.58
Minnesota	168	100.00	100.00	3,400	2.84	92.58	7.43
Mississippi	121	100.00	100.00	3,000	0.96	92.33	5.63
Missouri	136	100.00	100.00	2,800	1.39	94.09	10.63
Montana	200	99.86	98.41	2,900	4.03	92.04	6.47
Nebraska	169	100.00	100.00	2,900	4.73	93.78	7.42
Nevada	100	99.70	97.35	3,200	1.93	92.86	8.87
New Hampshire	96	100.00	100.00	3,000	4.15	92.22	11.57
New Jersey	116	100.00	100.00	2,900	7.08	92.32	10.71
New Mexico	136	99.09	99.40	3,900	5.74	91.29	6.30
New York	174	99.08	99.67	4,600	3.10	91.33	16.61
North Carolina	159	100.00	100.00	4,900	2.06	92.09	11.66
North Dakota	209	99.99	99.47	2,600	7.90	93.50	5.62
Ohio	194	100.00	100.00	4,200	5.75	93.25	8.72
Oklahoma	149	100.00	100.00	2,900	4.34	92.52	9.37
Oregon	144	99.10	99.26	3,300	2.13	92.34	8.93
Pennsylvania	168	100.00	100.00	4,400	3.10	91.91	12.73
Rhode Island	61	100.00	100.00	3,000	1.16	92.66	13.70
South Carolina	116	100.00	100.00	3,100	5.30	93.72	4.14
South Dakota	261	100.00	100.00	3,300	3.22	94.70	5.66
Tennessee	123	100.00	100.00	3,200	6.31	91.99	4.63
Texas	236	99.09	99.63	8,800	6.04	93.74	2.91
Utah	125	100.00	100.00	3,400	3.70	92.06	6.05
Vermont	124	100.00	100.00	2,300	2.76	93.05	12.94
Virginia	108	100.00	100.00	3,100	3.62	93.66	7.73
Washington	140	100.00	100.00	3,600	2.11	92.07	9.42
West Virginia	117	100.00	100.00	3,100	1.45	92.44	7.19
Wisconsin	179	100.00	100.00	4,000	2.20	93.81	13.64

[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.

[2] Department of Defense Education Activity.

NOTE: 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), 2011 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
Wyoming	108	100.00	100.00	2,400	1.96	92.67	10.77
BIE	116	83.16	84.68	1,100	2.01	89.62	12.89
DoDEA [2]	72	98.56	95.31	2,000	3.26	91.78	7.80
Trial Urban (TUDA) Districts							
Albuquerque	34	100.00	100.00	1,400	7.29	88.93	9.06
Atlanta	26	100.00	100.00	1,600	3.55	92.38	6.29
Austin	24	100.00	100.00	1,700	9.01	93.15	4.96
Baltimore City	75	100.00	100.00	1,300	16.95	88.94	3.31
Boston	45	100.00	100.00	1,400	9.66	89.97	14.19
Charlotte	38	100.00	100.00	1,600	2.04	92.97	9.66
Chicago	117	100.00	100.00	2,100	2.28	94.92	15.81
Cleveland	71	100.00	100.00	1,200	5.21	91.23	24.77
Dallas	41	100.00	100.00	1,600	5.95	92.60	4.86
Detroit	63	100.00	100.00	1,900	7.96	85.41	7.95
Fresno	23	100.00	100.00	1,500	1.90	92.21	6.37
Hillsborough	50	100.00	100.00	1,500	1.77	94.45	20.97
Houston	50	100.00	100.00	2,500	6.40	94.12	3.87
Jefferson County, KY	46	100.00	100.00	1,600	6.83	91.58	6.06
Los Angeles	71	100.00	100.00	2,300	2.04	91.67	8.84
Miami	86	100.00	100.00	2,800	3.77	92.91	15.63
Milwaukee	59	100.00	100.00	1,400	3.33	90.89	27.81
New York City	91	100.00	100.00	2,500	2.63	91.54	22.49
Philadelphia	60	100.00	100.00	1,500	4.69	91.11	19.54
San Diego	31	100.00	100.00	1,300	1.49	95.63	9.03
District of Columbia (TUDA)	50	100.00	100.00	1,600	3.86	87.69	19.31
National private	930	74.40	69.89	8,600	0.47	94.77	4.32
Catholic	332	93.23	92.09	4,500	0.38	95.36	3.29
Non-Catholic private	598	57.54	58.75	4,100	0.56	94.22	5.29
Lutheran	141	92.73	91.45	1,000	0.48	95.12	2.79
Conservative Christian	149	72.51	74.78	1,200	0.29	93.55	2.00
Other private	308	45.71	45.85	1,900	0.69	94.39	7.03

[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.

