New Mexico

School Grading Technical Guide

Calculation and Business Rules

Hanna Skandera
Secretary Designate

Pete Goldschmidt, Ph.D.
Director of Assessment and Accountability

Data Planning & Analysis
New Mexico Public Education Department
Cindy Gregory, Ph.D. Chief Statistician
cindy.gregory@state.nm.us

Preface

School Grading was mandated by New Mexico state lawmakers in 2011 where basic requirements were established for schools to achieve an A, B, C, D, or F for annual accountability [§22-2-1, §22-2-2, and §22-2E-1 to §22-2E-4] [6.19.8.1 NMAC – N, 12-15-11]. This School Grading Technical Guide supplements these documents by providing detailed decision rules for each indicator, statistical treatment of data, and calculation parameters. These business rules apply to New Mexico public schools, and do not apply to private, Bureau of Indian Education (BIE), or home schools that are not within the jurisdiction of the New Mexico Public Education Department (PED).

New Mexico's school grading model is currently being reviewed by the U.S. Department of Education to serve as the state's ESEA accountability method for future years. The state's ESEA Flexibility Request is available on the website at http://www.ped.state.nm.us/waiver/index.html and outlines the underlying framework for the system, in particular that schools will be monitored on three general factors: current performance; growth; and other academic factors considered important for student learning.

In this inaugural year of school grading (2011-12) certain data constraints apply:

1) Preliminary grades will be based on data based on the 2008-2009, 2009-2010, and 2010-2011 school years. School grades for subsequent years will be based on the current school year (e.g. 2011-2012) and two prior years of data.
2) Graduation rates are restricted to 4-year and 5-year cohort rates; 6-year rates will be added in subsequent years as data become available.
3) Achievement, current standing and growth components are restricted to reading and mathematics. If resources become available in future years to restore science or other assessments, these content areas may be added to the model.

New Mexico is adopting the nationally recognized *Common Core State Standards* (National Governors Association Center for Best Practices and Council of Chief State School Officers), and is actively participating in the *Partnership for Assessment of Readiness for College and Careers* (PARCC) consortium of states. At the time of implementation, anticipated in 2015, specific features of the school grading system may require change (e.g. cut points adjustment to account for new assessments, using additional grades in HS to calculate student growth, etc); however, the underlying framework (i.e. school grades based on current standing, growth, and other indicators) will remain the same.

Please submit questions or requests for revision and clarification to:

Cindy Gregory, PhD
Chief Statistician, Public Education Department
Jerry Apodaca Bldg, Rm 126
300 Don Gaspar
Santa Fe, New Mexico 87501-2786
Cindy.Gregory@state.nm.us

Revision History

Date	Description of Major Changes	Reference	Author
	(intentionally blank)		

I. Definitions and Abbreviations

Terms used in this document and their meaning to school grading

A. **FAY** (Full Academic Year) indicates whether a student has been enrolled at a single location for one year. The school or LEA where the student was enrolled is accountable for the student's instructional legacy, since the student was present for one year's worth of growth and learning.

B. **Feeder schools** are schools that have no tested grades. For elementary these are schools with grades K through 2 (testing begins in the 3rd grade). For high schools in school year 2010-11, that includes schools with grades 9, 10, or 12 (high school testing is in grade 11 only). Beginning in 2011-12 high schools will begin testing in the 10th grade as well.

C. **Priority schools** are identified by the combination of school grade and points earned. To be eligible, these schools must receive Title I allocations and also be ranked in the lowest 5% of all schools statewide.

D. **Focus schools** are identified by the combination of school grade and points earned. To be eligible, these schools must receive Title I allocations and also be ranked in the lowest 10% of non-Priority schools statewide.

E. **Reward schools** are identified by the combination of school grade and points earned. To be eligible, these schools must receive Title I allocations and also be ranked in the highest 5% of all schools statewide.

F. **Snapshots** are the fixed dates required for all districts to submit data to the PED data warehouse called STARS (Student Teacher Accountability Reporting System). These dates are fixed at

1. Second Wednesday of October (known as 40th day; abbreviated as 40D)
2. December 1 (known as 80th day; abbreviated as 80D)
3. Second Wednesday of February (known as 120th day; abbreviated as 120D)
4. End of Year, variable but principally in June (known as EOY)
5. Additionally, a specialized snapshot occurs during the terminal week of testing (known as the *Assessment Snapshot*); timing is variable but generally occurs in March and April.

G. **VAM** - value-added modeling isolates the school's contributions to student performance from factors outside the school's control that are known to affect student test performance.

H. **Conditional Status** represents the current standing of a school, acknowledging differences in student factors that are outside of a school's control. This is estimated simultaneously with School Growth using a mixed effects Value Added Model (VAM).

I. **School Growth** represents the change in performance of successive cohorts of students over time. This growth is based on different students each year. This is estimated using a mixed effects VAM.

J. **Student Growth** represents the average of individual student growth over three years (current and two prior years). Student growth is estimated using a mixed effects panel growth model.

