Offeror Name: Interactive Achievement

Proposed Assessment Name: Interactive Achievement Student Growth Assessments

Content Area(s) and Grade Level(s) Assessed: Math K – 8, Algebra, Geometry, Algebra II
English K – 11
Science K – 8, Earth Science, Biology, Chemistry
History K – 8, WHI, WHII, VA/US History

Section 1: Overview of Tests

Requirement:
1.1 Describe the specific grade(s) and subject area(s) covered by each assessment and provide an overview of the content and skills measured. Include the types of test items used, the mode(s) of delivery, the availability of equivalent forms, including short forms or screeners (if available) and a test blueprint for each test being proposed.
The Interactive Achievement Student Growth Assessments are available in the four core content areas (Math, English, Science, and History), in kindergarten through end of course subjects. Assessments are available for math in grades K – 8 as well as Algebra, Geometry, and Algebra II; for English in grades K – 11; for science in grades K – 8 as well as Earth Science, Biology, and Chemistry; and for History in grades K – 8 as well as World History I, World History II, and Virginia/US History. The content and skills measured by the assessments are derived directly from the Virginia Standards of Learning and Curriculum Frameworks and are specifically outlined in the attached blueprints. All Interactive Achievement Student Growth Assessments include selected response items and are deliverable via paper and pencil or online. Each assessment is comprised of a single form and does not include short forms or screeners. A test blueprint is provided for each of the 49 Interactive Achievement Student Growth Assessments in Appendix A.

Requirement:
1.2 Provide evidence of alignment of test items to the Virginia Standards of Learning (SOL) for existing assessments. For assessments developed in response to the RFP, provide a plan for assuring the alignment of test items to the SOL.
All test items were written according to the principals of universal item design and were specifically written for alignment to the Virginia Standards of Learning (SOL). Interactive Achievement's Content Specialists

Requirement:
are highly trained item writers and assessment authors as well as experienced Virginia educators. The attached blueprints demonstrate clear evidence of item alignment to the SOL.

Section 2: Technical Characteristics

Requirement:
2.1 Provide evidence of content, construct, concurrent, and predictive validity as appropriate. Include validity evidence that supports the use of scores from the proposed assessment in teacher evaluation, addressing specifically the validity of using assessment results to support inferences about effectiveness of teacher in producing growth in student performance (if available).

All Interactive Achievement Student Growth Assessments are authored by Interactive Achievement's team of highly trained Content Specialists who are item writers and assessment authors as well as experienced Virginia educators. All Student Growth Assessments are constructed to specifically address the Virginia Standards of Learning (SOL) and the essential skills and knowledge outlined in the Curriculum Frameworks. By specifically constructing assessments to address the target content, as opposed to merely aligning existing content to the SOL, the Interactive Achievement Growth Assessments demonstrate evidence of content, construct and predictive validity.

Since 1993, the team behind SAS® EVAAS® reporting has worked with educators, administrators and policymakers to provide the most reliable, comprehensive and accessible set of reporting tools to drive innovation and improve education. SAS' value-added and projection reporting is available statewide in Tennessee, North Carolina, Ohio, and Pennsylvania and to individual districts and schools in many other states, including Arkansas, Arizona, California, Colorado, Connecticut, Delaware, Georgia, Illinois, Indiana, Louisiana, Missouri, New Jersey, New York, South Carolina, Texas, and Virginia.

SAS works with its education partners to determine the appropriate data inputs, modeling choices, policy decisions, release dates, and other customization for its EVAAS reporting. Upon receipt of clean, usable data, SAS engages in a variety of quality checks and processing to create a longitudinal database of students' test scores. The database is used for analytic modeling, and the resulting analyses are displayed via a customized, secure web application. The typical turnaround for the first delivery of our full standard solution is eight to ten weeks to allow for sufficient time for policy decisions by our education partners. In subsequent analyses, EVAAS reporting is typically delivered within six to eight weeks.

The SAS EVAAS models can utilize a variety of testing sources, such as the end-of-grade and end-of course tests typically administered. SAS is familiar with Virginia's Standards of Learning (SOL) tests as a result of our current work with Local Education Agencies (LEAs) in Virginia. In fact, SAS EVAAS is a flexible solution that has the capability of providing value-added results for many different types of

Requirement:

assessments including the following:

- Standardized state-wide assessments for End-of-Grade and End-of-Course subjects
- National Norm-Referenced Tests (NRTs)
- ACT/PLAN/EXPLORE
- SAT/PSAT
- Or any other additional testing, such as Interactive Achievement Student Growth Assessments, given certain criteria are met that are described below

SAS EVAAS can utilize tests with vertical scaling, as well as those that are horizontally scaled. The SAS modeling approaches can utilize historical data, even when tests change over time, assuming they meet our criteria for modeling. Prior to using assessment data with EVAAS, SAS works with the educational partner to determine if the assessments meet three characteristics:

1. Scales must have sufficient stretch to show measurable differences among students at the very top and bottom of the distributions.
2. Scales must be reliable from year to year.
3. The test must be highly correlated with the curricular objectives.

With SAS EVAAS reporting, educators are not held accountable based on a single test score. Our modeling approaches use up to five years of data for an individual student by simultaneously analyzing scores across all grades and subjects. This minimizes the influence of measurement error and increases the precision of the growth and projection estimates.

After qualifying the proposed assessment(s), the SAS EVAAS team, in partnership with VA DOE or LEA, evaluates the data for the three characteristics identified above as well as any other anomalies that might cause concerns with the overall analyses. SAS formally recognizes that the use of this qualified data can not only provide indications of student progress, but also teacher effectiveness. In addition, the U.S. Department of Education has emphasized the importance of value-added assessment models in evaluating teaching effectiveness as part of its criteria in distributing stimulus funds through such programs as the *Race to the Top* program.