[2] Department of Defense Education Activity.

NOTE: 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), 2011 Assessment.

Participation, Exclusion, and Accommodation Rates for Grade 8 Science for the 2011 Assessment

The following table displays the school- and student-level response, exclusion, and accommodation rates for the grade 8 science 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 science assessment, by school type and jurisdiction: 2011

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	**8,600**	**97.32**	**88.19**	**141,000**	**1.57**	**92.87**	**10.58**
National all[1]	8,600	97.32	88.19	141,000	1.57	92.87	10.58
Northeast all	1,400	95.45	79.60	24,000	1.36	92.02	15.71
Midwest all	2,400	98.76	91.49	32,400	1.60	93.40	10.41
South all	2,700	97.65	89.96	47,700	1.70	93.10	9.71
West all	2,100	96.86	89.18	35,100	1.49	92.68	8.38
National public	7,500	99.54	99.57	138,000	1.69	92.81	11.10
Alabama	125	100.00	100.00	2,600	1.08	93.22	4.13
Alaska	167	99.90	97.93	2,400	1.08	89.90	16.37
Arizona	136	99.02	99.08	2,700	0.88	93.19	9.33
Arkansas	126	100.00	100.00	2,700	0.95	94.08	11.63
California	257	100.00	100.00	2,800	1.76	92.96	7.76
Colorado	130	84.41	87.04	2,200	0.90	92.56	10.32
Connecticut	119	100.00	100.00	2,500	1.32	91.41	12.55
Delaware	68	100.00	100.00	2,600	1.69	92.10	12.20
District of Columbia	102	100.00	100.00	3,000	1.49	87.56	17.52
Florida	233	100.00	100.00	2,700	1.24	93.10	16.28
Georgia	132	100.00	100.00	2,800	1.55	92.81	8.43
Hawaii	81	100.00	100.00	2,700	1.99	92.59	10.80
Idaho	113	100.00	100.00	2,800	1.46	93.06	6.66
Illinois	223	100.00	100.00	3,900	1.14	94.03	12.40

[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.

[2] Department of Defense Education Activity.

NOTE: 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), 2011 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
Indiana	113	100.00	100.00	2,600	1.29	93.83	12.92
Iowa	138	100.00	100.00	2,500	0.94	92.75	14.34
Kansas	148	100.00	100.00	2,600	1.41	94.45	9.06
Kentucky	156	100.00	100.00	3,800	2.72	93.04	8.19
Louisiana	163	100.00	100.00	2,600	1.21	93.36	13.21
Maine	143	100.00	100.00	2,600	1.83	92.66	13.97
Maryland	176	99.05	98.82	2,600	1.93	92.54	10.87
Massachusetts	154	99.46	98.47	2,700	3.20	92.20	15.97
Michigan	180	100.00	100.00	2,600	2.74	92.28	8.36
Minnesota	168	100.00	100.00	2,900	1.96	92.13	8.48
Mississippi	121	100.00	100.00	2,500	0.92	92.49	6.22
Missouri	136	100.00	100.00	2,400	1.23	93.44	9.93
Montana	200	99.86	98.41	2,500	1.53	91.02	9.08
Nebraska	169	100.00	100.00	2,500	1.44	94.57	11.59
Nevada	100	99.70	97.35	2,700	1.23	93.07	11.17
New Hampshire	96	100.00	100.00	2,600	2.15	90.78	13.11
New Jersey	116	100.00	100.00	2,500	1.22	91.77	17.29
New Mexico	136	99.09	99.40	3,300	1.75	91.94	10.28
New York	174	99.08	99.67	3,900	1.39	91.24	18.41
North Carolina	159	100.00	100.00	2,900	1.61	92.21	12.12
North Dakota	209	99.99	99.47	2,200	3.22	94.63	10.14
Ohio	194	100.00	100.00	2,600	2.13	92.62	12.38
Oklahoma	149	100.00	100.00	2,400	2.86	92.26	10.04
Oregon	144	99.10	99.26	2,800	1.55	92.66	10.18
Pennsylvania	168	100.00	100.00	2,600	1.03	93.28	14.73
Rhode Island	61	100.00	100.00	2,600	0.65	92.15	14.31
South Carolina	116	100.00	100.00	2,700	1.19	94.22	9.39
South Dakota	261	100.00	100.00	2,800	1.22	95.08	8.01
Tennessee	123	100.00	100.00	2,700	1.43	92.35	10.05
Texas	236	99.09	99.63	3,200	2.36	93.04	7.69
Utah	125	100.00	100.00	2,900	1.83	91.80	9.47
Vermont	124	100.00	100.00	2,000	1.39	93.95	14.15
Virginia	108	100.00	100.00	2,600	2.68	94.01	9.96
Washington	140	100.00	100.00	3,000	1.88	91.86	9.68
West Virginia	117	100.00	100.00	2,700	1.60	93.48	9.14

