New Mexico

A-F School Grading Technical Guide

Calculation and Business Rules

For Schools and Districts

NMPED
Public Education Department

Christopher N. Ruszkowski

Acting Secretary

Assessment and Accountability Division

Lisa Chandler, Director

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]. The *School Grading Technical Guide* provides detailed decision rules for each indicator, statistical treatment of data, and calculation parameters. These business rules apply to New Mexico public and charter schools and do not apply to private schools, Bureau of Indian Education schools, home schools, or other schools that are not within the jurisdiction of the New Mexico Public Education Department (PED). In addition to schools, districts are held to similar accountability standards.

New Mexico's school grading model was first approved in 2012 by the U.S. Department of Education (USDE) and was recently one of the first to be approved under requirements of the *Every Student Succeeds Act* (ESSA). Under ESSA, school accountability will undergo slight revisions of student success indicators and measures of academic progress beginning in 2019. The full plan is available on the website at http://ped.state.nm.us/ped/ESSA_docs/FINAL_NMESSAPlan.pdf .

Please submit questions or requests for revision and clarification to Ryan Tolman (ryan.tolman@state.nm.us).

I. Definitions and Abbreviations

Terms used in this document and their meaning in school grading follow:

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

> If Full Academic Year (FAY)=Yes, the accountable school is the FAY school; or
> IF FAY=No, the accountable school is the location where the student was tested.

Aggregate describes presentations of data where individual students are combined in a group. Typical aggregates are English learners or students with disabilities. Common practice uses the term "disaggregate" to refer to reporting information in these smaller student groups.

Bonus Points are extra points awarded to schools that worked hard to involve students and parents in education, reduce truancy, and promote extracurricular activities. Points are based on the school's use of Next Step Plans (NSP) and Student Assist Teams (SAT), truancy statistics, and participation in activities recognized by the New Mexico Activity Association. The rubric for earning bonus points is available on the PED's website at http://aae.ped.state.nm.us/. These points are added to a school's total points prior to determining the school's overall grade.

CDF (Cumulative Density Function) is the probability (from 0 to 1.0) that an observation will be found at a value less than or equal to the value of that observation. In other words, it represents the sum of the area to the left of a density curve for a given value. To illustrate, the graph below left shows several examples of probability distributions. These could be distributions of scores, points, or rates. Note that three of the distributions center around zero, and one (green) has a mean of -2. The gold, red, and blue distributions have been standardized to a mean of zero, but they could just as easily be means of 55 on a test that totals 100 points.

The CDF graph on the right shows each of the probability distributions on the left graphed in cumulative (additive)
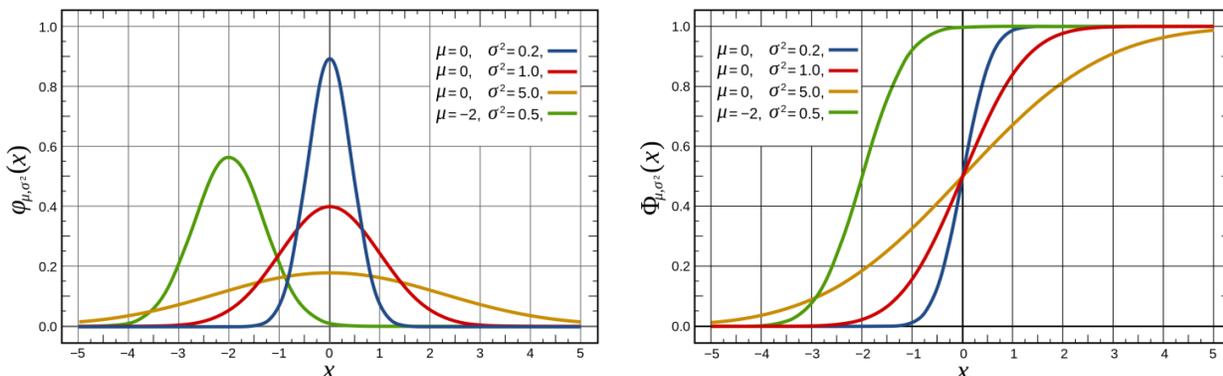
fashion, that is, the values to the left of any given point on the x-axis are accumulated until the entire group of probabilities sums to 1.0 (or 100%). Where the distributions are peaked (blue), the slope near the mean changes quite rapidly, and where the distributions are flatter (gold), the slope changes slowly along a broader range of values.

Typically, the CDF distribution is used to convert scores on the horizontal x-axis to points on the vertical y-axis such as value added-model (VAM) score to Current Standing points. One can see that a small gain in the x-axis will

result in a large gain in points on the y-axis when the score is near the middle of the distribution. The conversion tables were established and set from the first operational year of school grading in 2012. Now that they are fixed, scores from future years can be judged on a stable framework that allows all schools to progress to a higher letter grade.

Current Standing is a two-part measure of the status of a school in the current year. The two parts are composed of the percentage of students who are proficient and a score based on a growth model that accounts for prior scores.

Duplicate refers to two test records that have the same student identification (ID). Duplicates may occur when the same student was tested twice, such as once in English and again in 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 data review. Each student can contribute only one test score for each content area. Where duplicates cannot be resolved, both scores are removed from the accountability file.

ELA (English Language Arts) is a broad term encompassing reading, writing, speaking, listening, and language literacy. For the purposes of school grading, ELA represents components of reading and writing that are assessed within the assessment battery.

FAY (Full Academic Year) indicates whether a student has been enrolled at a single location for one year. The school or local education agency (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. FAY is also used as a measure of student educational stability and risk. The portion of a school's population that is not FAY (more mobile) is used to adjust predicted values in the growth calculation.

Growth Models (see *Value-Added Models*).

HED (Higher Education Department) is a separate cabinet-level agency in New Mexico that governs state institutions of higher education. The PED limits accountability to only K-12 institutions but works closely with the HED for the reporting of some post-secondary measures.

IStation is an assessment of early literacy for students in grades kindergarten through second grade. The assessment is administered in both English and Spanish versions and is the sole source of achievement for students in these grades.

LEA (Local Educational Agency) traditionally refers to the 89 school districts that manage more than 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.

Model determines the school's points and framework for receiving a letter grade, and each school is assigned to either high school (HS) or elementary/middle school (EL) based on the grade span of the school. Where the grade configuration does not clearly dictate a model, the available indicators for each model are examined to maximize the data available for the school's grade. For example, high schools that also have middle school grades are reassigned to the EL model when they do not have cohort measures and growth for their high school grades.

NMAPA (New Mexico Alternate Performance Assessment) is the assessment in mathematics, English language arts (ELA), science, and social studies that is administered for students with profound cognitive disabilities.

Normalized refers to the statistical treatment of a distribution of scores to adjust the values measured on different scales to a common scale, often prior to averaging or combining them. Typical examples of normalized scores are *z* scores, where the mean of the distribution is zero and variation in the positive and negative direction is described in standard deviation units (see *CDF*).

One-Percent Rule relates to the assessment of students with the most severe cognitive disabilities. Generally, these students represent approximately 1% of all students in the state. Set by the USDE, the one-percent rule requires that the percentage of students taking an alternate assessment cannot exceed one percent of the total number of assessed students 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 higher are converted to *Not Proficient*. This rule is applied only to multi-school LEA accountability.

Opportunity to Learn (OTL) represents the learning environment schools provide. It is determined from student attendance and scores on a student or parent survey administered annually. The survey measures the extent to which classroom teachers demonstrate instructional practices known to facilitate student learning.

PARCC (Partnership for Assessment of Readiness for College and Careers) is the regular assessment of English language arts and mathematics for students in grades 3 through 11. It is aligned to *Common Core State Standards*, and was administered for the first time in New Mexico in 2015. The PARCC assessment includes new end-of-course assessments in high school mathematics (i.e. geometry) and adds the 9th grade to school accountability.

Proficiency refers to whether the student has achieved a benchmark score on various criterion-based assessments. For example, the PARCC scores classify students in one of five levels, with levels four and higher considered college and career ready, on grade level, or proficient. The scores required to be considered proficient were determined by national school officials, subject matter experts, and certified teachers in the inaugural years of each assessment. The standards are revisited every three to five years for possible re-calibration.

Q refers to the student's quartile status for school grading. Every student is classified as being a member of either the Q1 (the bottom quartile) or Q3 (the top three quartiles) subgroup. By separating students into higher and lower performing groups, schools can see if their practices are benefitting both groups. New Mexico law requires that accountability highlight the lowest performing 25% of the student population.

SAM (Supplemental Accountability Model) schools are a subset of graded schools that serve students at risk of academic failure. Schools qualify by having a high proportion of returning adults or a high proportion of students with disabilities and by publically declaring the school mission and goals for these students. School grading rules are moderated slightly for SAM schools, and mission-specific measurements are incorporated into the determination of the school's letter grade.