For external validation of SAS EVAAS modeling, SAS refers you to the myriad studies and other academic and scientific literature supporting the use of statistical tools to analyze student and teacher performance with the goal of using such analysis in efforts to improve student and teacher performance. By way of example, the following are a sample of such available literature.

- On the choice of a complex value-added model, please consult: McCaffrey, D. F., Han, B. and Lockwood, J. R. (2008). “Value-Added Models: Analytic Issues.” A paper prepared for the National Research Council and the National Academy of Education, Board on Testing and Accountability Workshop on Value-Added Modeling, November 13 and 14, 2008, Washington D.C.
- On the advantages of the longitudinal, mixed model approach, please consult: Lockwood J. R.

Requirement:

and McCaffrey D. F. (2007). "Controlling for Individual Heterogeneity in Longitudinal Models, with Applications to Student Achievement." Electronic Journal of Statistics, Vol. 1, 223-252.

- On the insufficiency of simple value-added models, please consult: McCaffrey, D. F., Han, B. and Lockwood, J. R. (2008). "From Data to Bonuses: A Case Study of the Issues Related to Awarding Teachers Pay on the Basis of the Students' Progress." Paper presented at the conference on Performance Incentives: Their Growing Impact on American K-12 Education, Feb. 28-29, National Center on Performance Incentives at Vanderbilt University.

Recently, WestEd, a non-profit education research and development agency, reviewed and evaluated possible models for use in North Carolina's accountability model, and its final recommendation were models provided by SAS. For more information please refer to http://www.ncpublicschools.org/docs/stateboard/highlights/2012/02highlights.pdf

Furthermore, SAS EVAAS reporting has passed rigorous scrutiny by several federal agencies:

- Variations of the SAS EVAAS individual student projections provided herein have been approved as NCLB AYP growth model augmentations in Tennessee, Ohio, and Pennsylvania. Through four separate Peer Review Committees, the U.S. Department of Education concluded that an EVAAS projection made three years before a student takes a test is more accurate than just considering the student's score in the year before that future test. In other words, EVAAS projections to an eighth grade math test for current fifth grade students are more closely correlated to students' actual performance than their seventh grade math scores. http://www2.ed.gov/admins/lead/account/growthmodel/oh/index.html
- Additionally, the SAS EVAAS student projection methodology was reviewed by the Government Accounting Office (GAO) in 2006 and is referenced in its July report to the House Education Committee, which is available on the official GAO website at: http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=gao&docid=f:d06661.pdf
- Four different US Department of Education Peer Review Committees have approved SAS reporting for use in growth waivers in NCLB. Ohio's growth model waiver provides more details at http://www2.ed.gov/admins/lead/account/growthmodel/oh/ohadd2.doc

In addition to being used in educator evaluation systems in North Carolina, Ohio, and Tennessee, EVAAS is used as one measure of teaching effectiveness in districts across 16 other states. District partners are both large and small, ranging from Houston Independent School District in Houston in Texas with over 200,000 students and almost 300 schools, to Bells City School, a district in Tennessee with less than 500 students in one school. In all of these places, SAS EVAAS has worked closely with policymakers to ensure that the models implemented by the program apply sound statistical reasoning to evaluation or compensation decisions involving EVAAS reporting and minimize the risk of misidentifying teachers and schools.

Requirement:
2.2 Provide evidence of reliability, both for the total test and for any subtests for which scores are reported. Include estimates of error in measurement.
All Interactive Achievement Student Growth Assessments undergo measures of reliability for total test and for SOL Reporting Category (subset for which scores are reported) as part of the assessment development process and criteria for meeting the strict requirements of application of value-added measures by SAS EVAAS. As a component of the joint effort by IA and SAS, the standard measure of error that can be expected when delivering the assessments is being studied over the course of the 2012-2013 school year and will be measured against the results of the 2013-2014 school year and beyond. By normalizing the assessment results and ensuring that there is adequate performance ranges, the results are expected to provide evidence that students can reliably expect to perform in the same percentile on the proposed assessments from year to year. Conversion of student performance to scale scores for the purpose of applying EVAAS measures will also provide the appropriate data for which to conduct measures of reliability and estimate error in measurement.

Requirement:
2.3 Provide evidence that the assessment is appropriate for use with student subgroups, including English language learners and student with disabilities. Include documentation that the assessment does not exhibit bias toward any major subgroups (e.g., through an analysis of differential item functioning). In addition, provide a sensitivity review to demonstrate the assessment tasks and items are designed to be accessible and fair for all students.
All Interactive Achievement Student Growth Assessments undergo analysis of appropriateness for use with student subgroups, including English Language Learners and students with disabilities as part of the assessment development process and criteria for meeting the strict requirements of application of value-added measures by SAS EVAAS. All Interactive Achievement test items are developed using best practices and follow the principles of universal design. Universal design provides a rigorous framework by which Interactive Achievement items are developed that includes multiple stages of review. This framework includes bias and sensitivity review. From a reporting perspective, the SAS EVAAS web-based application allows users to drill down to various student subgroups by achievement level so that educators can assess their success with different kinds of students. From a statistical modeling perspective, the goal of value-added modeling is to provide reliable estimates of student progress, and it is critical that these estimates reflect educators' actual impact, independent of