K. **Elementary or Middle School (EL/MS)** is defined by the grade span of the school and requires that the EL/MS grading model be applied. Schools with the highest grade of 10 or less are considered elementary/middle for school grading.

L. **High School (HS)** is defined by the highest grade of the school and requires that the HS grading model be applied. Schools with the highest grade of 11 or higher are considered HS for school grading.

M. **One Percent Rule**, set by the US Department of Education, requires that the percentage of students scoring *Proficient* or *Advanced Proficient* on an alternate assessment cannot exceed one percent of the total number of students tested in the LEA. If the LEA violates this rule, a random selection of students equal to the excess above 1%, who took the alternate assessment and scored *Proficient* or *Advanced Proficient* must be converted to not proficient. This rule is applied only to LEA accountability.

N. **Opportunity to Learn** represents:

1. Students attendance; and
2. Scores on a 10 item survey administered students annually during standardized testing. The survey measures the extent to which classroom teachers demonstrate instructional practices known to facilitate student learning.

O. **Accountable school** denotes the location where the student's scores are assigned for accountability. The assignment follows this hierarchy:

1. If FAY=Yes, the accountable school is the FAY school; or
2. IF FAY=No, the accountable school is the location where the student was tested.

P. **Duplicate** refers to two test records that have the same student ID. Duplicates may occur when the same student was tested twice, such as English and Spanish; or when a student moves during the test window and is retested at a different school. More commonly, different students are mistakenly identified as being the same, which happens with the accidental mislabeling of a test, or with an incorrectly bubbled ID. Duplicates also occur when two students within the state are unknowingly sharing the same ID. All of these conditions must be reconciled during biodata review. Each student can contribute only one test score for each content area.

Q. **LEA** (Local Educational Authority) traditionally represents the 89 school districts that manage over 800 schools in New Mexico. More recently, the term also applies to a growing number of State-authorized charter schools that operate independently of any district.

R. **Subgroups** are the student groups disaggregated for accountability reporting, as required by ESEA. A single student can contribute to several subgroups, and only A through l and 2 are used in school grading:

a) All students (reported as "All Students")
b) Caucasian/White-Non Hispanic (reported as "Caucasian")
c) Black-Non Hispanic (reported as "African-American")
d) Hispanic (reported as "Hispanic")
e) Asian/Pacific Islander (reported as "Asian")
f) Native American (reported as "American Indian")
g) English Language Learners (reported as "English Language Learners"; abbreviated as "ELL" where necessary)
h) Special Ed, Not Gifted (reported as "Students with Disabilities"; abbreviated as "SWD" where necessary)
i) FRL, free or reduced lunch program (reported as "Economically Disadvantaged"; abbreviated "ED" where necessary)
j) Gender (reported as "F" or "M")
k) Migrant (Title 1C) (reported as "Migrant")
l) FAY, full academic year, reported as "Y" or "N"

2. The school grading system identifies a separate subgroup for students who are in the bottom quartile of their school's performance in year one of the three years used to calculate school grades (see IV b).

II. Data Sources

A. **School Attributes**

1. The school file lists all open public schools and locations in New Mexico with enrolled students in any grades K through 12. The purpose of this file is to finalize which individual schools receive a rating, and their characteristics that impact calculations. Occasionally, schools merge, change configuration, change name, or are considered a "program" rather than a school (and vice versa). Such changes are finalized prior to school grading. Each location is classified as:

    a) Public school; if students take the test at a school program where they are enrolled for only part of the day, their home school must be identified and their scores attributed to their home school. Scores of students in transient programs (i.e. programs at different schools in which the student is enrolled for several weeks or months) may also be assigned to their home school, after approval by PED.

    b) Locally-authorized charter school.

    c) State-authorized charter school.

    d) Off-site program (correctional facilities, treatment centers, homebound/hospitalized). Students in Off-Site programs, such as treatment centers, correctional facilities, or hospitals generally are excluded from school level calculations and counted only at the LEA level for rating. However, if the student qualified as FAY at a school prior to entering the program and testing, their test will count at the FAY school. Off-site programs are not rated.

    e) State-supported (Juvenile Justice, School for Visually Impaired, School for Deaf). New Mexico Military Institute is exempted from rating by statute.