SBA (Standards Based Assessment) is the regular assessment of science for students in grades 4, 7, and 11 and of Spanish language arts for students in grades 3-8 and 10-11. Science scores are not yet used for school grading.

School Growth measures whether a school's scores tend to go up or down in comparison with previous performance.

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 the following:

- Second Wednesday of October (known as 40th day; abbreviated as 40D)
- December 1 (known as 80th day; abbreviated as 80D)

- Second Wednesday of February (known as 120th day; abbreviated as 120D)
- End of Year, variable but principally in June (known as EOY)

Additionally, a specialized snapshot occurs during the last week of testing (known as the "Accountability", or "4/20" snapshot); timing is variable but the snapshot generally occurs in March or April.

Standard Setting refers to the process of setting cut scores. The A-F grading framework was set using student data from 2011 and 2012. At that time the PED normed each indicator, anchoring the grade of "C" at the average for all schools statewide. For example, a school that achieves a "C" in *College and Career Readiness* has scored near the average of all other schools in the state. A school that receives an "A" or "B" for any indicator performed above the average of how other schools performed on that indicator in 2011 and 2012. Since the time of the standard setting, the scale has not changed.

Status refers to schools in some form of improvement or reward. The four status categories from lowest to highest performing follow:

- Priority (5% of schools)
- Focus (10% of schools, not in Priority status)
- Strategic (10% of schools, not in Priority or Focus status)
- Reward (5%)

Priority and *Focus* categories will be replaced under ESSA as *Comprehensive School Improvement* (CSI) and *Targeted School Improvement* (TSI) with slightly different definitions.

Student Growth represents how much individual student achievement grows compared with other students. Student growth is estimated using a multilevel (mixed effects) regression model.

Subgroups are the student groups aggregated for accountability reporting, as required by USED. A single student can contribute to several subgroups. While data for each subgroup are reported, only the student characteristics of FAY and Q1/Q3 are used in school grading calculations.

All students (reported as "All Students")

Caucasian/White-Non Hispanic (reported as "Caucasian" or "White")

Black-Non Hispanic (reported as "African-American" or "Afr Amer")

Hispanic (reported as "Hispanic" or "Hisp")

Asian/Pacific Islander (reported as "Asian")

Native American (reported as "American Indian" or "Am Indian")

English Language Learners (reported as "English Language Learners"; abbreviated as "ELL" or "EL")

Special Ed, Not Gifted (reported as "Students with Disabilities" or "SWD")

FRL, free or reduced price lunch program (reported as "Economically Disadvantaged" or "ED")

Gender (reported as "F" or "M")

Migrant (Title 1C) (reported as "Migrant")

FAY (reported as "Y" or "N")

Q1, quartile 1, lowest performing 25% of students

Q3, highest three quartiles, remaining students who are not Q1

Test Code indicates the specialty within a particular assessment. The broader concepts of English language arts or mathematics each contain domains such as third-grade math, algebra I, or tenth-grade English. These specialty areas are applied during growth modeling to ensure that students are being compared to their academic peers.

Valid Test refers to a test that can be successfully scored. Tests are invalidated for a number of reasons (see *Conditioning of Data, Test Completion Code*), and these invalid tests adversely affect the school's participation rate.

VAM (Value-Added Modeling) is a statistical procedure that isolates the school's contributions to student performance from factors outside the school's control that are known to affect student test performance. VAM provides a truer picture of the school's impact (value added) on student achievement than just using achievement scores. This is also referred to as growth modelling.

II. Data Sources

School Attributes. The school file lists all open public and charter schools in New Mexico with enrolled students in any grades K through 12. The purpose of this file is to finalize the list of schools that 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:

1. Public school. If students take the test at a school or program where they are enrolled for only part of the day, their parent school must be identified and their scores attributed to that 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 parent school, after approval by PED.

2. Locally authorized charter school. If the authorizer is one of the 89 recognized multi-school districts, the school is under the jurisdiction of the authorizer.

3. State-authorized charter school. If the authorizer is the Public Education Commission, the school is considered independent from a district and is under the jurisdiction of the state authority.

4. Off-site program. Students in off-site programs (i.e., correctional facilities, treatment centers, homebound, or hospitalized) are generally excluded from school-level calculations and included at the LEA level for rating. However, if the student qualified as FAY at a school prior to entering the program and testing, his or her test will count at the FAY school. Off-site programs are not rated.

5. State-supported school. State-supported schools receive funding from multiple sources and may not be fully under the jurisdiction of the PED. Examples include *Juvenile Justice Schools,* the *School for the Blind and Visually Impaired*, and the *School for the Deaf*. The *New Mexico Military Institute* receives private funding, operates a junior college in addition to high school, and is exempted from rating by New Mexico statute.

Additionally, schools are characterized by

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

7. SAM school (Y/N)

8. Level (elementary/middle or high school)

9. New or Reorganized (impacts inheritance of school accountability, FAY, and other calculations)

10. 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.

11. Graduation is computed annually by the Accountability and Assessment Division at PED. The file lists rates and counts by subgroup, school, and LEA for 4-year, 5-year, and 6-year cohorts. Schools with any high school grade (9, 10, 11, or 12) will receive a rate. Details of the calculation are fully covered in the PED's *Graduation Technical Manual*.

12. Historic data are provided by the Accountability and Assessment Division from prior years and include figures that are required for the current year's calculations.

Student Attributes. The student file lists all students assessed in the current year, their demographics, historic scores, and relevant accomplishments. The purpose of the student file is to calculate the parameters used to grade the student's current school.

1. Mathematics and ELA proficiency scores are supplied by the vendors that administer the assessments used in School grading.

2. Subgroup membership is assigned from the data submitted by districts to STARS snapshots. Where students are missing from the 4/20 snapshot, the nearest snapshot date where the student is found is used.

1. Surveys are completed by students and by parents for younger grades. The surveys consist of 10 questions with six-step scales, and questions on both parent and student versions are very similar. For 2016 these surveys are also used in Educator Effectiveness, where students may fill out multiple surveys for multiple teachers. Surveys are administered anonymously through a secure online portal through a third party vendor. Survey questions are found on the PED website at http://ped.state.nm.us/ped/NMTeachDocs/Toolbox/StudentSurveyWithCrosswalk.pdf and

    http://ped.state.nm.us/ped/NMTeachDocs/Toolbox/ParentSurveyWithCrosswalk.pdf.

2. Student/Parent Engagement does not have a specific data file associated. This information arrives through data submitted to PED by the LEA and is part of Bonus Points.

3. ACT student level data are submitted by LEAs through STARS regular submissions throughout the year.

4. SAT student level data are submitted by LEAs through STARS regular submissions throughout the year.

5. PSAT student level data are submitted by LEAs through STARS regular submissions throughout the year.

6. AP (Advanced Placement) student level data are submitted by LEAs through STARS regular submissions throughout the year.

7. 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.

8. Career Technical Education is partly determined by course enrollment 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, a student meets the success benchmark.

9. <u>Accuplacer</u> student level data are submitted by LEAs through STARS regular submissions throughout the year.

10. <u>COMPASS</u> student level data are submitted by LEAs through STARS regular submissions throughout the year.

11. <u>Aspire (formerly PLAN)</u> student level data are submitted by LEAs through STARS regular submissions throughout the year.

12. <u>IB (International Baccalaureate)</u> student level data are submitted by LEAs through STARS regular submissions throughout the year.

13. Standardized Subject Tests (SATSUB) data are supplied by LEA's during annual data exchange.

14. TABE data are supplied by LEA's for SAM schools during annual data exchange.

15. Work Keys data are supplied by LEA's for SAM schools during annual data exchange.

16. ASVAB data are supplied by LEA's for SAM schools during annual data exchange.

III. Data Validation

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, the following:

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, Accountability, 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 reports

IV. Conditioning of Data

Assessment Scores. A subset of records in the test data file (valid tests) and content within records (ELA, mathematics) is used for school rating. Results for Spanish, English, and Braille administrations are included, as are the alternate assessments. The following guidelines apply to the selection and cleaning of those records:

1. Remove records of students from non-PED schools, such as home-schooled students who were permitted to test at the discretion of the LEA or BIE students who are not held to PED accountability.

2. Reconcile Test Completion Code (TC), Scaled Score (SS), and Proficiency Level (PL). The TC is a field marked 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, i.e., SS and PL. This reconciliation is performed by the Accountability and Assessment Division and details can be supplied upon request. Test Completion codes have these meanings:

TC=0   Tested all sessions and received a valid score

TC=1   Withdrew before testing; excused

TC=2   Received a non-allowed modification; invalidate test (SS=99)

TC=3   Exempt from ELA/reading assessment because of language exception; excused

TC=4   Medical emergency exemption; excused

TC=5   Parental refusal; invalidate test; (SS=99)

TC=6   Incomplete testing; invalidate test; (SS=99)

TC=7   Testing irregularity; invalidate test; (SS=99)

TC=8   Absent; invalidate test; (SS=99)

3. A single student can have a valid mathematics test (TC=0) and an invalid ELA/reading test (TC=5). This split testing impacts participation rates for each content area.