Requirement:
their students' entering achievement or background. In other words, educators should not be accountable for the things they cannot control and should be accountable for the things they can control. A key principle of SAS EVAAS value-added modeling is to follow the progress of individual students. Consequently, the models include a student's entire testing history in multiple subjects (multivariate) over multiple years (longitudinal). However, socioeconomic (SES) and demographic (DEM) variables are not explicitly included in these models, either at the student level or at any higher (classroom, school, community) level, leading to concerns about unfairness. At the student level, by including all of a student's testing history, each student serves as his or her own control. To the extent that SES/DEM influences persist over time, these influences are already represented in the student's data. This negates the need for SES/DEM adjustment and was confirmed empirically by Ballou, et al. (2004) and by Lockwood and McCaffrey (2007) who conclude: *William Sanders … has claimed that jointly modeling 25 scores for individual students, along with other features of the approach, is extremely effective at purging student heterogeneity bias from estimate teacher effects … The analytical and simulation results presented here largely support that claim.* On a philosophical level, the question educators should ask is whether they should have lower expectations for a student from an economically disadvantaged family than one from an economically advantaged family, even when the two students have identical test scores and academic history. By adjusting for these variables, one is directly assuming that there will be different expectations for two students with the same prior achievement pattern who come from different SES/DEM communities. The use of SES/DEM adjustments at the student level has largely been discouraged among statisticians and policy makers involved with value-added modeling, including the policies developed for Adequate Yearly Progress in growth model augmentations for No Child Left Behind. While technically feasible, the adjustment of student characteristics in sophisticated modeling approaches is not necessary from a statistical perspective and may have unintended negative policy consequences. While some models make adjustments to mask their estimates' correlation with students' background, other models such as those used by SAS EVAAS can utilize each student's full testing history, including those with missing test data, such that each student serves as his or her own control.

Requirement:
2.4 Provide evidence that the assessment includes items of varying difficulty to ensure accurate measurement of student achievement across the ability continuum, including the tails of the score distribution.
Current statistical analysis of the first round of assessment delivery suggest that the Interactive

Requirement:
Achievement Student Growth Assessments include items of varying difficulty to ensure accurate measurement of student achievement across the ability continuum, including the tails of the score distributions. By normalizing the data collected across the varying divisions in Virginia, the initial distributions have shown sufficient stretch to show measurable differences among students at the very top and bottom of the distributions. This provides early evidence that the proposed assessments meet this specific characteristic required for the SAS modeling approaches.

Section 3: Use of Assessment as a Measure of Growth

Requirement:
3.1 Provide evidence that the scores resulting from the assessment have been used as measures of growth by other local or state education agencies.
Interactive Achievement has provided assessment and reporting services to a combination of more than 90 Virginia school divisions for the past 6 years. IA is partnering with SAS to provide evidence that the scores resulting from the proposed assessments can be used as measures of growth and to inform educator evaluation. SAS EVAAS is contracted through the State Education Agencies in North Carolina, Ohio, Pennsylvania, and Tennessee. Additionally, SAS EVAAS is contracted by individual districts and schools in many other states, including: Arkansas, Arizona, California, Colorado, Connecticut, Delaware, Georgia, Illinois, Indiana, Louisiana, Missouri, New Jersey, New York, South Carolina, Texas, and Virginia. In 2011, SAS received test data from 2,125 LEAs, which constituted 6.7 million students. Assessment data used in SAS EVAAS value-added and projection modeling can include: ➢ Standardized state-wide assessments for End-of-Grade and End-of-Course subjects ➢ National Norm-Referenced Tests (NRTs) ➢ ACT/PLAN/EXPLORE ➢ SAT/PSAT ➢ Or any other additional testing, such as Interactive Achievement Student Growth Assessments, given certain criteria are met

Requirement:
3.2 Describe the methodology used to measure growth. For example, does the assessment employ a vertical scale, use a computer-adaptive model to measure growth over time, or employ some other

Requirement:
methodology. Does the methodology allow for the longitudinal measure of growth across academic years? What about the measurement of required growth on the proposed assessment to reach proficient on the statewide assessments (the Standards of Learning tests) in a specified amount of time? Include standard setting studies or other analyses conducted to establish measures of growth.

SAS EVAAS Value-Added Methodology

SAS EVAAS value-added models use multivariate, longitudinal data structures to provide precise and reliable measures of the academic progress of populations of students. This enables them to measure the influence of educational entities, or districts schools and teachers, on student level academic gains. The SAS EVAAS statistical methodology and models are outlined in detail in the SAS white paper on SAS EVAAS Statistical Models (2010)[1] and Chapter 13, "The Tennessee Value-Added Assessment System," in Grading Teachers, Grading Schools (1997).[2]

SAS EVAAS uses two main categories of value-added approaches, comprised of district, school and teacher models. One type is used for tests given in consecutive grades, like Math and Reading tests, and the other is used when a test is given in non-consecutive grades such as Science. Regardless of the modeling approach used, the progress of every student is important. In other words, every student counts, regardless of their achievement level or socioeconomic status or demographic background. In order to provide the most reliable value-added metrics, the estimates are provided by test, grade, subject and year for the students associated with a particular district, school or classroom. However, the less formal diagnostic reports allow users to drill down to various subgroups by achievement level so that educators can assess their success with different kinds of students.

SAS EVAAS Multivariate Response Model (Used for Tests Given in Consecutive Grades):
SAS employs three separate analyses for the multivariate response model (MRM) approach, one each for districts, schools and teachers. The district and school models are essentially the same. They perform well with the large numbers of students that are characteristic of districts and most schools. The teacher model uses a different approach that is more appropriate with the smaller numbers of students typically found in teachers' classrooms. All three models are statistical models known as linear mixed models and can be further described as repeated measures models.