    Additionally, schools are characterized by:

    f) Title I status (S=Schoolwide, T=Targeted, N=Not receiving Title I funds)

    g) Alternate school (Y/N)

    h) Level (elementary/middle, or high school)

    i) Tier I (Y/N)

    j) Tier II (Y/N)

    k) Tier III (Y/N)

    l) SIG – receiving federal School Improvement Grant funding

    m) New or reorganized (impacts inheritance, FAY, and other calculations)

2. Attendance is extracted from the data submissions by districts at 40D, 80D, and 120D snapshots of the current year. A rate is computed for every subgroup, and includes all grades K-12 that are served by the school. Feeder schools must be included. The calculation of attendance is fully covered in PED's *Attendance Technical Guide.*

3. Graduation is provided by the Data Analysis and Planning unit at PED. The file lists rates and counts by subgroup, school, and LEA for 4-year, 5-year, and beginning in 2012 6-year cohorts. Schools with any high school grade (9, 10, 11, or 12) will receive a rate. The calculation of graduation is fully covered in PED's *Graduation Technical Manual.*

4. School rating and figures from prior years are required for the current year's calculations.

B. Student Attributes

1. Mathematics and Reading Proficiencies are supplied by the vendor that administers the standards based assessment to grades 3-8, 10, and 11.
2. Subgroup membership is assigned from the data submitted by districts to the *Assessment Snapshot*, scheduled the final week of standardized testing. Where students are missing from this snapshot, the nearest snapshot date where the student is found will be used.
3. Opportunity to Learn survey item responses are supplied by the vendor that administers the survey during standardized testing.
4. Student/Parent Engagement does not have a specific data file associated. This information arrives through a dossier submitted to PED by the LEA. PED authorities then rate each for point values utilizing a formalized consensus model.
5. ACT student level data are supplied by the vendor annually in summer following all test sessions. The file is limited to students that authorized release of scores to their high school and other post-secondary institutions.
6. SAT student level data are supplied by the vendor annually in summer following all test sessions. The file is limited to students that authorized release of scores to their high school and other post-secondary institutions.
7. PSAT student level data are supplied by the vendor annually in summer following all test sessions. The file is limited to students that authorized release of scores to their high school.
8. AP student level data are supplied by the vendor annually in summer following all test sessions. The file is limited to students that authorized release of scores to their high school and other post-secondary institutions.
9. Dual Credit data are supplied by a cooperative agreement between PED and HED. The data are limited to students who have enrolled and earned credit in post-secondary institutions governed by HED.
10. Career Readiness is partly determined by course enrollments and course grades extracted from data submissions by districts at 40D, 80D, and 120D snapshots of the current and prior years. Definitions established for Carl Perkins Grant funding then classify students as "Concentrators" or "Completers". When completers graduate with a regular diploma they meet the *success* benchmark.

C. LEA Attributes

1. LEA rating and figures from prior years are required for the current year's calculations.
2. Attendance similar to schools is aggregated from data submissions by districts at 40D, 80D, and 120D of the current year.
3. Graduation is provided by the Data Analysis and Planning unit at PED. The file lists rates and counts by subgroup, school, and LEA for 4-year, 5-year, and beginning in 2012 6-year cohorts.

III. Data Validation

A. Verification of preliminary files insures consistency with prior years and completeness. Discrepancies are presented to suppliers of the data source for resolution. Data checks include, but are not limited to:

1. Correct grade ranges for all schools
2. Schools to be rated
3. New or reorganized schools with inherited grading histories or FAY anomalies
4. Notable variation in the size of any subgroup over the prior year
5. Any variation in student subgroup membership among 120D, Assessment, and EOY snapshots
6. Verification of prior year's ratings

7. Verification of the appropriate assignment and completeness of graduation and attendance
8. Verification of level (elementary/middle, or high school)
9. Verification that all rated schools are represented in the vendor test files
10. Verification that all students tested are represented in school rating and GAP results

IV. Conditioning of Data

A. <u>Assessment Scores</u>. A subset of records in the test data file (valid tests), and content within records (Reading, Math), is used for school rating. Results for Spanish, English, and Braille administrations are included, as are both the SBA and NMAPA tests from both vendors. The following guidelines apply to the selection and cleaning of those records.

1. Remove records where *Exclude from Summaries* = Yes. These tests belong to students from non-PED schools.
2. Reconcile Test Completion Code (TC), Scaled Score (SS), Proficiency Level (PL). The TC is a field that is bubbled by the testing administrator at the time of testing to indicate whether the test was successfully administered. Often it is bubbled incorrectly, overridden by scoring (student fails the *Attemptedness Rule*), or overridden by a sanction imposed by PED for a testing irregularity. Because TC is used for various counts, it must match the results of scoring, Scaled Score (SS) and Proficiency Level (PL). This reconciliation is performed by Data Planning and Analysis, and detail can be supplied upon request.
    a) TC=0 Tested All Sessions and received a valid score (SS=0 to 80) (PL=1, 2, 3, or 4))
    b) TC=1 Withdrew before testing; remove test
    c) TC=2 Received a non-allowed modification; invalidate test (SS=99) (PL=5)
    d) TC=3 Exempt from READING (language); remove READING test
    e) TC=4 Medical exemption; remove test
    f) TC=5 Parental refusal; invalidate test (SS=99) (PL=5)
    g) TC=6 Incomplete testing; invalidate test (SS=99) (PL=5)
    h) TC=7 Testing irregularity; invalidate test (SS=99) (PL=5)
    i) TC=8 Absent; invalidate test (SS=99) (PL=5)
3. Note that a single student can have a valid MATH test (TC=0) and an invalid READING test (TC=5). This impacts participation rates for each content area.
4. Note that a student can take the test in more than one school (i.e. MATH in school X, and READING in school Z). Special rules apply and are explained in *Calculations*.
5. Limit records to students in eligible grades (3 through 8, 10, and 11). Occasionally students in other grades are tested because assessments are also used in high school graduation.
6. Use enrollment data to manually update missing values in student characteristics (i.e. ethnicity).
7. Determine the treatment of duplicate records (see Definitions: Duplicate). This process utilizes a complex set of rules performed by Data Planning and Analysis, together with district officers and the test vendors. Partial documentation can be supplied on request.
8. Assign FAY from enrollment data.
    a) FAY =YES if a student is enrolled at the 120th day of the prior school year, and the 40th, 80th, and 120th day of the current school year. There are exceptions to this rule:
    b) Students in <u>transition grades</u> (the lowest grade in the school's grade span) are FAY=Yes provided they meet the following conditions:

(1) Enrolled 40D, 80D, 120D of the current year, AND

(2) Enrolled 120D of the prior year in the same LEA as the transition school. This rule applies to locally authorized charter schools but not to state authorized charter schools.

c) Students in reorganized schools in the current year are FAY=Yes under the same provision as transition grades (A.) provided they are in a lower grade that is new to the school. For example, if a school that previously served grades 7-8 adds a 6th grade, both 6th grade and 7th grade students must meet conditions 1 and 2 to be considered FAY=Yes.

d) Students in new schools are FAY=Yes under the same provisions as transition grades (A.). For example a new school that serves grades 6, 7, and 8 requires only that students meet conditions 1 and 2 to be considered FAY=Yes.

e) State charter schools follow the same options in A, B, and C, but without the requirement for LEA membership in the prior year (A.2.).

9. Assign subgroup membership from snapshot data.

10. Transform prior year scaled scores.

a) SBA Proficiency scores from years prior to 2011 utilized an expanded vertically aligned scale, with scores ranging from 200 to 950. In 2011 the assessment was rescaled to utilize a vertically moderated scale that was uniform for all grades. The new scaled scores range from 0 to 80, with proficiency anchored at 40.

b) Scores prior to 2011 were transformed to the current scale bridge information supplied from the SBA vendor. In addition, the NMAPA scaled scores were transformed to match the new scale, utilizing a linear transformation. Details of the transformation algorithm are provided on request.

B. Student matching

1. Separately for each student file, the dataset is aggregated to the school level to calculate the cut score of the 25th percentile in math and reading separately (so it is possible for a student to be in the bottom quartile in one subject but not the other).

2. Given the cut score for each school, a student is then identified as Bottom Quartile (BQ) (ie, the variable =1 if a student is in the bottom quartile and 0, otherwise).

3. The three above files are merged by student ID and the most recent year school ID is used as the school of record for that student. The student's Bottom Quartile indicator is used from the most prior year file.

4. Missing individual student background information in any year is replaced with information from one of the other years, if available, otherwise, mean replacement is used. Test scores are not replaced, but students with incomplete data remained in the analysis.

C. Feeder Schools

1. Feeder schools are only evaluated at the school level. They are not included in LEA and state analyses.
2. Participation rates are not calculated for feeder schools.
3. Attendance is generated for feeder schools the same as for non-feeder schools.
4. Performance:
    a) Wherever possible, feeder school ratings are based on test results from students who graduated from these feeder schools into a tested grade (i.e. 3rd grade for a K-1 school, or 11th grade for a 9-10 school).
    b) The scored results of 3rd graders and (potentially 4th graders) and 11th graders will be accumulated in a specialized file used for rating feeder schools. Students can contribute their scores to as many as three feeder schools (for example, if a student attended schools each consisting of a single grade level, K, 1, and 2).
    c) Current Standing estimates were apportioned based on the percentage of students that the feeder school students represented of the school it fed into E.g., if students from a feeder school fed into two schools with (tested grades) and each received half of the students, then the feeder school score on school growth would be 50% from each school.
    d) For individual student growth, the growth estimates of the students were aggregated back to the feeder school and averaged to create a individual growth for bottom quartile (if they were bottom quartile students) and non-bottom quartile students.
    e) School growth estimates were apportioned based on the percentage of students that the feeder school students represented of the school it fed into. For example, if students from a feeder school fed into two schools with tested grades and each received half of the students, then the feeder school score on school growth would be 50% from each school.
    f) Exited students used to assign accountability for feeder schools must reside currently at a school that is in the same LEA as the feeder school.
    g) A small number of feeder schools exist for which alumnae cannot be found in the assessed population, and for which no feeder pattern or LEA affiliation can be identified for a potential data substitution. These are typically new schools that add one grade each year and have not yet reached a tested grade. Students are anticipated to continue at the feeder school until they reach a tested grade.