Filtering, Student Identification, and Assignment of Accountable School. These rules are required to define the *Accountability dataset* that is unique to school grading and general reporting.

1. Reassign tests to the location where a student is FAY=Yes. For a small number of students who move during the test window, the location of the assessment does not match the location where the students were fully enrolled the prior year (FAY). These students are reassigned to the FAY school for accountability.

    A student can take the test in more than one school (e.g., mathematics in school N, and ELA in school P). Special rules apply and are explained in *V. Calculations*.

2. Limit records to students in eligible grades 3 through 11 and also to students in later grades in high school who have never been tested for accountability. Occasionally students in other grades are tested because assessments are also used in high school graduation, and these students are removed from the accountability determinations and calculations.

3. Use enrollment data to update missing values in student characteristics (i.e., ethnicity).

4. Determine the treatment of duplicate records. This process uses a set of rules performed by the Accountability and Assessment Division, together with district officers and test vendors. Where duplicates cannot be resolved, both duplicate pairs are removed from the accountability dataset.

5. Determine the treatment of invalid student IDs. A small number of tests are unidentified each year with either a missing or invalid ID. After all attempts to identify these students manually have failed, the tests will be included with the location where they were submitted, with a dummy id assigned by the PED.

6. Assign subgroup membership and demographics from snapshot data.

FAY is determined entirely from enrollment submissions. For verification, an LEA may run reports in STARS that show a student's snapshot history. FAY is determined secondarily by the grade configuration of the accountable school.

1. <u>FAY=Yes</u> 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:

   A. Students in <u>transition grades</u> (e.g., 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.

   B. Students in <u>reorganized schools</u> in the current year are FAY=Yes under the same provision as transition grades 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 A.1) and A.2) to be considered FAY=Yes.

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

   D. State-authorized charter schools follow the same options above but without the requirement for LEA membership in the prior year.

2. <u>FAY=No</u> when a student misses any single snapshot in the series. Mobile students are not dismissed from school grading proficiencies. However, a school's expected growth outcome (VAM) is adjusted slightly to account for student mobility.

Student matching. A longitudinally linked data file includes every student's scaled score and proficiency level for the current year and the scaled scores that they received in ELA and mathematics for the two most recent prior years within the last four years. Thus, students who are missing a prior year of data as a result of a prior grade level being untested (such as 9th grade prior to 2015) will not be removed. Students with ambiguous student IDs or who are identified in an unresolved duplicate pair (see *Data Conditioning,* Duplicates) are not included in the linked dataset. Missing data are estimated using the average prior scores for each student's grade level and school.

Determine Q1 or Q3. Students are assigned as either Q1 or Q3 in the current test year. This determination comes from the following decision rules:

1. All students in the test file for the current year are aggregated at the school level. This aggregate includes all students who took any assessment in the current year and disregards a student's grade level. For example, it is possible to have more 4th graders than 3rd graders in the Q1 group for a single school.

2. All of the school's scaled scores are ranked from highest to lowest. A cut score is established for each school that divides the bottom 25% (Q1) from the remaining students (Q3) in mathematics and in ELA separately. These cut scores are unique to each school and are recomputed each year.

3. After the cut score for each school is established, each student is then classified as either Q1 (below the cut) or Q3 (above the cut) for the current school year. Every student fits into either Q1 or Q3 groups and no student is excused.

The student's Q status is determined separately for ELA and for mathematics, so it is possible for a student to be Q1 in mathematics and Q3 in ELA at the same time.

Categorize students into Q1 or Q3 Subgroups. The Q status of both prior and current years determine a student's Q1 or Q3 membership for the current school grading calculations. The student's assignment for the current year's accountability is derived from the following decision rules:

1. The student is matched to his or her prior scores regardless of school or grade level. Matching occurs for the current and prior two years, yielding up to three data points per student.

2. The student's assignment to either the Q1 or Q3 subgroup is determined by their <u>first in the series</u> of the matched scores.

To illustrate, Jack's assignment would come from his score of 55 from the prior year, Alicia's from the current year, and Tom's and Javier's from 2 years ago.

MATHEMATICS

	2 Years Ago	1 Year Ago	Current Year	**Subgroup Assignment**
Jack	(no test)	**55 (Q3)**	62 (Q3)	*Q3*
Alicia	(no test)	(no test)	**33 (Q3)**	*Q3*
Tom	**13 (Q1)**	25 (Q3)	26 (Q3)	*Q1*
Javier	**0 (Q1)**	(no test)	33 (Q3)	*Q1*

In the subsequent year this three-year window will shift, with the earliest year's scores dropping off and the second year scores becoming the first in the series. In this example, all 4 students would be assigned to the Q3 subgroup in next grading season.

The student's accountable school is used as the school of record for school grading, representing the final year of growth and achievement. In the example, it is evident that Javier's school will receive considerable credit for his dramatic growth from 0 to 33 scaled score points, which will be reflected in *Student Growth Lowest Quartile (Q1)* since he belonged to that subgroup.

To summarize, for longitudinal comparisons every student's scores from up to two prior years are used. Their Q1/Q3 membership is determined by the first score in any progression.

- If the student has three years of test scores, their quartile membership is determined from the first y ear.
- If a student has only two years, their membership is determined by the prior year.
- If the student is new to PED schools, then membership comes from their current year. The current year counts as the third of the three-year progression.

A new set of students will enter the progression every year, in every school. Likewise, the years of progression will move up for every student, so that a student who had scores from grades 3, 4, and 5 last year will now be using their scores from grades 4, 5, and 6 this year, which will redefine their membership in Q1 and Q3.

A school's quartile memberships shift yearly as well as bringing in new students, but in general, the school should see some stability in their overall impact on *Student Growth*. If the school is consistently growing the achievement of their lower performing (Q1) students, it will show each year even though some students leave and new students enter.

Scaled Score Standardization. School grading is based on achievement scores from several tests and vendors. The differing assessments have different score ranges and cut scores. In order to combine assessments for grading, each is transformed to a scale from 0 to 200 where 0 is the minimum for the test, 100 is the cut-score for proficiency, and 200 is the maximum. Between these anchor points the transformation is linear.

The transformation is:

if ($x_i \leq cs$) $y_i = 100 (x_i - min) / (cs - min)$

if ($x_i \geq cs$) $y_i = 100 + 100 (x_i - cs) / (max - cs)$

Where: x = the original scaled score
        cs = the proficiency scaled cut score for x
        min = the minimum original scaled score for x
        max = the maximum original scaled score for x

Prior Achievement Scores for Growth. Most students have two prior ELA and math scores, one each from the previous two years for the corresponding subjects. There are exceptions.

1. For early grades KN-2 students are assessed in ELA only so there is no growth for math for these students. Further, because cognitive development is often more rapid at these ages than for older students, the first ELA prior is average of the beginning-of-year (BOY) score and middle-of-year (MOY) score from the current year and the second ELA prior is the end-of-year (EOY) score from the previous year. Because kindergarten does not have a previous year, the first ELA prior comes from BOY and the second ELA prior score is MOY score for the current school year.

2. Because math does not get assessed until third grade, the third and fourth grade students do not have a complete set of prior math scores. For these students, the first and second ELA prior scores are used to predict the current math score.

3. In cases where a student is missing a score from the prior two years, then scores from the prior three and four years are used.

Imputing Missing Values. Students within the PED system are rarely missing prior scores; however a small number of students that are new to PED schools will not have prior scores. If a student group is missing any of the necessary prior scores the average score for students in their grade and school is used.

V. Calculation Parameters

Software. Calculations are performed with IBM SPSS© Version 23 available at http://www-01.ibm.com/software/analytics/spss/. Additional calculations utilize R[©] (the R Foundation for Statistical Computing) available at https://cran.r-project.org.

Rounding. No rounding occurs until final points and rates are computed. All computations prior to reporting use unrounded figures out to 12 decimal places. Final rounding occurs to the second decimal place (e.g., 92.27%) unless otherwise indicated. The terminal digit of 5 is rounded up.

Points. Points for the School grading indicators are achieved by one of two methods:

1. Comparison to a pre-existing standard or

2. Relative to the performance of other New Mexico schools.

Point distributions and cut scores were fixed from preliminary grades established in 2011 and 2012. The first point assignment method is used for *Current Standing (Proficiency)*, *OTL Student Survey*, *Graduation*, and *Attendance*. The second point assignment method is used for remaining indicators.

Letter Grades. The cut points, or point boundaries leading to the letter grades of A, B, C, D, and F, were determined by the overall point distributions in the state (elementary/middle schools as one group and high schools as another) in the first year of school grading. The "A" boundary was equated at the 90th percentile of school points. The "C" boundaries were placed around the 50th percentile. The letter grades for all the supplementary indicators of school grading were based on the 90th, 70th, 50th, and 30th percentiles, with the exception of Opportunity to Learn*,* which was based on 90th, 80th, 70th, and 60th, for A, B, C, D, and F, respectively. The overall grade is based on the 90th, 70th, 40th, and 10th percentiles (values rounded). Overall and indicator letter grade point frameworks are appended to this document.