Depending on the LEA's preference, the progress of districts and schools can be compared to the average progress in the nation (for NRTs), to the average progress in the state (for SOL), or to the average progress within a group of progress in the district (for SOL or Interactive Achievement Student Growth

[1]Wright, S. Paul, John T. White, William L. Sanders, June C. Rivers (2010). *SAS white paper on SAS EVAAS Statistical Models* [White Paper]. Retrieved from http://www.sas.com/resources/asset/SAS-EVAAS-Statistical-Models.pdf.

[3] Sanders, William L., Arnold M. Saxton, Sandra P. Horn (1997). The Tennessee Value-Added Assessment System. In *Grading Teachers, Grading Schools,* edited by Jason Millman, 137-162. Thousand Oaks: Corwin Press. Refer to http://www.sas.com/govedu/edu/sanderssaxtonhorn.pdf.

Requirement:

Assessments). Furthermore, in addition to reporting the estimated mean scores and mean gains produced by these models, EVAAS can report (1) cumulative gains across grades (for each subject and year), (2) three-year- average gains (for each subject and grade), and optionally (3) composite gains across subjects.

Teacher Reporting:

The teacher estimates use a more conservative statistical process where each teacher is assumed to be the district average in a specific year, subject and grade until the weight of evidence pulls him/her either above or below that district average. This model could also be implemented such that the teacher was assumed to the average across all districts. Furthermore, the teacher model is a "layered" model, which means that the current and previous teacher effects are incorporated. Each teacher estimate takes into account all the students' testing data over the years, the percentage of instructional responsibility the teacher has for each student, the impact of previous teachers on the current year students, and adjustments for future teachers (meaning, when next year's student scores are obtained, the previous year's teacher estimates can be refined with this additional information). All of this provides a wealth of protection against misclassifying each teacher estimate.

Because teacher effects are treated as random effects in this approach, their estimates are obtained by shrinkage estimation, technically known as best linear unbiased prediction or as empirical Bayesian estimation. This means that *a priori* a teacher is considered to be "average" (with a teacher effect of zero) until there is sufficient student data to indicate otherwise. This method of estimation protects against false positives (teachers incorrectly evaluated as effective) and false negatives (teachers incorrectly evaluated as ineffective), particularly in the case of teachers with few students.

Similar to the district and school reporting, the teacher model provides estimated mean gains as well as (1) cumulative gains across grades (for each subject and year), (2) three-year- average gains (for each subject and grade), and optionally (3) composite gains across subjects.

Advantages of the SAS EVAAS MRM Approach:

There are numerous advantages to the SAS EVAAS MRM value-added modeling, some of which are listed below:

- All students are included in the analyses, even if they have missing test scores. All of each student's testing history is included without imputing any test scores.
- By including <u>all</u> students in the analyses, even those with a sporadic testing history, it provides the most realistic estimate of achievement available for a district or school.
- It minimizes the influence of measurement error by using up to five years of data for an individual student. Analyzing all subjects simultaneously increases the precision of the estimates.
- Because the influence of measurement error is minimized, there is no need to adjust the estimates for socio-economic factors.
- It allows educators to benefit from all tests, even when tests are on differing scales.

Specific to the teacher model, there are distinct advantages of the SAS EVAAS approach:

Requirement:

- Allows for comparisons among teachers at different schools, within an entire district or across the state.
- Allows for the possibility that a teacher may be very effective in one subject and very ineffective in another.
- Accommodates the teaching scenarios where more than one instructor has responsibility for a student's instruction in a particular grade/subject, such as team teaching.
- The EVAAS teacher model for tests, such as the state CRT, allows teacher effects to accumulate over time. That is, how well a student does in the current subject/grade/year depends not only on the current teacher, but also on the accumulated knowledge and skills acquired under previous teachers.
- Using a shrinkage estimation, the model protects against misclassifying teachers as effective or ineffective, particularly in the case of teachers with few students.

Additional information can be found at: http://www.sas.com/resources/asset/SAS-EVAAS-Statistical-Models.pdf

SAS EVAAS Univariate Response Model (Used for Tests Given in Non-Consecutive Grades):
Tests that are not given each year require a different modeling approach. The SAS EVAAS approach that must be used with these tests is called the Univariate Response Model (URM). The URM is, in traditional terminology, an analysis of covariance (ANCOVA) model. In simple terms, this model compares students' predicted scores for a particular subject/year with their observed scores. The model is a function of the difference between these two scores for all their students. The predicted scores is based on each student's prior testing history (all grades, subjects and years), and a student will receive a prediction only if he or she has three prior test scores in any subject or grade to ensure reliability due to the measurement error in the prior testing history being used to make the prediction.

The URM has similar models for district and school analyses, and a slightly different model for teachers that allows multiple teachers to share instructional responsibility.

Advantages of the SAS EVAAS URM Approach:
There are many advantages to the SAS EVAAS URM approach, some of which are listed below:

- It does not require students to have all predictors or the same set of predictors, so long as a student has at least three prior test scores in any grade/subject.
- It minimizes the influence of measurement error by using all available data for an individual student. Analyzing all subjects simultaneously increases the precision of the estimates.
- Because the influence of measurement error is minimized, there is no need to adjust the estimates for socio-economic factors.
- It allows educators to benefit from all tests, even when tests are on differing scales.
- By using shrinkage estimation, it protects the effects from misclassification as effective or ineffective.
- Accommodates the teaching scenarios where more than one instructor has responsibility for a student's instruction in a particular grade/subject, such as team teaching.

Requirement:

Additional information can be found at: http://www.sas.com/resources/asset/SAS-EVAAS-Statistical-Models.pdf

SAS EVAAS Projection Methodology

SAS EVAAS projection methodology complements the value-added reporting tools while using the same student assessment data provided by the VA DOE and/or LEA.