V. Evaluation Parameters

A. Rounding

1. No rounding occurs until the final rates are reported. All computations prior to reporting utilize unrounded figures.
2. Final rounding occurs to the first decimal place (i.e. 92.2%) unless otherwise indicated.
3. Data files provided to PED will include non-rounded figures, or if necessary, rounded to the 4th significant digit beyond what is reported (i.e. 92.234821 in the previous example)
4. The terminal digit of 5 is rounded up.
5. Participation and attendance are not rounded and are truncated to the nearest significant digit.

VI. Calculations, School

A. Schools Rated

1. School calculations include public, locally-authorized charter, state-authorized charter, feeder, and state-supported schools.
2. School calculations exclude off-site locations, programs, and students tested in those locations. These students are rolled up to LEA accountability.

B. Participation

1. Rates are computed for these subgroups that have 40 or more students (see *Definitions*, *Subgroups*):
   a) All Students
   b) Caucasian/White-Non Hispanic (ethcode ='C')
   c) Black-Non Hispanic (ethcode='B')
   d) Hispanic (ethcode='H')
   e) Asian/Pacific Islander (ethcode='A')
   f) Native American (ethcode='I')
   g) English Language Learners (ellstatus=1 or ellstatus=2 or ellstatus=3)
   h) Special Ed-Not Gifted (spedcode=Y)
   i) FRL=Yes when student is either Free (F) or Reduced (R)
2. Participation rates are not computed for feeder schools.
3. The participation rate is a percentage:
   a) Numerator READING: count if PL is 1, 2, 3, or 4 (valid score). Add students for whom the TC is equal to 3, which means they are counted for participation even though they do not have a scored test.
   b) Numerator MATH: count if the PL is 1, 2, 3, or 4 (valid score) for each subtest.
   c) Denominators READING, MATH: count all students in that subgroup, school, or district from the Assessment snapshot.
   d) Compute the percentage. For reporting a truncated calculation to the whole number, do not round.
4. Participation Averaging, Two Years. If a school does not meet the 95% rate for all subgroups with 40 or more students, compute for each eligible subgroup an unweighted average. For example the two year rate would be:
   a) (current year participation percentage + previous year participation percentage)/2
   b) Note that this calculation does not use the number of students, but only the percentage rates from the years being considered.
   c) Two year averaging results should be rounded (see *V. Calculation Parameters*).
5. Participation Averaging, Three Years. If a school does not meet participation, via the target (95%) or via the two-year average, compute the three year participation average, in the same manner:
   a) (current year percentage + prior year percentage + year antecedent to prior year percentage)/3.
   b) Three year averaging results should be rounded (see *V. Calculation Parameters*).

C. Attendance

1. A rate is computed school wide for every school that is graded, including feeder schools and high schools. There are no excluded grades.

2. Ineligible student attendance records must be removed prior to calculation:
    a) Days Present < 0 or missing
    b) Days Enrolled < 0 or missing
    c) Days Present > Days Enrolled
    d) Attendance records for students without a corresponding record in Student Snapshot at the same location
3. The rate is the average of all individual attendance rates of every student in 40D, 80D, and 120D snapshots:
    a) Numerator: days attended accumulated across all enrollment periods (40th, 80th, 120th day snapshots) for a single student.
    b) Denominator: days enrolled accumulated across all enrollment periods (40th, 80th, 120th day snapshots) for a single student.
    c) Compute the percentage rate per student using the above numerator and denominator.
    d) Compute the school average: sum percentage rates for all students and divide by the total number of students.
    e) Note that a single student can contribute to the rates of more than one school if they are mobile.
4. When a school is missing a reporting period (40th, 80th, or 120th) compute the attendance rate in the same manner, only excluding the missing reporting period from the days attended and days enrolled.
5. Attendance is rounded to two decimal places for reporting (i.e. 91.59%).

D. Graduation

1. Graduation rates are one-year lagged. That is, the rates that are published in May are for the cohort that graduated by August 1 of the prior year. Calculation of 4-year and 5-year cohort graduation utilizes the *Shared Accountability* method and is described fully in the *Graduation Technical Manual* on the PED website.
2. A rate is generated for every school that has any grade 9, 10, 11, or 12. For the purposes of school grading, rates are counted only for high schools (see *Definitions, High School*).
3. For new high schools that do not yet have a graduating cohort class, the district mean will be substituted. Where a district affiliation is not evident, the state mean will be substituted.