Growth. The grading model relies on a year's worth of growth, which on the New Mexico scale is equal to a growth index of 0. Growth for each student is measured in relation to how a particular student scored in the current year compared to his or her academic peers. Academic peers are students who scored about the same in the two prior years in ELA and mathematics. A student who scored the same as the average of his or her academic peer group has made one year's worth of growth. A benefit of such a model is that it is easy to determine if students are demonstrating more or less than a year's worth of growth simply by whether the growth score is positive or negative. Another advantage of this scale is that the standard error of measurement is both small and stable across the grade levels.

Value-Added Modeling (VAM) is used to estimate certain elements of school improvement. The VAM framework was established for use at PED in 2011 for application to future years. To determine the anchors, each school's VAM estimate was derived from their position in a distribution of all schools. This position was then used to assign point boundaries (see *Definitions, CDF*). The VAM distribution and its associated cut points from the base year of 2011 were frozen for use in the evaluation of future years. In 2015, the value-added models were adjusted to accommodate New Mexico's shift to the PARCC test, but the cut points remain consistent with those established in 2011.

To arrive at the anchor values, the VAM estimates are converted to *t* scores, and cut points are established at the point where a set proportion of schools fall below. VAM was estimated simultaneously for *School Growth* and *Current Standing* using methods that take into account prior achievement.

Reading (ELA) and Mathematics. These two subject areas are equally weighted throughout school grading for all grade levels. A school's grading indicators (e.g., *Current Standing*) summarize both content areas taken together and equally weighted. The subject areas are detailed separately in the remainder of the repot to help inform factors leading to overall performance. With the introduction of new assessments in 2015 reading is synonymous with English language arts (ELA).

Schools Rated. Grades are calculated for public, locally-authorized charter, and state-authorized charter schools. School calculations exclude off-site locations, programs, and students tested in those locations. These students are rolled up into LEA accountability or are reassigned to an accountable school where possible.

VI. Calculations

Each indicator of school grading is assigned its own grade. The points from indicators are then summed together with bonus points for a final overall grade. The indicators that make up each of the two models are outlined in the model summaries appended to this document.

A. Current Standing is computed identically for both EL and HS models, although each receives different

point totals, 40 and 30 points respectively. Current Standing has two components: 1) an unconditional proficiency status model and 2) a conditional value-added model (VAM, also referred to as a growth model). Unconditional proficiency status is the familiar concept of the percentage of students who achieve proficient or advanced at the most recent testing occasion. The conditional growth or VAM accounts for the school's size, student mobility, and previous scores and measures how much each school is above or below their predicted value. A strong positive relationship exists between a school's proficiency rate and a school's average scaled score.

1. Proficiency (sometimes referred to as "status") refers to the percent of students who are proficient or above in the current reporting year and uses the goal of 100%. The resulting percentage is rounded for reporting while the unrounded figures are used to compute points. Those familiar with AYP will recognize proficiencies as the basis for that accountability model, and *Current Standing* is the only portion of school grading where this concept is used. Proficiency percentages account for valid tests only, after invalid tests and students exempted from testing are removed.

   Points for proficiency are awarded based on an equally weighted combination of mathematics and ELA, except in grades KN-2 where there are only ELA scores. Each are computed as the proportion of students who are proficient multiplied by the number of available points. For example, in 2017 elementary mathematics is worth 10 points; a school with 50% proficient would score .50 X 10 =5 points.

2. Current Standing VAM is calculated based on students' scaled scores and is used to isolate the school's effect on student achievement from the pre-existing differences in school setting and population. Unlike proficiency, VAM calculates growth for all the students. The school's size (enrollment), prior achievement scores, alternate assessment indicator (whether a student is taking alternate assessment), and mobility (proportion of the student body that is FAY) is used to predict the school's score. Schools composed of students who exceed expectations have higher value-added scores than schools composed of students who fall below expectations. The procedure that is used to compute these scores is called Hierarchical Linear Modeling (HLM, also known as multilevel or mixed effects regression), and is described in the technical details below. A positive score from VAM indicates that a school has exceeded expectations based on the parameters described above. A negative score means that a school has fallen below the expectations, while a score of zero indicates that she school has met the expectations.

> **TECHNICAL DETAIL**

The *Current Standing* VAM/growth model uses a multilevel (HLM or mixed-effects) regression model that recognizes two levels, students (level 1) who are then nested within schools (level 2). Levels are referred to with subscripts such that $score_{ij}$ refers to the *i*th student's score and that student is in the *j*th school.

The variables used to predict these scores are shown in the following equation. The growth model estimates values for all B coefficients in the equation and predicts an individual student's score. Separate VAM models are run for each test code (see *Definitions*, *Test Code*), which are then combined into a weighted average by subject.

$$SubjectScore_{ij} = B0 + B1\ SubjectP1_{ij} + B2\ SubjectP2_{ij} +$$

$$+\ B3\ FAY_{ij} + B4\ ALT_{ij} + B5\ n_j + B6\ MOB_j + u_j +\ e_{ij}$$

Where:

$SubjectScore_{ij}$ = the predicted score for student *i* in school *j*

B0 = the intercept. A student having zero values on all of the remaining variables in the equation would have this value as his or her predicted score, which is also the grand mean across all schools.

$SubjectP1_{ij}$, $SubjectP2_{ij}$ = the student's two prior subject scores (i.e., ELA or math) that came from any four prior years, the most recent valid scores available. Prior scores were derived within the same content area (ELA or mathematics) with the exception of 3rd and 4th graders as described above.

FAY_{ij} = whether or not the student is considered full academic year status (FAY).

ALT_{ij} = whether or not the student took an alternate assessment.

n_j = the size of the school, represented by the number of students assessed in grades K-8 (EL model), and grades 9-11 (HS model).

MOB_j = the percent of students in the school who are not FAY (or are mobile).

Uj=school level random effect

e_{ij}=student level random effect

The variables, *B0* through $B6MOB_j$, form the *fixed* component of the model, which is linear and additive. There are no interaction terms in the equation, nor is a curvilinear relationship recognized.

After accounting for the fixed component of the model, u_j, a random component, which represents each school's unique contribution, and which is not explained by school level predictors *(i.e., n_j and MOB_j),* is considered. This component reflects the extent that the average scaled score for a school is greater than, equal to, or less than the state averaged scaled score. The Empirical Bayes (EB_j) estimates are estimated from knowing the mean and standard deviation of u_j.

A final component, e_{ij}, represents the student contribution that reflects those aspects of a scaled score that we cannot control for, after the fixed student-level effects (i.e., student's prior Math/Read scores, FAY_{ij}, ALT_{ij}) are accounted for.

<u>Empirical Bayes</u> residuals are estimated from knowing the mean and standard deviation of u_j. From this model each school earns an EB residual which is comprised of the fixed and random (school effects) components, and the residual can be negative, positive, or zero:

- Positive means students in the school tended to score above their predicted values.
- Negative means students in the school tended to score below their predicted values.
- Zero means students in the school, on average, performed as predicted.

Three steps convert EB_j into points for school grading:

1. Standardize EB_j by dividing the value by the standard deviation of the EB_j :

    $z_j = (EB_j - Mean\ EB_j\ /\mathrm{sd}(EB_{jj}))$

2. Transform standardized EB_j values (z_j) into a probability ranging from 0 to 1, where $z_j = 0$ corresponds to .5:

    $VAMScore_j = \mathrm{CDF}(z_j)$

    where $\mathrm{CDF}(z_j)$ is the cumulative normal distribution function which maps z_j values of -3, -2, -1, 0, 1, 2, and 3 onto: .001, .02, .16, .50, .84, .98, and .999, respectively.

3. Multiply $VAMScore_j$ by the maximum number of points for the indicator:

    7.5 X $VAMScore_j$ for EL Model

    5 X $VAMScore_j$ for HS Model

The VAM index is computed separately for ELA and mathematics, and both are weighted equally in the final assignment of points.

B. School Growth is computed identically for both EL and HS models and yields 10 points toward the total

score (5 each for ELA and mathematics) for a school. *School Growth* is different from *Current Standing* because it predicts the mean score for a school from the mean values of students' prior scores along with school characteristics.

Growth is calculated using the mean from previous scores to predict the mean score in the current year. The school's difference above or below their predicted mean is weighted by the number of students tested. The weighted average is a measure of each school's growth.

> **TECHNICAL DETAIL**

The calculation of *School Growth* uses a multiple linear regression model. The unit of analysis is the school, not the student. Using the same conventions as in *Current Standing VAM*, the mean for the *j*th school is noted as $MeanScore_j$. Equation 2 is similar to equation 1, except the model predicts the mean scores for the school rather than an individual student so all predictors are at one level (school). As with the *Current Standing VAM*, the model presumes that relationships are linear and that there are no interactions among variables.