The SAS EVAAS projection methodology[3] provides educators with estimates of individual students' likelihood of reaching future academic targets. Depending on the testing available, these targets could include next-grade proficiency levels, end-of-course tests, requirements for high school graduation, exemption from post-secondary remedial courses, and success in a variety of majors using SAT or ACT benchmarks.

A student receives a projection for a specific grade/subject/test by comparing his or her testing history to those of the most recent cohort of completers for that specific grade/subject/test. The specific modeling approach is very similar as to what was described in detail for the SAS EVAAS URM methodology. Refreshed each year, these individual student projections provide meaningful indicators for student placement as well as measures of resource requirements to meet specific academic goals.

Educators can use these projections as the most reliable metric to identify the appropriate academic experience for each student, regardless of achievement level. Students with a low probability of being proficient on future tests can be identified years in advance so that they can receive targeted interventions before they take the test. Early high achieving students can be encouraged to enroll in appropriate rigorous coursework, like 8th grade Algebra or Advanced Placement classes, to continue their early success and ensure college readiness.

Variations of the SAS EVAAS individual student projections provided herein have been approved as growth model augmentations in Tennessee, Ohio, and Pennsylvania. Additionally, the SAS EVAAS student projection methodology was reviewed by the Government Accounting Office (GAO) in 2006 and is referenced in its July report to the House Education Committee, which is available on the official GAO website at: http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=gao&docid=f:d06661.pdf

It is important to note that the growth/projection methodology is not based on the assumption that every student must have the same set or number of predictors. In those instances where data are missing (e.g., out-of-district transfer students), the missing data will not be imputed. Instead, the projections for any student, regardless of the number of prior test scores, would be obtained by using the appropriate relationships between the student's available testing history and the subject/grade that is being projected

[3] Sanders, William L,. S. Paul Wright, June C. Rivers (2006). Measurement of Academic Growth of Individual Students toward Variable and Meaningful Academic Standards. In *Longitudinal and Value Added Modeling of Student Performance*, edited by Robert W. Lissitz. Maple Grove: JAM Press. Refer to http://www.sas.com/govedu/edu/wrightsandersrivers.pdf.

Requirement:
based on the most recent set of student completers of that subject/grade. This flexible way of handling missing data reduces the possibility that highly mobile populations are not excluded from the growth model projections. Also important to note is that, while the methodology adequately handles missing data, projections are made only for those students who have at least three prior test scores in any grade or subject. This minimum number of predictors is required to sufficiently dampen the effects of measurement error in this type of model. **Advantages of the SAS EVAAS Projection Approach:** The projections derived from the SAS EVAAS methodology offer several distinct statistical and psychometric advantages over simplistic modeling, which is less likely to provide reliable information: ➢ Unlike some proposed growth and projection models, there is no requirement that the tests scores for predictors and responses be vertically linked. In fact, they need not even have to be from the same test company or even in the same subject. The important feature is that the predictors be good predictors of the responses. This provides an enormous amount of flexibility in the choice of what could be projected and which predictors to use in making the projections and avoids many of the psychometric issues raised as criticisms of other growth models. Also, this gives policymakers more flexibility in the future as to its testing options. ➢ Even if vertically linked scales are deployed, there is no assumption required about the overall shape of the growth curve. ➢ Missing values are easily handled so that different students can have different sets of predictors. ➢ This approach allows each student in each district to have individual projections because of the efficiency of the computational strategy deployed. ➢ Each projection takes into account the amount of data that the individual student has when providing a measure of certainty for that projection. Additional information can be found at: http://www.sas.com/resources/asset/SAS-EVAAS-Statistical-Models.pdf

Requirement:
3.3 Describe the methodologies used to control item exposure so that the accuracy of students' scores is not impacted by multiple exposures to the same items.
Interactive Achievement Student Growth Assessments are only delivered and scored via the onTRAC™ instructional improvement system. onTRAC is a secure, web-based application that requires user authentication for access. onTRAC is a hierarchical, permissions-based application that validates user access via multiple levels of credentialing. The Interactive Achievement Student Growth Assessments are

Requirement:
made available only to system administrators in school divisions. System administrators have multiple measures for controlling item exposure so that the accuracy of students' scores is not impacted by multiple exposures to the same items. These measures include: ➢ Designation of "reserved" status at both the item and assessment level from within the system, both of which require explicit permissions for access ➢ Test window designation, which restricts general users from creating test sessions outside of the defined test window ➢ Secure report designation, which restricts general users from viewing the content of test items when accessing reports but does not restrict access to student results and other data ➢ Restriction of students being assigned test sessions for the same assessment more than once

Requirement:
3.4 Describe the procedures used to validate the measures of growth.
Many steps are taken before and after the analysis to validate the measures of growth. First, before the analysis is run, the scaling of the tests is validated given the criteria for use in value-added models that was described above. The requirement of sufficient stretch has to be met to ensure valid and reliable measures of growth can be obtained for students of all achievement levels. The test must also be highly correlated with curricular objectives or else no difference would be seen in the resulting value-added measures. Student level data is checked beforehand to see if outliers exist. If there are clear outliers, then those data are removed from the models. In creating the value-added models, parameters and relationships are checked very carefully. Relationships are checked between scores at the individual student level to ensure this is a sufficient relationship to create these value-added models. This is important for predicting scores using the URM methodology and when measuring from one point in time to the next in the MRM methodology. The different subjects and grades all have some relationship between each other and that is checked prior to creating the value-added measures. Model parameters of variability among districts, schools, and teachers are checked to ensure there is enough variability to see measurable differences among these entities. After models have been built and they look satisfactory, different aspects of the results are checked for quality and assurance, which is also a part of the validation. With the MRM and URM reporting, value-added measures are compared to arithmetic or raw measures of progress for the group of students to ensure everything appears reasonable based on the individual student level data. If anything does look out of the ordinary, those particular measures are examined more thoroughly to understand exactly why there are larger differences. Next, last year's measures for each school or teacher are compared to this year's measures to see if there

Requirement:
are any drastic changes. If there are significant or large changes, those particular value-added measures are examined to determine if the student level data support the difference. Finally, the value-added measures are reported through the web-based application back to users of the system. Before this is delivered, these measures are checked again by a Quality Assurance team to ensure the right measures are used the reporting that is seen by the educators.