[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.

[2] Department of Defense Education Activity.

NOTE: 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), 2011 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
Wisconsin	179	100.00	100.00	2,400	1.91	93.21	13.62
Wyoming	108	100.00	100.00	2,000	1.31	92.27	11.27
BIE	116	83.16	84.68	112	0.00	88.14	6.07
DoDEA [2]	72	98.56	95.31	1,700	1.33	94.33	9.55
National private	930	74.40	69.89	903	0.18	93.65	4.74
Catholic	332	93.23	92.09	476	0.35	94.27	4.65
Non-Catholic private	598	57.54	58.75	427	0.00	93.03	4.84

[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.

[2] Department of Defense Education Activity.

NOTE: 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), 2011 Assessment.

http://nces.ed.gov/nationsreportcard/tdw/weighting/2011/resp_excl_accomm_rates_g8science_2011.aspx

Nonresponse Bias Analyses for the 2011 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 2011 assessment indicated a need for school nonresponse bias analyses for private school samples in grades 4 and 8 (operational subjects), for public school samples for the Bureau of Indian Education (BIE) in grades 4 and 8, and for private school samples in grades 8 and 12 (computer-based writing). Student nonresponse bias analyses were necessary for students in Detroit for grade 8 public schools (mathematics). Additionally, a student nonresponse bias analysis was required to handle the special case of session nonresponse in the science sample in grade 8 Colorado public schools. Thus, six separate school-level analyses and two 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 2011 NRBA report (818KB).

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, and size of school (categorical).

Based on the school characteristics available, for the private school samples at grades 8 and 12, 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 4 private school samples. 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 overrepresent 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. 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. For Colorado, additional school characteristics were examined: Census region, urban-centric locale, size of school (categorical), and state-based achievement (categorical).

Based on the student characteristics available, for the grade 8 Detroit student samples, there does not appear to be evidence of substantial potential bias resulting from student nonresponse. The same result can be concluded for grade 8 Colorado student samples, when considering student characteristics. However, analyses of the school characteristics suggest that a potential for nonresponse bias remains. Please see the full report for more details.

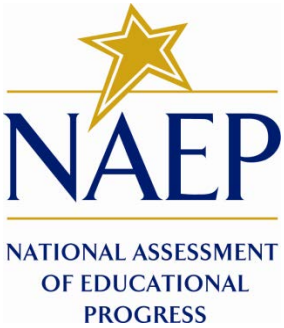
NATIONAL CENTER FOR EDUCATION STATISTICS
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

Appendix C

2018 Sampling Memo

National Assessment of Educational Progress (NAEP)
2018 and 2019

OMB# 1850-0928 v.5



March 2017

Westat | An Employee-Owned Research Corporation

1600 Research Boulevard
Rockville, MD 20850-3129
tel: 301-251-1500
fax: 301-294-2040
www.westat.com

Date: February 22, 2017 **Memo:** 2018-m03v01psu/m01v01s

To:

Bill Ward, NCES	David Freund, ETS
Amy Dresher, ETS	Cathy White, Pearson
Ed Kulick, ETS	Keith Rust
Dianne Walsh	Dwight Brock
Rick Rogers	Lauren Byrne
William Wall	John Burke
Rob Dymowski	Joel Wakesberg
Lisa Rodriguez	Lloyd Hicks
Chris Averett	Sipeng Wang
Kavemuii Murangi	Jason Schuknecht

From: Leslie Wallace

Reviewer: David Hubble

Subject: Sample Design for 2018 NAEP – Overview

I. Introduction

For 2018, the sample design involves several components, all of which are national assessments of one kind or another.