E. Current Standing

1. Proficiency refers to the percent of students who are proficient or above in the current reporting year. The four categories of proficiency are:

    Beginning Step (PL=1)

    Nearing Proficient (PL=2)

    Proficient (PL=3)

    Advanced Proficient (PL=4)

    a) Examinees whose tests were invalidated for various reasons (see *IV. Conditioning of Data, TC Codes*) did not receive a score but are still counted for accountability (PL=5).
    b) Numerator: sum of students with PL=3 or 4
    c) Denominator: sum of students with PL=1, 2, 3, 4, or 5
    d) The resulting percentage is rounded to one decimal place (i.e. 65.3%)

2. Conditioned Status refers to an adjustment of status that takes the school's student characteristics into account. Conditioning is accomplished by the statistical application of VAM to school outcomes, isolating the characteristics of the students from the school's effect on achievement. The result is a truer picture of the school's impact (value added) on student achievement.

   a) The conditioning variables include are identified for each student:

      (1) Gender
      (2) ELL
      (3) FRL
      (4) SWD
      (5) Ethnicity (African American, Hispanic, Native American, Asian)
      (6) FAY
      (7) Bottom Quartile

   b) Conditional Status is estimated simultaneously with school growth and is detailed in F.

F. School Growth refers to the ability of a school to increase grade-level performance. For example, did the performance of 3rd graders improve over the three prior years?

Using the merged file described in IV b., missing individual student background information in any year is replaced with information from one of the other years, if available, otherwise, mean replacement is used. Test scores are not replaced, but students with incomplete data remained in the analysis.

Separate files are created for math and reading (math and reading models are run separately).

The datasets (math and reading) are transposed from the "wide" format to the "long" format where each student could have multiple records (up to three).

These files are then aggregated to cohort/ year to form a year level file with only one record per year (i.e. there were 3 records per school.

The "long" file was also aggregated to the school level (2011 school ID), aggregating all of the student variables to the school to be used for peer effects. Student background variables were indicator coded and grand mean centered[1,2]:

$$SS_{itk}^{s} = \pi_{0tk} + \pi_{1tk}(\text{AfricanAmerican}_{itk} - \text{AfricanAmerican}\ldots) + \pi_{2tk}(\text{Hispanic}_{itk} - \text{Hispanic}\ldots) + \pi_{3tk}(\text{Asian}_{itk} - \text{Asian}\ldots) + \pi_{4tk}(\text{Native}_{itk} - \text{Native}\ldots) + \pi_{5tk}(\text{Female}_{itk} - \text{Female}\ldots) + \pi_{6tk}(\text{FRL}_{itk} - \text{FRL}\ldots) + \pi_{7tk}(\text{SWD}_{itk} - \text{SWD}\ldots) + \pi_{8tk}(\text{ELL}_{itk} - \text{ELL}\ldots) + \pi_{9tk}(\text{FAY}_{itk}) + \pi_{10tk}(BQ_{itk}^{s}) + \sum_{g=1}^{G} \pi_{(10+g)tk}(\text{GRADE}_{itk}^{g} - \text{GRADE}\ldots) + e_{itk} \quad (1)$$

In equation (1), the scale score in subject *s* (math or reading) for student *i* at time *t* in school *k* is a function her background characteristics, FAY, whether or not they were in the bottom quartile, and grade. The subscripts *itk* indicates student *i* at time *t* in school *k*, while the subscript, .. indicates the grand mean of the sample, i.e. the New Mexico State Average. There are G= 5, with grade 3 being the "left-out" grade. Also, grade is recoded to equal the actual grade value -3, so that the intercept represents 3rd grade. For elementary/middle school $\pi_{10tk}(BQ_{itk}^{s})$ is excluded from the model. As well as the corresponding specification displayed in eq. (3).

The two coefficients of interest are and are modeled as random effects:

[1] For the 2010-2011 HS calculation, Grade only represents one parameter, grade 11.
[2] BQ is not included in the elementary/middle school model.

$$\pi_{0tk} = \beta_{00k} + \beta_{01k}(Year_{tk}) + r_{0tk}$$
$$\pi_{10tk} = \beta_{100k} + \beta_{101k}(Year_{tk}) + r_{10tk} \quad (2)$$

The remaining coefficients in (1) were treated as fixed effects. Year is reverse-coded, such that 0=2011 and -1 =2001, etc. In this way β_{00k} is the current year status. We are further interested in whether we can identify unique school effects associated with the (the β's). These unique effects form the basis for the value added estimates and are estimated as random effects.

$$\beta_{00k} = \gamma_{000} + X\Gamma + U_{00k}$$
$$\beta_{01k} = \gamma_{010} + U_{01k}$$
$$\beta_{100k} = \gamma_{1000} + U_{100k}$$
$$\beta_{110k} = \gamma_{1010} + U_{101k} \quad (3)$$

Where $X\Gamma$ is the set of individual student characteristics (and coefficients) in (1) aggregated to the school as proxy for peer effects. That is, the school proportion of

(1) Gender
(2) ELL
(3) FRL
(4) SWD
(5) Ethnicity (African American, Hispanic, Native American, Asian)
(6) FAY

The variables listed above were also grand-mean centered. The primary interest is in U_{00k} and U_{01k} for elementary/middle schools and additionally U_{101k} for high schools, these represent the unique school effects for current year Conditional Status, School Growth, and School Growth for the Bottom Quartile, respectively[3].