Section 4: Test Administration Procedures

Requirement:
4.1 Describe the administration procedures necessary to produce growth scores. For example, is the assessment designed to be administered multiple times during the year or administered once in the fall and once in the spring?
Interactive Achievement Student Growth Assessments are designed to be administered once in the fall and once in the spring. To ensure maximum reliability and to ensure the assessments meet the stringent criteria of SAS EVAAS reporting, the assessments are delivered under prescriptive directives. These directives include measures involving assessment security and administration windows. Interactive Achievement and SAS propose that the VA SOL tests and Interactive Achievement Student Growth Assessments will serve as inputs into the SAS EVAAS modeling. Thus the new Interactive Achievement tests will be administered during a set window at the end of the school year. The measure of growth for these tests will show how much progress students made from the beginning of the year to the end of the year in that assessment area. The windows of administering the test will ensure a valid and reliable growth measure as well. Prior to using assessment data with EVAAS, SAS works with LEAs to determine if their assessments meet three characteristics to determine if they are appropriate to use in a longitudinally linked analysis. It is also important to be sure that the appropriate value-added model does not make assumptions, which are not met by these scales. Typically, scales must meet the below three requirements to be used in most types of analysis: 1. Scales must have sufficient stretch to show measurable differences among students at the very top and bottom of the distributions. 2. Scales must be reliable from year to year. 3. The test must be highly correlated with the curricular objectives.

Requirement:
SAS checks the first two requirements every year using the distribution of scale scores along with the sample of student test data that is sent each year before the full test data is given. SAS will aid each LEA in checking these distributions and provide reporting back to the LEA around the distribution of scale scores each year. These checks are vital to providing a valid and reliable value-added analysis. When necessary, SAS offers options to alleviate the concerns on the first requirement. In general, some approaches exist that take out the differences of scales from year to year. As you will see below, some value-added models, such as predictive-type models do not depend on test scaling. In fact different tests can be used in different years with different scales to provide the most comprehensive use of data. If the distributions do not meet the second requirement, then SAS can alter models to eliminate any effect of year to year change in test scales. Most state and district administrators design/select tests that assess what students are expected to learn during their standard course of study; therefore, the third requirement is typically met in practice. However, if the third requirement is not met, then all districts, principals, and/or teachers would not be measurably different in terms of value-added measures. SAS monitors this correlation annually. EVAAS follows the progress of each student over time and uses all of the available assessment data so that educators are not held accountable for any single test score. Incorporating more test data will dampen measurement error. SAS then uses the modeling approaches described in Section 3.2 to provide a gain/growth effect of the district, school, and teacher's impact on students' progress.

Requirement:
4.2 Describe any processes used for pre-identifying and/or registering students for testing. Include what data, including the State Testing Identifier, are collected for each student, how data are collected or transmitted, and how data are maintained and securely managed.
All students registered in a K – EOC core-content course are pre-identified for testing. Students are registered for testing through the onTRAC™ application by course enrollment and the creation of a test session. onTRAC™ requires that all student records include the unique State Testing Identifier. Student data can be collected in one of two ways for each student. First, the student data can be collected online via IA's online testing interface iTest, which transmits all student responses via a secure SSL connection. Second, student data can be collected from student answer sheets that are scanned via IA's secure scanning interface, iScan. In both instances data is securely collected, transmitted and managed.

Requirement:
4.3 Describe all materials needed for test administration and how school divisions will order and obtain sufficient quantities. Include details of test booklets and answer documents for paper/pencil testing (if applicable), test administration manuals, etc. If applicable, identify any test administration materials school divisions would be responsible for supplying locally (manipulatives, copies of test materials, etc).
Materials needed for test administration are dependent upon each division's/school's desired delivery method. If divisions/schools prefer online delivery, no specific materials are required for administration. All test instructions and assessment manipulatives are made available online through the onTRAC™ application. If divisions/schools prefer paper/pencil delivery, student answer sheets and scanners are required for collection of data. In the case of paper/pencil delivery, divisions/schools would be responsible for supplying their own student answer sheets and scanners, both of which can be purchased through IA or other 3rd party vendors (i.e. Scantron).

Requirement:
4.4 Provide examples of the test administration manuals to be used with the assessment(s).
Interactive Achievement Student Growth Assessments do not require test administration manuals to be used with the assessments.