1. Operational Digitally-based Assessment (DBA) in Civics at grade 8;
2. Operational DBA in U.S. History at grade 8;
3. Operational DBA in Geography at grade 8;
4. Operational assessment in Technology and Engineering Literacy (TEL) at grade 8;
5. Pilot tests at grades 4, 8, or 12:
    a. Math DBA at grade 12
    b. Reading DBA at grade 12
    c. Science integrated DBA at grades 4, 8, and 12;
6. Bridge Paper-based Assessment (PBA) in Civics at grade 8;
7. Bridge PBA in U.S. History at grade 8;
8. Bridge PBA in Geography at grade 8;

In addition, the following special studies will be conducted:

9. Oral Reading Fluency (ORF) at grade 4

10. Reading Scenario-based Task (SBT) Special Study at grades 4, 8, and 12.

11. NTPS-NAEP Linking Study at grades 4, 8, and 12. Note that this study will likely involve a subsample of the NAEP schools at each grade. Details are unclear at this time, so the study is not discussed in the remainder of this memo.

The target sample sizes of assessed students for the various components are shown in Table 1 (which also shows an estimate of the required number of participating schools). All of these assessments are to take place in the typical NAEP testing window of late January to early March 2018. Note that the Pilot assessments and special studies are conducted in public schools only, whereas the Social Studies and TEL assessments are conducted in both public and private schools.

Table 1. 2018 NAEP Sample Sizes (Public and Private)

	Session*	Public school students	Private school students	Total students
Grade 4				
Science Integrated DBA (P)	DA	16,100	0	16,100
Reading SBT (S)		2,200	0	2,200
ORF (S)	DB	2,000	0	2,000
Total		20,300	0	20,300
Schools				520
Grade 8				
Civics DBA (O)	DO	7,200	800	8,000
Geography DBA (O)		7,200	800	8,000
U.S. History DBA (O)		9,000	1,000	10,000
Civics PBA (B)	PC	7,200	800	8,000
Geography PBA (B)	PH	7,200	800	8,000
U.S. History PBA (B)		9,000	1,000	10,000
TEL (O)	DL	14,400	1,600	16,000
Science Integrated DBA (P)	DA	15,900	0	15,900
Reading SBT (S)		2,200	0	2,200
Total		79,300	6,800	86,100
Schools				2,050
Grade 12				
Math DBA (P)	DA	10,500	0	10,500
Reading DBA (P)		4,500	0	4,500
Science Integrated DBA (P)		17,500	0	17,500
Reading SBT (S)		2,200	0	2,200
Total		34,700	0	34,700
Schools				870
GRAND TOTAL		134,300	6,800	141,100
Schools				3,440

(O) = Operational, (B) = Bridge, (P) = Pilot, and (S) = Special Study

* The session designations are not final.

II. Assessment Types

From a sampling and operations point of view, many types of assessment sessions can be distinguished. The detailed target counts of assessed students are provided in Table 1.

1. The DBA Civics, U.S. History, and Geography spiral is only at grade 8. This spiral must be assessed in a different physical session from the others, but will be in the same schools as the PBA Civics, U.S. History, and Geography session types (see immediately below). The session has a target of 26,000 assessed students (8,000 Civics, 10,000 U.S. History and 8,000 Geography).

2. The PBA U.S. History and Geography spiral is only at grade 8. This session spiral has a total target of 18,000 assessed students (10,000 U.S. History and 8,000 Geography). The PBA Civics session is also only at grade 8. This session has a target of 8,000 assessed students. Both of these PBA Social Studies sessions will be conducted in the same schools as the DBA Social Studies session described above.

3. The Technology and Engineering Literacy (TEL) operational assessment for grade 8 will be computer delivered. Because of the different delivery method, this assessment must be in separate sessions. In fact, out of concern for overburdening schools conducting Social Studies assessments, an additional set of Primary Sampling Units (PSUs), with minimum overlap with Social Studies PSUs, will be used for conducting TEL with a target of 16,000 assessed students.