We use the empirical Bayes estimates of U_{00k}, U_{01k}, and U_{100k} and use those to determine the percentile rank each school's U is in the state. For each of the three effects, we first normalize (using a *t*-distribution) and estimate the cdf (cumulative density function) for the normalized score. This, "percentile" is multiplied by the points available for a subject and grade component. For example for conditional status in elementary school math:

If the cdf t_{U00k} returns .80 then,
Points for Math for conditional standing = .8 X 7.5 = 6.0.

For high schools in 2011, school growth was computed separately for the highest three quartiles (high performers) and bottom quartile of students (see next two sections) and reported there. Starting in 2012-2013, school growth will be estimated in elementary/middle, and high Schools in the same way, which is that there will only be one school growth estimate (per subject) per school.

G. <u>Student Growth</u>

Using individual student growth to monitor school performance uses a mixed effects growth model. In 2010-2011 this applies to elementary and middle schools. Using the same "long" files as described earlier, we first aggregated to a student file to form a student level file with only one record per student.

The "long" file was also aggregated to the school level (2011 school ID), aggregating all of the student variables to the school to be used for peer effects.

The individual student growth model uses only three independent variables .

[3] Recall the for Elementary/Middle school there are not a separate effects for BQ students.

$SS_{tik}^{s} = \pi_{0ik} + \pi_{1ik}Year_{tik} + \pi_{2ik}FAY_{tik} + e_{tik}.$ (4)

Where at time *t*, for student *i* in school *k* in subject *s*, the score is a function of initial status, π_{0ik}, growth per year, π_{1ik} ,FAY status, π_{2ik}, and error, e_{tik}. Year is coded as 2011 = 2 and 2010 =1, etc.

We allow the initial status and growth to vary randomly among students, while setting FAY as a fixed effect, hence:

$\pi_{0ik} = \beta_{00k} + \beta_{01k}(BQ_{ik}) + r_{0ik}$
$\pi_{1ik} = \beta_{10k} + \beta_{11k}(BQ_{ik}) + r_{1ik}$ (5)

And, of course we are interested in whether there are unique school effects associated with schools:

$\beta_{00k} = \gamma_{000} + U_{00k}$
$\beta_{01k} = \gamma_{010} + U_{01k}$
$\beta_{10k} = \gamma_{100} + U_{10k}$
$\beta_{11k} = \gamma_{110} + U_{11k}$ (6)

For the individual Student Growth calculations for Bottom Quartile and non-Bottom Quartile (aka Highest Performing) we use the Empirical Bayes (EB) Coefficient for growt, that is, the predicted growth. This is estimated as:

$\beta^{*}_{01k} = FV_{01k} + U_{01k}$ and
$\beta^{*}_{11k} = FV_{11k} + U_{11k}$,

for the highest performing and bottom quartile students, respectively. Hence, the EB coefficient, β^{*}, is equal to the fitted value, *FV*, and a unique school effect, *U*. The fitted value is the estimated school effect based on the school variables in the model and *U* is the EB residual.

We use the EB coefficients of β^{*}_{01k} and β^{*}_{11k} to determine the percentile rank each school's growth (for highest performing and bottom quartile) is in the state. For each of the two effects, we first normalize (using a *t*-distribution) and estimate the cdf (cumulative density function) for the normalized score. This, "percentile" is multiplied by the points available for a subject and grade component. For example, for Student Growth in elementary school math:

If the cdf $t_{\beta^{*}01k}$ returns .80 then,
Points for Math for conditional standing = .8 X 10 = 8.0.

This model would be run separately for math and reading, generating separate unique school estimates that are separately estimate the percentile.

H. <u>Highest Quartile Student Growth</u> Students assessed in a single year were cleaved into two groups, the highest three quartiles (high performers) and the bottom quartile.

1. See IV b. conditioning data, student matching.
2. Elementary and Middle Schools:
    a) See VI G. for a description of the growth model used to calculate student growth.
    b) β^*_{01k} is the basis student growth calculations for the highest performers.
3. High Schools
    a) For the inaugural year of school grading individual student growth was not available, since the only high school grade tested in prior years was grade 11.
4. The method described for school growth (see prior section *School Growth*) was duplicated for 11th graders whose scaled score placed them in the highest 75% of the assessed population for that school in the current year.
    a) U_{01k} is the basis for School Growth, high performers.
5. The method described for school growth in high schools for the highest performers will change to the method described for student growth in elementary and middle schools (IV b.) in the 2012-2013 school year.

I. <u>Lowest Quartile Student Growth</u> Students assessed in a single year were cleaved into two groups, the highest three quartiles and the lowest quartile. This indicator applies to students whose scaled score placed them in the lowest 25% of the assessed population for that school in the current year.