Requirement:
4.5 Describe all technology requirements related to school personnel managing the administration of tests and to students completing tests if assessments include technology-based delivery. Include the minimum and recommended hardware and software requirements and network requirements for test administration by school personnel and test delivery to students. Include how assessments are hosted (e.g., locally, vendor, 3rd party). Provide examples of user interfaces for test administration by school personnel and test delivery to students. Include descriptions or examples of test navigation and any test tools (e.g., calculator, ruler, highlighter) available to students for testing.
Online delivery of the Interactive Achievement Student Growth Assessments requires the following minimum and/or recommended hardware, software and network requirements:

Requirement:

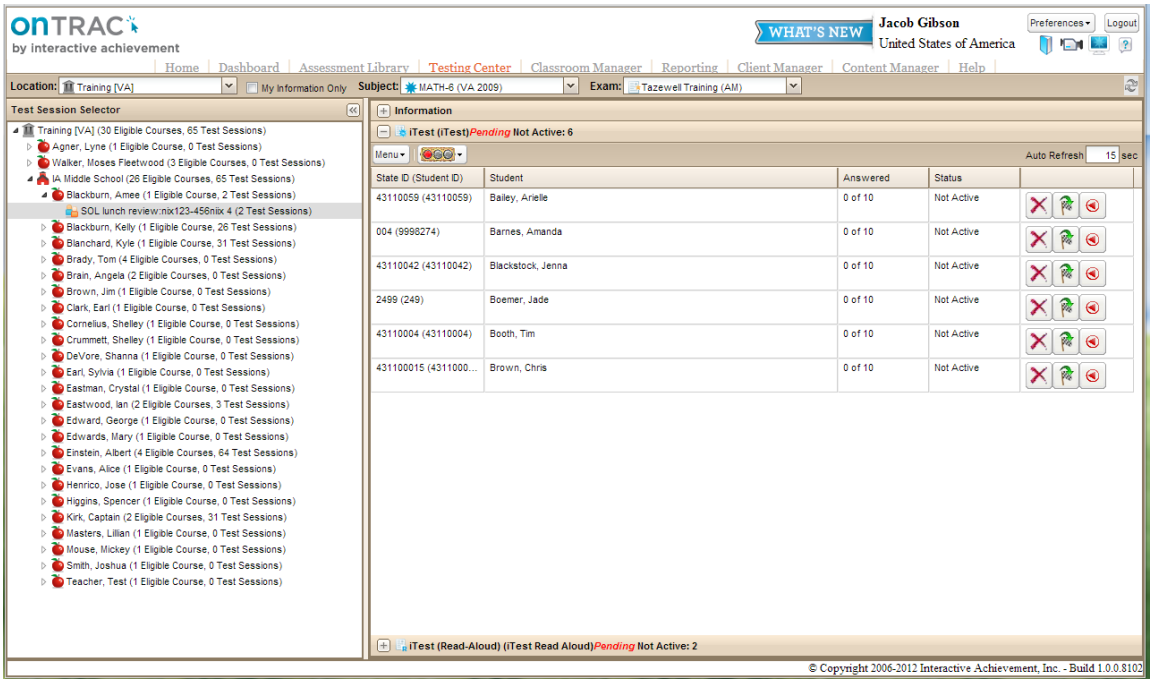
Web Hosted Services Require…

- Mozilla Firefox™ or comparable web browser
- High-speed internet service

iTest Testing Software Require…

- Windows XP, Windows 2000, Windows Vista, or Mac OS X or higher
- Flash v8 or higher
- iTest software installation
- High-speed internet service

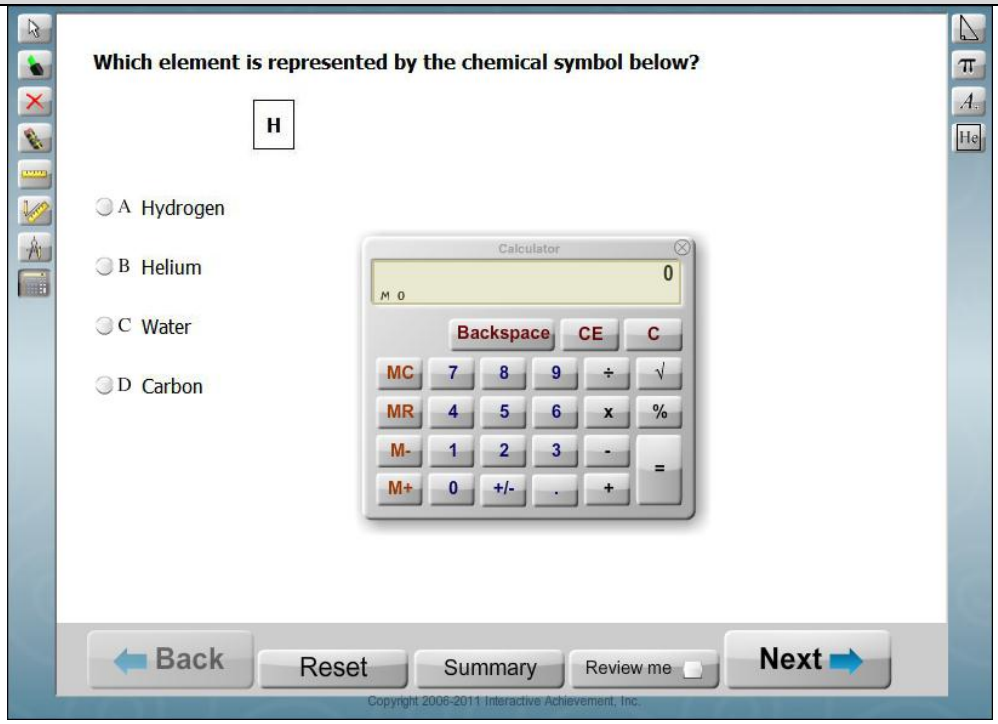
onTRAC™ operates as a software as a service (SaaS) model and all assessments are hosted by Interactive Achievement. The following screen capture represents a sample user interface for test administration (test session creation and monitoring) by school personnel.



School personnel have the ability to create test sessions for course sections; assign students to paper/pencil, online, or read-aloud test sessions; grant permission for student access to assessments; monitor student progress for online administrations; and reset test sessions.

onTRAC delivers assessments online utilizing Interactive Achievement's online student test-taking interface: iTest. iTest provides students with a cutting edge online testing experience that includes enhanced technology and tools (see below) while mirroring the Virginia SOL testing experience. Available tools include, but are not limited to: ruler, highlighter, eraser, protractor, compass, straight edge, and calculator.