4. The DBA Pilot/SBT at grades 4, 8, and 12 will be tablet delivered and conducted only in public schools. Because of the different delivery method and out of concern for overburdening schools conducting social studies and TEL assessments, a third set of PSUs will be used for conducting the DBA Pilot/SBT. One session will be administered at each grade. At grades 4 and 8, the session will be a DBA Science Pilot/SBT spiral, with targets of 18,300 assessed students at grade 4 and 18,100 assessed students at grade 8. At grade 12, the session will be a DBA Science/Reading/Math/SBT spiral, with a target of 34,700 assessed students.

5. The ORF Special Study at grade 4 will be conducted in some of the DBA Pilot (public) schools. This will be a separate session with a target of 2,000 assessed students.

III. Primary Sampling Units Selection and Overlap Control

There are three separate PSU samples for 2018: one each for Social Studies, TEL, and DBA Pilot/Special Studies. As the U.S. History, Geography, and Civics assessments are national, with a total original sample size of assessed students of about 52,000 at grade 8, for reasons of operational efficiency in conducting the assessments a sample of PSUs was selected, and all sampled schools will be drawn from within the sampled PSUs. With a smaller sample size of about 16,000 assessed students for the computer delivered TEL assessment in grade 8, a separate sample of PSUs was selected with the largest PSUs being in common to both PSU samples. Finally, with a total sample size of 73,100 assessed students across grades 4, 8, and 12 for the DBA Pilot and Special Studies, a third sample of PSUs was selected with the largest PSUs being in common in all three PSU samples.

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 Census region 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 Social Studies samples. The 29 largest 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. For the Social Studies sample, 76 such strata were formed and a single PSU was selected from each stratum for a total of 105 PSUs. For the TEL sample, the same certainty PSUs were selected. However, the 76 strata were formed and paired and a single PSU was selected from one stratum in each of the 38 pairs for a total of 67 PSUs. For the DBA Pilot/Special Studies samples, the same certainty PSUs were again selected and the 76 strata were formed and paired. However, a single PSU was selected from each of the strata in the 38 pairs not used for the TEL sample for a total of 67 PSUs. The three PSU samples were selected in such a way as to minimize overlap between them. This was done to reduce the chance that a school is selected for more than one of the Social Studies, TEL, or DBA Pilot/Special Studies assessments. Due to the fact that three PSU samples were selected and that one of them consisted of 105 PSUs, overlap among the three PSU samples, though minimized, was not entirely avoided. Five noncertainty PSUs overlap between the Social Studies and TEL samples, and (a different) five noncertainty PSUs overlap between the Social Studies and DBA Pilot/Special Studies samples.

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. Note that the DBA Pilot and Special Studies designs do not include private school components as the assessment goals could be better met through other means in this case.

Explicit stratification will take place at the PSU level. For schools within PSUs, stratification gains will be achieved by sorting the school file prior to systematic selection. As in past national samples, the expectation is that, within the set of certainty MeSA PSUs within a census region, PSU will not necessarily be the highest level sort variable. Thus, type of location will be used as the primary sort variable. 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, 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 three different kinds of oversampling of public schools. First, in order to increase the likelihood that the results for American Indian/Alaskan Native (AIAN)[1] students can be reported for the operational samples, we will oversample high-AIAN public schools for Social Studies and TEL at grade 8. That is, a public school with over 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. 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 Social Studies and TEL that are not oversampled for AIAN enrollment will be oversampled for Black and Hispanic enrollment. That is, as used in past national assessments, a public school with over 15 percent 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.

The second kind of oversampling to be implemented will be oversampling of public schools based on National School Lunch Program (NSLP) eligibility in order to accommodate the ORF Special Study. That is, for the grade 4 DBA Pilot/Special Study samples, a public school with over 75 percent student eligibility for the NSLP will be given twice the chance of selection of a public school of the same size with a lower percentage of NSLP student eligibility.

The third kind of oversampling to be implemented will be oversampling of high-minority public schools for the grade 8 DBA Pilot/SBT and grade 12 DBA Pilot/SBT samples. That is, as used in past national assessments, a public school with over 15 percent 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 not implement the oversampling of Black and Hispanic students at the student level in schools not being oversampled at the school level as has been done in the past because such student-level oversampling is incompatible with the digital mode of assessment.