1. Groups were identified separately for Reading and for Math.
2. Elementary and Middle Schools
    a) See VI G. for a description of the growth model used to calculate student growth.
    b) β^*_{11k} is the basis for calculating student growth for the bottom quartile of students.
3. High Schools
    a) For the inaugural year of school grading individual student growth was not available, since the only high school grade tested in prior years was grade 11.
4. The method described for school growth (see prior section *School Growth*) was duplicated for 11th graders whose scaled score placed them in the lowest 25% of the assessed population for that school in the current year.
    a) U_{101k} is the basis for School Growth bottom quartile students.
5. The method described for school growth in HS for the bottom quartile will change to the method described for student growth in elementary and middle school (IV b.) in the 2012-2013 school year.

J. <u>Opportunity to Learn</u>

1. Is based on Attendance as described in II. 3; and,
2. An Opportunity to Learn Survey administered to students taking the SBA, beginning in 2012. The survey consists of 10 questions related to the opportunities teach provide students to learn the materials necessary to be successful on the assessments.
    a) Survey responses are on a Likert-type scale.
    b) Elementary students are asked to think about the teacher that taught them the content on which they are being tested.
    c) High School students are asked to consider the teachers they currently have, in general, and respond about the opportunities they provide.

K. College and Career Readiness

PARTICIPATION is determined by the percent of enrolled students (10th, 11th, or 12th graders) eligible for that indicator, who show evidence of a career or a college preparatory path. Career paths are established through course enrollment leading to an industry recognized certification. College paths are established through a student's taking a recognized academic precursor to post-secondary education.

SUCCESS is determined by the percent of students who attempted any of these indicators, and who met the success criterion.

- Achieving a '3' on an Advanced Placement (AP) exam in a core academic area (11th, 12th graders)
- Achieving *College Readiness* benchmark scores on the four content areas of the ACT. Each content area contributes 1/4 to the potential points. (11th, 12th graders)
    - English Composition
    - Social Sciences [Reading]
    - College Algebra [Mathematics]
    - Biology [Science]
- Achieving *College Readiness* benchmark scores on the three content areas of the PSAT. Each content area contributes 1/3 to the potential points. (10th graders)
    - Reading
    - Mathematics
    - Writing
- Completing all course requirements for *Career Readiness* with a "C" or better, and graduating with a regular diploma in 4 years. (12th graders)

Student attempts will be pooled as the denominator, and student successes will be pooled as the numerator for the final calculation. Students may make multiple attempts, with multiple indicators, and the single most successful indicator will be retained.

Point Boundaries for All Indicators

Elementary and Middle Schools

Indicator	Grade	Points*
Current Standing	A	30.6 or above
	B	23.8 to 30.5
	C	18.9 to 23.7
	D	14.6 to 18.8
	F	14.5 or below
School Growth	A	8.9 or above
	B	6.6 to 8.8
	C	5.0 to 6.5
	D	3.4 to 4.9
	F	3.3 or below
Growth of Highest Performing Students	A	13.7 or above
	B	8.6 to 13.6
	C	5.8 to 8.5
	D	3.0 to 5.7
	F	2.9 or below
Growth of Lowest Performing Students	A	18.6 or above
	B	16.5 to 18.5
	C	14.2 to 16.4
	D	11.5 to 14.1
	F	11.4 or below
Opportunity to Learn	A	9.0 or above
	B	8.0 to 8.9
	C	7.0 to 7.9
	D	6.0 to 6.9
	F	5.9 or below
Overall Grade	A	75.0 or above
	B	60.0 to 74.9
	C	50.0 to 59.9
	D	37.5 to 49.9
	F	37.4 or below

* Points are rounded for tables for simplicity. However in calculations, figures were carried out to 6 or more decimals. Therefore, letter grades at the highest and lowest boundary of a point span may not be apparent because of rounding. Unrounded figures are available upon request from PED's Data Planning and Analysis Bureau.

High Schools

Indicator	Grade	Points*
Current Standing	A	18.8 or above
	B	14.2 to 18.7
	C	10.9 to 14.1
	D	9.0 to 10.8
	F	8.9 or below
School Growth	A	This indicator was combined with the next two indicators in 2011. It will be reported separately in 2012.
	B	
	C	
	D	
	F	
School Growth of Highest Performing Students	A	13.9 or above
	B	10.9 to 13.8
	C	6.8 to 10.8
	D	3.8 to 6.7
	F	3.7 or below
School Growth of Lowest Performing Students	A	12.4 or above
	B	8.4 to 12.3
	C	6.3 to 8.3
	D	5.1 to 6.2
	F	5.0 or below
Opportunity to Learn	A	9.0 or above
	B	8.0 to 8.9
	C	7.0 to 7.9
	D	6.0 to 6.9
	F	5.9 or below
Graduation	A	16.2 or above
	B	13.6 to 16.1
	C	12.1 to 13.5
	D	10.0 to 12.0
	F	9.9 or below
Career College Readiness	A	13.6 or above
	B	10.0 to 13.5
	C	8.6 to 9.9
	D	6.1 to 8.5
	F	6.0 or below
Overall Grade	A	75.0 and above
	B	65.0 to 74.9
	C	50.0 to 64.9
	D	35.0 to 49.9
	F	34.9 and below