$SubjectMeanScore_j = B0 + B1\ MeanPrior1Subject_j + B2\ MeanPrior2Subject_j + B3\ MeanFAY_j + B4\ n_j + Bs\ Test_j + e_j$

The calculations include the categorical variable $Test_j$ in the model and allowing each test code (see *Definitions, Test Code*) to have its own B value. The important result is the residual e_j. The residual for each test is weighted by the count of tests, which are then aggregated test code and averaged for school growth (SG_j).

Calculating SG_j points is similar to EB_j for *Current Standing VAM*:

1. Standardize SG_j value:

    $z_j = (SG_j - MeanSG_j)/\text{sd}(SG_j)$

2. Transform z_j into a variable (probability) that can range from 0 to 1:

    $SGScore_j = \text{CDF}(z_j)$

    where CDF(z_j) is the cumulative normal distribution function that maps z_j values of -3, -2, -1, 0, 1, 2, and 3 onto .001, .02, .16, .50, .84, .98, and .999, respectively.

3. Multiply $SGScore_j$ by the maximum number of points for the indicator, for example, 5 X $SGScore_j$.

The number of points is computed separately for ELA and mathematics, and both are weighted equally in the final assignment of points.

C. Student Growth is similar to *Current Standing VAM* except that scores are calculated separately for two

student subgroups, Q1 and Q3 (see *IV. Conditioning of Data, Determine Q1/Q3*). The points for these two groups are weighted equally. Because there are fewer students in the Q1 than there are in the Q3 subgroup, the students in the Q1 subgroup have more influence, per student, on this portion of the school's overall points.

Student Growth is estimated with mixed effects (HLM or multilevel) regression model that is conducted for each test code for Q1 and Q3 subgroups separately. The value-added scores for each assessment are aggregated using a weighted average. The HS model yields Q1/Q3 point totals of 10/10, while the EL/MS model yields 20/20. Each set of points is divided equally between ELA and mathematics.

Similar to *Current Standing VAM* and *School Growth*, Student Growth is expressed in a difference from expected. For example, if a school has a positive Q1 student growth index the school's students are now performing better than expected, while a negative index means that these students' achievement is less than predicted.

TECHNICAL DETAIL

The scores for this indicator are derived in a manner similar to *Current Standing VAM*, and *School Growth*, but separately for Q1 and Q3 subgroups. The scores are based on the same HLM regression as *Current Standing VAM*:

$$SubjectScore_{ij} = B0 + B1\ SubjectP1_{ij} + B2\ SubjectP2_{ij} +$$

$$+ B3\ FAY_{ij} + B4\ ALT_{ij} + B5\ n_j + B6\ MOB_j + u_j + e_{ij}$$

D. Opportunity to Learn (OTL) is based on attendance and a classroom survey administered to

students or parents. High schools can earn 8 total points (3 for attendance, 5 for the survey) and elementary/middle schools can earn 10 points (5 for each).

Attendance is computed schoolwide for every school, including schools with early grades KN-2 and high schools. Rates are presented as schoolwide percentages, and calculations are detailed in VI. C. Percentages are rounded to two decimal points prior to computing points, and are further rounded to integers for reporting.

1. The target for attendance is 95%. That is, all students enrolled, whether for a short time or a long time, are in school for an average of 95% of their days enrolled. Schools with average rates less than 95% are given partial credit, and schools with greater than 95% can earn more than the maximum allotted points.

2. Ineligible student attendance records are removed prior to calculation:

   - Days Present < 0 or missing
   - Days Enrolled < 0 or missing
   - Days Present > Days Enrolled
   - Students without a corresponding record in *Student Snapshot* at the same location

3. The rate is calculated as the unweighted average of individual student's attendance, which accrue from 40D, 80D, and 120D snapshots. The student's numerator is the total days attended accumulated across three enrollment periods (40th, 80th, 120th day snapshots). The student's denominator is the days enrolled for the same periods. After each student's attendance rate is computed, all student rates are averaged for the school summary.

4. A single student can contribute to the rates of more than one school if the student is mobile.

5. When a school is missing a reporting period (40D, 80D, or 120D snapshots) because of submission error, the rate is computed in the same manner, only excluding the missing reporting period from the days attended and days enrolled. Schools may check their attendance rates by running specialized reports in STARS.

6. The points awarded for attendance result directly from the percentage. To illustrate:

Attendance for an elementary school = 80%

Points = .80 X 5 / .95= 4.21

The school earned 4.21 out of 5 points even though it did not reach the target attendance rate of 95%. The attendance rate is multiplied by 5 total points, and then divided by the target goal of 95%.

Classroom Surveys. The parent and student surveys consist of 10 questions about practices known to promote successful learning. Students beginning in third grade are asked to consider the teachers they currently have and respond about the practices these teachers exhibit. Parent surveys are provided for parents for children in grades KN-2. Students are instructed to fill out the survey for all their teachers. Parent and student surveys have the same structure, and are combined to calculate a school's overall survey score worth 4 points.

1. Survey responses are on a 6-step Likert scale, coded 0-5 from least to most beneficial.

2. Responses are summed for each survey, yielding scores ranging from 0 to 50.

3. The target score of 45 yields 100% of the five total points possible. Schools that average higher than 45 on the total score can earn slightly higher than five points.

4. Scores for all students are averaged for each school. Points are derived from the ratio of the school's overall score to the target score of 45. This ratio is applied to the total possible points to arrive at the indicator score as in *Attendance*.

The points from attendance and the student survey are independently calculated, then summed to form the OTL total.

E. Graduation rates are one-year lagged. That is, the rates that are published in January are for the cohort that graduated by August 1 of the prior year. Calculation of 4-year, 5-year, and 6-year cohort graduation rates uses the *Shared Accountability* method that is described fully in the *Graduation Technical Manual* on the PED website, http://ped.state.nm.us/Graduation/index.html. The method is not repeated here but in general:

1. A rate is generated for every school that has any grade 9, 10, 11, or 12.

2. For new high schools that do not yet have a graduating cohort class, a hybrid school grading model is used. These schools are graded on the remaining non-cohort indicators and they are excused from *College/Career Readiness* and *Graduation*. The resulting total points are adjusted to reflect the reduced number of indicators, and the scale and maximum possible total points are the same.

3. The graduation goal is 100%. The model includes 4-year, 5-year, and 6-year rates, which produce a maximum of 8, 3, and 2 points respectively. Growth of the 4-year rate is worth an additional 4 points yielding a total of 17 possible points for graduation. The extended-year rates include only members of the prior 4-year cohort and do not allow new entrants in subsequent years.

4. Points are awarded when the rate is multiplied by the possible points for that category. For example, a 5-year rate of 80% is equal to .80 X 3 = 2.4 points.

Graduation Growth refers to annual increase in the 4-year graduation rate and is based on three years of data. Growth in the 4-year rate reflects the school's overall ability to help students complete their high school careers in a timely way. The goal is 90% of students graduating in four years, so any school that has a graduation rate of 90% in 2015 is awarded all four points. The slope is calculated (see below) and changed into points.

TECHNICAL DETAIL

Graduation Growth is based on the slope of the 4-year graduation rates for the past three years. The table below shows how these slopes are calculated for schools that have graduation rates for each year (Schools A and B), and for schools that have missing graduation rates (for example, new schools with only two years).

Table 1: Calculating Slopes from Four-Year Rates

	2013	2014	2015	Slope	Method
School A	50%	55%	60%	+5% per year	(2015 - 2013)/2
School B	60%	70%	50%	-5% per year	(2015 - 2013)/2
School C	-	55%	60%	+5% per year	2015 - 2014
School D	60%	-	50%	-5% per year	(2015 - 2013)/2
School E	-	-	40%	no slope	
- no graduation rate for that year					

Schools with only one rate (School E) have no slope. For these schools the points for their other graduation components are adjusted to account for the absence of growth.

Growth can be conceptualized as a regression line:

$$GradRate_{ij} = B0^{(j)} + B1^{(j)} Year_{ij} + e_{ij} \qquad (3)$$

Where:

$BO^{(j)}$ = the intercept for the individual school.

$B1^{(j)}$ = the slope for the individual school.

$Year_{ij}$ = the year.

The slopes depicted in Table 1 are a simplification of this method.

The slope is divided by the standard deviation of all slopes, resulting in some positive and some negative values. These values are then transformed using a CDF into a score that can range from 0 to 1. Points are derived by multiplying the CDF value by the points possible, with the qualification that any school where the rate is higher than the goal of 90% receives all points regardless of their slope.

SAM Graduation Rate

A non-cohort graduation rate augments the 4-year cohort rate to arrive at a compiled SAM graduation rate. The numerator and denominator are derived from non-cohort members since cohort members are already counted in the regular 4-year rate. Using this method, schools receive feedback on their success in graduating returning dropouts and adults whose cohort is no longer part of the accountability system.