The updated preliminary 2015/16 CCD and the updated 2015/16 PSS school files were approved for use by NCES. They serve as the basis for the public and private school frames for the 2018 NAEP.

V. New Schools

To compensate for the fact that the CCD file used to create the NAEP public school sampling frames is out of date at the time of frame construction, we will supplement the samples for the Social Studies and TEL assessments with a sample of new public schools for grade 8. New school samples will not be developed for the private school samples or the DBA Pilot/Special Studies samples.

The new school samples will be drawn using a two-stage design. At the first stage, a national sample of school districts will be selected from the Social Studies and TEL sample PSUs. The sampled districts will be asked to review lists of their respective schools and identify new schools. Frames of new schools for grade 8 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.

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

In keeping with the efforts at the PSU level to reduce potential overlap between the Social Studies and TEL samples, methods will be employed to reduce overlap during sample school selection within the PSUs that contain more than one sample. In addition to the overlap control efforts between the Social Studies and TEL samples, methods will be employed to reduce overlap during NAEP school selection with the International Computer and Information Literacy Survey (ICILS). The ICILS is a national sample of schools at grade 8 (not PSU-based) and is being conducted in the spring of 2018. With this approach we expect it to be possible to avoid any school overlap among the Social Studies, TEL, and ICILS school samples at grade 8.

Concurrent with the selection of the Social Studies and TEL samples, the DBA Pilot/Special Studies schools will be selected independently. No effort will be made to minimize overlap between the DBA Pilot/Special Studies samples at grade 8 and the Social Studies, TEL, or ICILS samples because the level of effort required to implement overlap control among four samples in order to avoid a few schools that might overlap is not justified for a pilot or special study. DBA Pilot/Special Studies schools at grade 8 that are also selected for Social Studies, TEL, or ICILS will be treated as nonrespondents and their substitutes will be recruited. Within the DBA Pilot/Special Studies sample, schools may sometimes be selected to participate at more than one grade.

The Keyfitz method will be used to compute conditional probabilities to reduce the overlap between the samples within grade 8. That is, in the Social Studies PSUs, the conditional probabilities of selection for the Social Studies schools will be based on the ICILS school sampling outcome. Also, in the 33 TEL noncertainty PSUs that do not overlap with Social Studies, the conditional probabilities of selection for the TEL schools will be based on the ICILS school sampling outcome. Finally, in the 29 certainty PSUs and the 5 noncertainty PSUs that overlap between Social Studies and TEL, the conditional probabilities of selection for the TEL schools will be based on the Social Studies and ICILS school sampling outcomes. Specifically, this will be done to reduce overlap between Social Studies and TEL sample schools, between Social Studies and ICILS sample schools, and between TEL and ICILS sample schools.

VII. Substitute Samples

Substitute samples will be selected for each of the 2018 samples in the following order for public schools: Social Studies, TEL, and then DBA Pilot/Special Studies. Within the DBA Pilot/Special Studies sample, the order for selecting substitute schools will be from “oldest” to “youngest”. That is, grade 12, 8, and then 4. The order for selecting substitute samples for private schools will be Social Studies and then TEL. This ordering of samples and grades is necessary since no school can be selected as a substitute more than once. It is more critical for operational samples to precede non-operational ones and higher grades to precede lower grades due to having fewer schools available to serve as substitutes at the higher grades. Selecting substitutes will done separately for both public and private schools. The general steps

for selecting substitutes are to put the school 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 2018 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 grade enrollment value. If estimated grade enrollment of both potential substitute schools differs from the original school by the exact same amount, the selection procedure randomly chooses 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, the few sampled private schools whose school affiliation is unknown will not get substitutes nor could 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

Student sample sizes 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 wish to avoid the situation where all but a few students (e.g., more than 90%, but fewer than 100%) are tested.

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

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

5. We wish to minimize the number of unique combinations of session types in the schools and to avoid three session types in a given sample school.