Both the 4-year cohort and non-cohort figures are combined for the final SAM graduation rate using the weighting scheme shown below. The SAM graduation rate is then used to assign points for SAM schools on the same point framework used for non-SAM graduation rates.

The formula for the SAM graduation rate follows:

$$\text{SAM Graduation Rate} = \left(4CGR * \frac{4CGR\ Denominator}{Total\ Denominator}\right) + \left(NCGR * \frac{Count\ of\ all\ 40D\ Seniors}{Total\ Denominator}\right)$$

Where: $4CGR = \left(\frac{Sum\ of\ fractions\ of\ snapshots\ of\ 4\ year\ cohort\ members\ who\ graduated}{Sum\ of\ fractions\ of\ snapshots\ of\ all\ 4\ year\ cohort\ members}\right) * 100$

$NCGR = \left(\frac{All\ non-cohort\ graduates}{All\ non-cohort\ students}\right) * 100$

$Total\ Denominator = Count\ of\ all\ 40D\ Seniors + 4CGR\ Denominator$

F. College and Career Readiness (CCR) scores are determined by the percentage of 4-year

graduation cohort members who show evidence of participating in college or career preparation, along with the proportion of those students meeting a benchmark. Evidence of participation and success is established through any of the 14 indicators available to high school students (below).

1. CCR is computed for all high schools that have members of the 4-year graduation cohort for that assessment year. CCR is not computed for elementary or middle schools.

2. Cohorts must have a minimum of three student records to be included. Otherwise the high school qualifies for the non-cohort hybrid model (see *Graduation 2.x*).

3. CCR calculations use the shared accountability method used for high school cohort graduation rates. For details see the *Graduation Technical Guide* at http://ped.state.nm.us/Graduation/index.html).

4. Any single student attempt (*Participation*) or single student success (*Success*) within the four years of the cohort is counted.

5. CCR is composed of *Participation* (5 points) and *Success* (10 points) yielding a total 15 points in the high school's overall grade. The school grading cut points for participation and success were derived from base rates normalized in 2012. This framework has been held stationary for evaluation of future rates.

6. The following CCR indicators are eligible. All students enrolled in grades 9 through 12 are eligible for participation in one or more of these programs:

- **PSAT** is the Preliminary SAT and is cosponsored by the College Board and National Merit Scholarship Corporation. The assessment yields scores in evidence based Reading and Writing and in Mathematics and offers benchmark scores that indicate college readiness in two age groups, sophomores and younger, and juniors and older.
- **SAT** is a widely used college admissions examination that measures the skills in Mathematics, Reading and Writing.
- **ACT** is a national college admissions examination that is recognized internationally. The ACT yields scores in four areas, English, Mathematics, Reading, and Science, and offers benchmark scores that indicate college readiness in each.
- **Concurrent Enrollment/Dual Credit** in an accredited New Mexico post-secondary institution offering college credit is counted as evidence of post-secondary preparation. All courses that are nonremedial are counted.
- **AP,** Advanced Placement, is a national qualifying examination aligned to 34 college-level courses. Most four-year colleges grant students credit, advanced placement, or both on the basis of the score on the AP exam for that subject. Students do <u>not</u> get credit for enrolling in a high school AP class. They must demonstrate participation and success in the national exam.
- **Career Program of Studies** is a sequence of high school courses that are recognized to lead to industry-recognized certification. Foundations for career readiness are built from the *Carl Perkins Vocational and Applied* Technology grant definitions. To be considered successful, the student must complete all coursework with a C or better and graduate from high school with a regular diploma.
- **ACT Aspire**, an assessment for 10th graders, is designed to utilize a student's achievement to guide career options and the remaining years of high school. The ACT Aspire has the capability to predict outcomes on the ACT. The test assesses English, Mathematics, Reading, Science and Writing
- **AccuPlacer** is a computer-adaptive college placement test offered by College Board that helps institutions of higher education place students in appropriate courses. Questions are chosen for each student on the basis of the answers to previous questions. The Accuplacer consists of Reading, Mathematics and Writing.
- **COMPASS** is a computer-adaptive college placement test offered by ACT that helps institutions of higher education place students in appropriate courses. Questions are chosen for each student on the basis of the answers to previous questions. The COMPASS provides scores in Reading, Writing and Mathematics.
- **IB,** or the International Baccalaureate program of studies, is a standardized and enhanced high school curriculum where students must demonstrate competency in six study areas or earn an IB diploma. The program originated in Sweden and grants credentials that are recognized outside the U.S.
- **SAT Subject Tests,** standardized subtests that complement the SAT, are usually taken to improve a student's credentials for admission to colleges in the United States. Each test is timed at one-hour, and tests are available in multiple subjects related to a student's interests or a college's requirements.

- **TABE**, Test of Adult Basic Education, is an assessment that measures a person's grade level in reading, mathematics, and language. This test is allowable only for designated SAM schools.
- **WorkKeys,** a job skill assessment created by ACT is used by businesses to measure workplace skills of job applicants, and by schools and colleges to help prepare students for the workplace. This test is allowable only for designated SAM schools and assesses Mathematics, Reading, Writing, Science and Social Studies.
- **ASVAB,** the Armed Services Vocational Aptitude Battery, provides a composite score called the Armed Forces Qualification Test (AFQT) score. The ASVAB scores identify occupations that best suit a candidate's abilities and can be used to qualify for enlistment. This test is allowable only for designated SAM schools.

Participation

1. Cohort members count as a participant when they attempt any one or more of the CCR indicators any time during their four-year tenure in high school. Students may make multiple attempts, with multiple indicators, in multiple years. However this results in only a single credit for participation.

2. SAM schools are allowed use of additional indicators ASVAB, WorkKeys, and TABE. These indicators are not available to other high schools.

Success

1. The success rate follows the same calculation as participation, resulting from weighted numerators and denominators from shared accountability. Students who achieve any one or more of the benchmarks (below) or higher are considered successful in the numerator, while students who attempted any program or assessment form the denominator. The success rate is the percent of participants (numerator for participation) that succeeded.

The benchmarks for success are shown in the table below.

AccuPlacer	Minimum Required Score
College-Level Mathematics	50
Elementary Algebra	80
Reading Comprehension	82
Sentence Skills	83
WritePlacer	6
ACT	
Mathematics	22
English Composition	18
Reading	22
Science	23
ACT Aspire	
Mathematics	432
English	428
Reading	428
Writing	428
Science	432
Advanced Placement (AP)	
Art History	3

Biology	3
Calculus AB	3
Calculus BC	3
Chemistry	3
Chinese Language and Culture	3
Computer Science A	3
European History	3
English Language and Composition	3
English Literature and Composition	3
Environmental Science	3
French Language	3
German Language	3
Government and Politics: Comparative	3
Government and Politics: United States	3
Human Geography	3
Italian Language and Culture	3
Japanese Language and Culture	3
Latin: Vergil	3
Macroeconomics	3
Microeconomics	3
Music Theory	3
Physics B	3
Physics C: Electricity and Magnetism	3
Physics C: Mechanics	3
Psychology	3
Spanish Language	3
Spanish Literature	3
Statistics	3
Studio Art: 2-D Design	3
Studio Art: 3-D Design	3
Studio Art: Drawing	3
United States History	3
World History	3
COMPASS	
Mathematics	52
Reading	88
Writing Essay (Scale 2–12)	9
Writing Essay (Scale 2–8)	7
Writing Skills	77
CTE Course Sequence	
Any PED-recognized CTE Pathway	C
Dual Credit	
Nonremedial Course	C
International Baccalaureate (IB)	
Mathematics	4
Literature (English or Spanish)	4
Language and Literature (English or Spanish)	4
Individuals and Society	4
Experimental Sciences	4
Arts	4

IB Diploma	24
PSAT-before November 2015	
Mathematics	47
Critical Reading	45
Writing	45
PSAT-before November 2015	
Mathematics	480
Evidence Based Reading & Writing	430
SAT-before March 2016	
Mathematics	500
Critical Reading	500
Writing	500
SAT-after March 2016	
Mathematics	530
Reading & Writing	480
SAT Subject Area Tests	
Mathematics Level 1	587
Mathematics Level 2	647
Literature	574
Chemistry	642
Ecological Biology	593
Molecular Biology	624
Physics	632
U.S. History	610
World History	589
French	601
French with Listening	626
German	608
German with Listening	594
Spanish	619
Spanish with Listening	640
Modern Hebrew	586
Italian	671
Latin	586
Chinese with Listening	739
Japanese with Listening	662
Korean with Listening	749

In addition SAM schools may apply the following assessments.

ACT WorkKeys	
Applied Mathematics	5
Listening for Understanding	4
Reading for Information	5
Business Writing	3
Applied Technology	3
Teamwork	4
Location Information	4
TABE (Complete Battery Subtests)	
Mathematics	506
Reading	518
Writing	524

		ASVAB (Comprehensive)					
	AFQT		31				

This example shows how students Joe and Jane contributed to the CCR rates for the multiple high schools they attended during their 4-year tenure. Note that each student's fractions will always add to 1.0 and are the equivalent of one student statewide. Note also that the attempt (participation) and success flags are distributed to all schools the student ever attended. This means that all high schools are held accountable for readiness no matter the school's grade configuration or the mobility of the student.

Student	High School	Count of Snapshots in School	Count of Snapshots in State	Fraction for that School	CCR Attempted* (Y/N)	CCR Success* (Y/N)	CCR Participation Numerator	CCR Success Numerator
Joe	Pine	4	16	0.25	Y	Y	0.25	0.25
Joe	Elm	8	16	0.50	Y	Y	0.50	0.50
Joe	Oak	4	16	0.25	Y	Y	0.25	0.25
Jane	Cedar	2	10	0.20	Y	N	0.20	0.00
Jane	Pine	5	10	0.50	Y	N	0.50	0.00
Jane	Oak	3	10	0.30	Y	N	0.30	0.00

* Multiple attempts or multiple successes do not change the student's weight in the calculation. The student's single best effort is used, and is distributed to all high schools attended.

<div style="border:1px solid;">

TECHNICAL DETAIL

</div>

The calculation of a school's points is illustrated for CCR participation, which is worth a total of 5 points:

1. StandardizedRate = (SchoolRate - Mean$_{SchoolRate}$)/(SD$_{SchoolRate}$)
2. NormalizedScore = CDF(StandardizedRate, N) where CDF is the cumulative distribution function (see Definitions) which ranges from 0 to 1.0.
3. PointsAwarded = NormalizedScore X 5

G. Bonus Points

Bonus points are derived from three sources, 1) the PED's data warehouse STARS, 2) the New Mexico Activities Association (NMAA), and 3) schools and districts use of *Next Step Plans* and *Student Assist Teams* submitted to an online system. Scores and points are given for these categories:
- Student and parent engagement (maximum 2 points)
- Habitual truancy improvement over the prior year (maximum 2 point)
- Activities (maximum 1 point)
- Use of online assessment format (maximum 2 points)

1. Parent and student engagement points are based on use of Next Step Plans and/or Student Assist Teams. Details are available at http://aae.ped.state.nm.us/.
2. Improvement in habitual truancy is calculated from STARS submissions from LEAs.

3. Student participation in athletic activities is derived from data supplied by NMAA, which includes counts of participants. This indicator is for high schools only, and elementary and middle school bonus points are adjusted for the exclusion of these points so that both models can earn up to five bonus points.

4. The use of online assessment format is computed as the percentage of students who took the assessments in computerized format as opposed to paper. The percentage is derived only for assessments that have multiple formats.

5. Bonus points are added to the school total points prior to determining the school's overall grade. Bonus points are capped at five.

H. Early Grades (KN-2)

1. <u>Participation</u> rates are not calculated for schools that do not have third grade or higher.

2. <u>Attendance</u> rates are generated for these using the same methods as for other schools.

3. *Current Standing*. *School Growth, and Student Growth* indicators are calculated identically, although the early literacy assessment is used as the primary achievement measure. The points a school earns for ELA in each component of school grading are doubled to account for the lack of mathematics.

I. Abbreviated Model

A small number of schools do not have enough data, either historical or current, to calculate a component. This may include lack of prior data for growth, or for current data the lack of examinees. For these schools, an alternate model that relies only on the remaining components is used. As with schools with only grades KN-2 the overall points are adjusted to meet the 100 point scale used for all of school grading.

On the summary of the school grading report card the indicator scores that are not missing should be interpreted the same as for any other school. The adjustment to a 100 point scale takes place only for *Overall Points*. Schools are not penalized for missing data that is deemed outside of their control.

J. Supplemental Accountability Model (SAM) Schools

SAM schools receive special considerations in VAM (*Current Standing*), *CCR*, *Graduation*, and *Participation*. Schools qualify for the SAM model under two conditions: 1) when the percentage of students with disabilities in the school is higher than typical (20% in 2013), or 2) when the percentage of students aged 19 or older is higher than typical (10% in 2013). Schools may also petition to become a SAM school if the publicized mission of the school is to serve students who have traditionally found a regular educational setting difficult, such as pregnant teens or behaviorally challenged students.

Schools that specialize in educating higher risk students typically experience higher student mobility that complicates the calculation of cohort rates and other time-dependent measures. Therefore, these schools are allowed to supplement their accountability with additional rigorous measures that demonstrate that they have met their mission. For example, the accountability for a school that specializes in credit recovery for returning dropouts will include their ability to build a student's credit bank and to graduate non-cohort members in the current year. These supplemental measures are negotiated with the PED and are in addition to, not in place of, the accountability measures required of non-SAM schools.

<u>Current Standing VAM</u>. When conditioning data for *Current Standing*, an offset is applied to level the playing field for SAM schools. The offset is based on the mean deviation of scaled scores for SAM schools.

The offset is recomputed every year and is dependent upon the ELA and mathematics performance of students in SAM schools. A value is added to the conditional mean for each SAM school for mathematics and ELA prior to dividing by the standard deviation.

Graduation. An auxiliary graduation rate is computed using a *senior completer* method which includes only 12th grade students who are not members of the 4-year cohort. The denominator is comprised of the count of 12th graders in the first enrollment snapshot (40D). The numerator is derived from the count of all non-cohort students who graduated by the end of the year (EOY snapshot). Note that these counts do not track the same students but are considered only samples of typical senior enrollment and of typical non-cohort graduates. It is possible for a school to have a graduate count at the end of the year that exceeds their senior population at the 40th day snapshot. This would occur when many students undertake brief periods of rapid credit recovery and graduate and exit throughout the year, while the school maintains a relatively stable student count.

The numerator and denominator are derived from only non-cohort members since cohort members are already counted in the regular 4-year rate. Using this method, schools receive feedback on their success in graduating returning dropouts and adults whose cohort has long since aged from the system. The rate calculations are covered in *E. Graduation.*

College and Career Readiness (CCR) rules were broadened for SAM schools to include additional career readiness indicators that were in keeping with the mission of the school. In addition to the measures available to other schools, WorkKeys, TABE, and ASVAB can be used by SAM schools.

Participation in assessments requires that 95% of enrolled students be assessed. With the high student turnover in most SAM schools, this rate becomes unrealistically high and unachievable. Therefore SAM schools were excused from a participation penalty, although PED will continue to monitor SAM schools to ensure that students are not systematically being excluded from assessment.

VII. Participation in Assessments

All enrolled students in eligible grades and courses must be tested with the appropriate regular or alternate assessment. The benchmark for participation is 95% of the *All Students* group. Schools that fail to meet the minimum of 95% have their letter grade reduced by one letter, and a penalty flag placed on their report. Schools that are in "F" status will remain "F" and will have a penalty flag placed on their report.

1. The participation rate is a percentage of all valid and scorable tests divided by a school's enrollment in eligible grades, or in the case of PARCC high school math, enrolled in eligible courses. The percentage is rounded to the nearest integer.

2. Two participation rates are calculated, one for ELA and one for mathematics. If either rate falls below 95% the school fails to meet the requirement.

3. For all schools, the unweighted average of rates from the current and prior two years is applied as a second tier of evaluation.

4. The denominator for participation is derived from enrollment counts of all students in the 120D, Accountability, and EOY snapshots. If the rate is met in one of the snapshot considerations, the school is considered to have met participation. Only students in eligible testing grades 3-11 are counted towards participation.

5. The numerator for participation is the count of all valid and scorable tests. Students who received an ELA exemption (TC=3) are counted for participation provided they have successfully completed the English language proficiency examination in lieu of the ELA assessment.

6. Students who are FAY=Y at one school (their *Accountable School*), but take the test at another school, are reassigned to their accountable school. Students who are FAY=N are assigned an accountable school based on their enrollment status closest to their testing date.

7. Examinees whose tests are invalidated for various reasons (see *IV. Conditioning of Data, TC Codes*) do not receive a score and do not count in the numerator for participation.

8. Examinees whose tests are exempted (*Medical Exemption, ELA Exemption, 9th Grade Spanish ELA Exemption, Course Less than 50% Complete Exemption*) are excused from both the numerator and the denominator of the participation rate.

9. Students who were identified as having transferred out of the system are excused from both the numerator and the denominator of the participation rate.

10. Participation rates are not computed for schools with only grades earlier than third grade. These schools, sometimes called "feeder schools" are exempted from the participation rate requirement.

11. A participation penalty is applied only to schools with 100 or more eligible students.

12. Schools designated as SAM schools are excused from the participation penalty.

VIII. LEA Grades

Local educational agency (LEA) report cards are required by federal statute: Elementary and Secondary Education Act (ESEA) section 1111(h)(2) as amended by the No Child Left Behind Act of 2001 and as amended by the Every Student Succeeds Act. Grades in the report cards have been calculated and reported since 2012 using the method that follows.