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

7. For the DBAs, we can use up to 25 tablets in a school at one time.

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

Social Studies: Grade 8 Schools

We will select all students, up to 75. In schools with more than 75 students, we will select 75. There are three session types: DBA Civics/U.S. History/Geography (DBA), PBA Civics (C) and PBA U.S. History/Geography (H/G). The proportion of students assigned is 1/2 for the DBA session, 9/26 for the H/G session and 2/13 for the C session. Schools will be assigned one or two sessions; no schools will be assigned all three sessions. Minimum session size is 12 within schools with 12 or more students. This assignment of the DBA, C, and H/G sessions, based on the number of students in the school, is detailed in Table 2.

Table 2. Grade 8 social studies school session allocations and proportions

	Grade 8		
Enrollment size	1 to 23	24 to 50	51 and higher
Probability of being assigned DBA and C	0	4/13	6/13
Proportion of sample students assigned to DBA (in schools with DBA and C)	NA	1/2	2/3
Probability of being assigned DBA and H/G with two DBA sessions	0	0	1/26
Proportion of sample students assigned to DBA (in schools with DBA and H/G and two DBA sessions)	0	0	2/3
Probability of being assigned DBA and H/G with one DBA session	0	9/13	1/2
Proportion of sample students assigned to DBA (in schools with DBA and H/G and one DBA session)	NA	1/2	1/3
Probability of being assigned DBA only	1/2	0	0
Probability of being assigned H/G only	9/26	0	0
Probability of being assigned C only	2/13	0	0

TEL: Grade 8 Schools

We will select all students, up to 30. In schools with more than 30 students we will select 30. All students will be assigned to the TEL assessment.

DBA Pilot/SBT/ORF: Grade 4 Schools

There are two session types: DBA Pilot/SBT (DBA) and ORF. The proportion of students assigned is about 9/10 for the DBA session and about 1/10 for the ORF session. Some of the schools will be assigned only DBA. In these schools, we will select all students, up to 50. In schools with more than 50 students we will select 50. All students in these schools will be assigned to the DBA session. The other schools have the potential for doing both DBA and ORF sessions. In these schools we will select all students up to 37. In schools with more than 37 students, we will select 37—25 for DBA and 12 for ORF. Minimum DBA session size is 12 within schools with 12 or more students. Maximum ORF session size is 12. In order to achieve the session size constraints, every school that has the potential for doing both sessions will not always be assigned both sessions. This assignment of the DBA and ORF sessions, based on the number of students in the school, is detailed in Table 3.

Table 3. Grade 4 school session allocations and proportions

	Grade 4				
Enrollment size	1 to 12	13 to 24	25 to 37	38 to 49	50 and higher
Probability of being assigned DBA only (and so selecting up to 50 students)	9/10	4/5	7/10	2/3	17/27
Probability of being assigned DBA and ORF (and so selecting up to 37 students)	0	1/5	3/10	1/3	10/27
Proportion of sample students assigned to DBA (in schools with DBA and ORF session types)	NA	1/2	25/37	25/37	25/37
Probability of being assigned ORF only	1/10	0	0	0	0

DBA Pilot/SBT: Grade 8 and 12 Schools

We will select all students, up to 50. In schools with more than 50 students we will select 50. All students will be assigned to the DBA Pilot/SBT session for grade 8 and 12.

IX. Weighting Requirements

Social Studies Samples

These samples will have a single set of weights for each subject (DBA Civics, DBA U.S. History, DBA Geography, PBA Civics, PBA U.S. History, and PBA Geography at grade 8) 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 grade and public/private.

For each subject, we will also provide weights for the combined DBA and PBA samples (Civics, U.S. History, and Geography at grade 8).

TEL Sample

As with the Social Studies samples, the TEL sample at grade 8 will be fully weighted.

Pilot Test and Special Studies Samples

As with the Social Studies and TEL samples, the ORF sample at grade 4 will be fully weighted.

We will not weight the students in the Pilot samples or Reading SBT studies at grades 4, 8, and 12. However, preliminary weights will be available for these samples.

Endnotes

[1]As states, districts, and schools are only required to report race/ethnicity data at the 7-category level (and specifically because this is how the data are recorded on the Common Core of Data, used as the sampling frame), the data used to oversample high AIAN percentage schools are the percent of students who are non-Hispanic AIAN students with no other race category. This is also the basis for the primary reporting results of AIAN students for NAEP. Note that the oversampling is at the school level, so that students who report multiple races, including AIAN, who are in schools with a high percentage of AIAN students, will also be oversampled. However, as noted, current NAEP primary reporting practices will not report such students as AIAN.