1. LEA grades are calculated and published annually after A-F School Grading appeals for schools have been fully adjudicated, typically in September.

2. Summary letter grades are calculated for multi-school LEAs. State-authorized charter schools qualify as a single-school LEA. Therefore these schools' A-F School Grading letter grade serves as their LEA rating and an additional score is not needed.

3. The LEA grade results from a weighted average of the schools within the district.

   A. The process weights only the schools' overall points and not individual component scores (e.g., *Q1 Student Growth*).

   B. Because the LEA's average is weighted by the school's size, larger schools contribute more to the LEA grade than smaller schools.

   C. The weight for each school is derived from the number of students in kindergarten through grade 12. Student counts are derived from the STARS 120D snapshot for the year of the accountability assessment.

   D. Weighting ensures that the LEA is held accountable for all of their students and that each student carries the same weight.

   E.

4. The LEA's weighted overall score is converted to a letter grade using a framework that is a compromise between the two models used in school grading. Elementary and middle schools have slightly different cut points for letter grades than do high schools. The LEA points are rounded to a single decimal (i.e., tenths) prior to the assignment of a letter grade.

LEA Letter Grade Framework

Grade	Overall Points
A	75.0 or More
B	62.5 to 74.9
C	50.0 to 62.4
D	36.2 to 49.9
F	36.1 or Less

5. Like schools, LEAs must meet the 95% participation threshold. Failure of an LEA to test the minimum percentage of eligible students results in lowering the letter grade by one. The participation rate calculation follows the procedure for schools except for the following:

   A. Students in off-site settings (i.e., correctional facilities, treatment centers, homebound, or hospitalized) are included.
   B. Three-year averaging is not applied.

   Participation failures are noted on the reporting of the LEA grades by an asterisk with a footnote: "This LEA's grade was lowered by one letter because it failed to test 95% of all students."

Illustration of LEA Grade Weighting

School	Y Overall Points	Z Enrollment (N)	Y x Z	Averaged Total Points Per Student	The final score of 65.0 yields an LEA letter grade of B.
Hilltop	60	100	6,000		
Elm	70	50	3,500	(13,000/200)	
Cedar	70	50	3,500		
		200	13,000	65.0	

The weighted mean score is 65.0 which produces the letter grade of B. An unweighted mean of the three schools would be 66.7 [(60+70+70)/3]. However, because Hilltop had twice as many students as the other two schools the weighted mean shifts slightly toward that school's overall score.

In practice, the overall points are computed using 14 decimal places until arriving at the final *Averaged Total Points per Student*, at which time the points are rounded to the nearest tenth prior to the assignment of a letter grade.

Appendix A: Elementary/Middle Model and Points

Overall Model and Points - Elementary and Middle Schools			Points	
Current Standing How did students perform in the most recent school year? Students are tested on how well they met targets for their grade level (Proficient).	Percent Proficient	25		40
	Value-added conditioning of proficiencies, accounting for school characteristics for the past three years.	15		
School Growth In the past three years did the school increase overall performance?	Value-added conditioning of performance, taking into account school characteristics for the past three years.	10		10
Growth of Higher Performing Students (Q3) How well did the school help individual students improve? The highest performing students are those whose prior scores placed them in the top three quarters (75%) of their school.	Individual student growth over the past three years is compared to the average for the state.	20		20
Growth of Lowest Performing Students (Q1) How well did the school help individual students improve? The lowest performing students are those whose prior scores placed them in the bottom quarter (25%) of their school.	Individual student growth over the past three years is compared to the average for the state.	20		20
Opportunity to Learn Does the school foster an environment that facilitates learning? Are teachers using recognized instructional methods, and do students want to come to school?	Attendance for all students	5		10
	Classroom/parent survey	5		
Total				100
Student and Parent Engagement Does the school involve students and parents in education, reducing truancy, and promoting extracurricular activities?	Bonus Points			+5

Appendix B: High School Model and Points

Overall Model and Points - High Schools		Points	
Current Standing How did students perform in the most recent school year? Students are tested on how well they met targets for their grade level (Proficient).	Percent Proficient	20	30
	Value-added conditioning of proficiencies, accounting for school characteristics for the past three years.	10	
School Growth In the past three years did schools increase overall performance?	Value-added conditioning of performance, taking into account school characteristics for the past three years.	10	10
Student Growth of Higher Performing Students (Q3) How well did the school help individual students improve? The highest performing students are those whose prior scores placed them in the top three quarters (75%) of their school.	Individual student growth over the past 3 years is compared to the average for the state.	10	10
Student Growth of Lowest Performing Students (Q1) How well did the school help individual students improve? The lowest performing students are those whose prior scores placed them in the bottom quarter (25%) of their school.	Individual student growth over the past 3 years is compared to the average for the state.	10	10
Opportunity to Learn Does the school foster an environment that facilitates learning? Are teachers using recognized instructional methods, and do students want to come to school?	Attendance for all students	3	8
	Classroom survey	5	
Graduation How does the school contribute to on-time graduation? On-time means within 4 years, and to a lesser extent, within 5 and 6 years for students who require longer.	Percent graduating in 4 years	8	17
	Percent graduating in 5 years	3	
	Percent graduating in 6 years	2	
	Value added conditioning taking into account school characteristics for the past 3 years.	4	
Career and College Readiness Are students prepared for what lies ahead after high school? Schools receive credit when students participate in college entrance exams and coursework leading to dual credit and vocational certification. The school receives additional credit when students meet success goals.	Percent of all students that participated in one of the alternatives	5	15
	Percent of participants that met a success benchmark	10	
Total			100
Student and Parent Engagement Does the school involve students and parents in education, reducing truancy, and promoting extracurricular activities?	Bonus Points		+5

Appendix C: Point Boundaries for School Grading 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
Student Growth of Higher Performing Students (Q3)	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
Student Growth of Lowest Performing Students (Q1)	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 practice, figures are carried out to six or more decimals. Therefore, letter grades at the highest and lowest boundary of a point span may not be apparent because of rounding.

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	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
Student Growth of Higher Performing Students (Q3)	A	6.8 or above
	B	4.3 to 6.7
	C	2.9 to 4.2
	D	1.5 to 2.8
	F	1.4 or below
Student Growth of Lowest Performing Students (Q1)	A	9.3 or above
	B	8.3 to 9.2
	C	7.1 to 8.2
	D	5.7 to 7.0
	F	5.6 or below
Opportunity to Learn	A	7.2 or above
	B	6.4 to 7.1
	C	5.6 to 6.3
	D	4.8 to 5.5
	F	4.7 or below
Graduation	A	15.3 or above
	B	13.6 to 15.2
	C	11.9 to 13.5
	D	10.2 to 11.8
	F	10.1 or below
College & Career Readiness	A	11.3 or above
	B	9.7 to 11.2
	C	8.2 to 9.6
	D	6.7 to 8.1
	F	6.6 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

History and Revisions

Date	Description of Major Changes	Reference	Author
2010	New Mexico adopts Common Core Standards		
December 2011	Preliminary grades supplemented but did not replace Adequate Yearly Progress (AYP)		
February 2012	A-F School Grading System authorized by USDE with certain modifications		
July 2012	Current Standing VAM eliminates student demographics	VI. A. 4.	Gregory
July 2012	All indicators normalized to 2011	V. VAM	Gregory
July 2012	School Growth VAM eliminates student demographics	VI. B.	Gregory
July 2012	Opportunity to Learn includes student survey	VI. D.	Gregory
July 2012	Current Standing includes 10th grade students	VI. A. 3.	Gregory
July 2012	Graduation adds 6-year rate; denominator to 100%	VI. E. 3.	Gregory
July 2012	Graduation growth uses 3+1 year model	VI. E.	Gregory
July 2012	College Career Readiness (CCR) uses Shared Accountability system	VI. F. 3.	Gregory
July 2012	CCR added additional indicators	VI. F. 5.	Gregory
July 2012	*No Cohort* option for qualifying schools	VI. E. 2.	Gregory
July 2012	Supplemental Accountability Model (SAM) for qualifying schools	VI. I.	Gregory
July 2012	Participation requirement added	VI.	Gregory
July 2012	Bonus Points added	VI. G.	Gregory
July 2012	Schools with grades KN-2 only, method alterations	VI. H.	Gregory
February 2014	Participation uses different denominator	VII.	Gregory
February 2014	SAM CCR rates include Shared Accountability	VI.I.	Gregory
February 2014	Revised growth to recognize HS 10th grade	VI.	Gregory
February 2014	Clarification and examples added	Throughout	Gregory
March 2015	First year, Partnership for Assessment of Readiness for College and Careers (PARCC) assessment		
December 2015	Aligned growth calculation with teacher evaluation model	VI.	Hall
July 2016	Simplified graduation growth	VI. E.	Wright
December 2016	LEA grading description added	VIII.	Gregory
August 2017	Current Standing returns to original weighting prior to PARCC	Appendices	Gregory