Requirement:



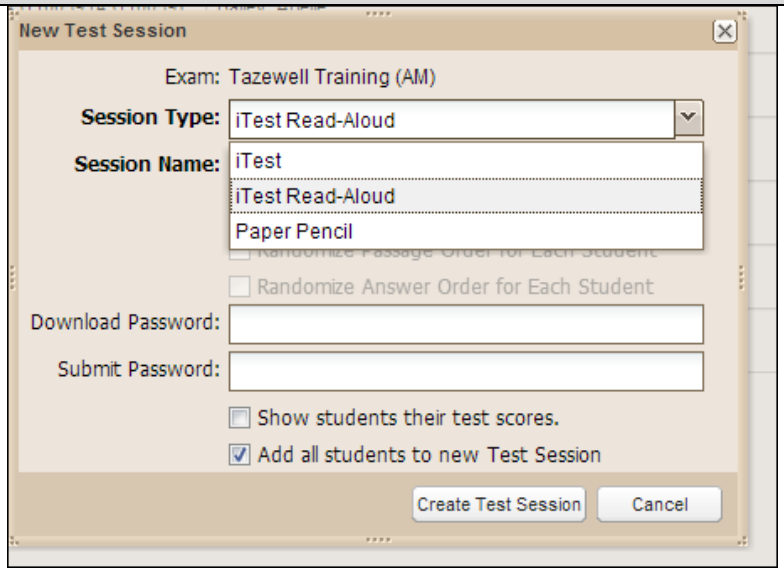
Requirement:

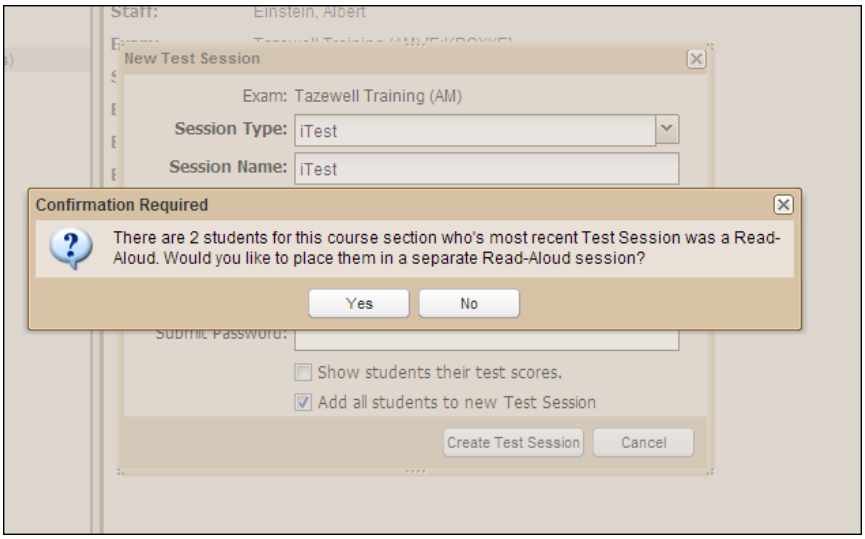
4.6
Describe accommodations available to students with disabilities and limited English proficient students. Include procedures related to the provision of accommodations to eligible students.

Accommodations are available to students with disabilities and limited English proficient students. These accommodations include:

- All paper/pencil administrations of assessments may be enlarged as needed according to the students' IEP.
- All online administrations may be scheduled as a read-aloud test session for accommodating students with disabilities and limited English proficient students (the following screen captures demonstrate the procedures related to the provision of accommodations to eligible students).

Requirement:





Requirement:

4.7
Describe procedures for completed student tests to be submitted for scoring and reporting purposes.

Student tests are submitted for scoring and reporting in one of two ways, depending on the method of delivery. If students are testing online then tests are submitted for scoring via a secure SSL connection and are scored and available online for reporting in real-time. If students are testing paper/pencil then tests are submitted for scoring and reporting by scanning student answer sheets through a scanner that is

Requirement:
connected to a computer with a secure SSL connection to the onTRAC™ application. Student test results are available online for reporting in real-time the moment they are scanned.

Section 5 : Scoring and Reporting

Requirement:
5.1 Describe scoring procedures for all item types and test forms administered, including implemented quality control measures.
Interactive Achievement Student Growth Assessments include selected response item types only. The submission procedures for these items types, whether collected paper/pencil or online, are outlined in section 4.7 above. The scoring procedures for all selected response items are 100% objective based on a correct vs. incorrect identification. No selected response items include rubrics or allow human scoring. All assessment items have a 1.0 weighting and are machine scored and therefore ensure concrete quality control measures for scoring all items.

Requirement:
5.2 Describe the type of reporting provided (e.g, static and/or dynamic, electronic and/or paper-based, item-level, strand-level, and/or test-level scoring). Include approximate timelines for score reports to be available to divisions, how score reports will be accessed and/or obtained, and samples of student, class, school, and division score reports and sample record layouts for electronic data files.
Interactive Achievement Student Growth Assessments include separate reporting interfaces. One interface provided by Interactive Achievement and one interface provided by SAS EVAAS. Interactive Achievement's reporting interface includes: • Static and dynamic reports • Online, electronic reports that are interactive as well as printable • Reporting at the division, school, teacher, class, and student levels • Reporting by AYP subgroup • Reporting by test level, reporting category (strand), and item level • Reporting by Bloom's Taxonomy levels • Item analysis reporting at the division, school, teacher, class and student levels

Requirement:

- Student performance by question, standard and Bloom's taxonomy levels

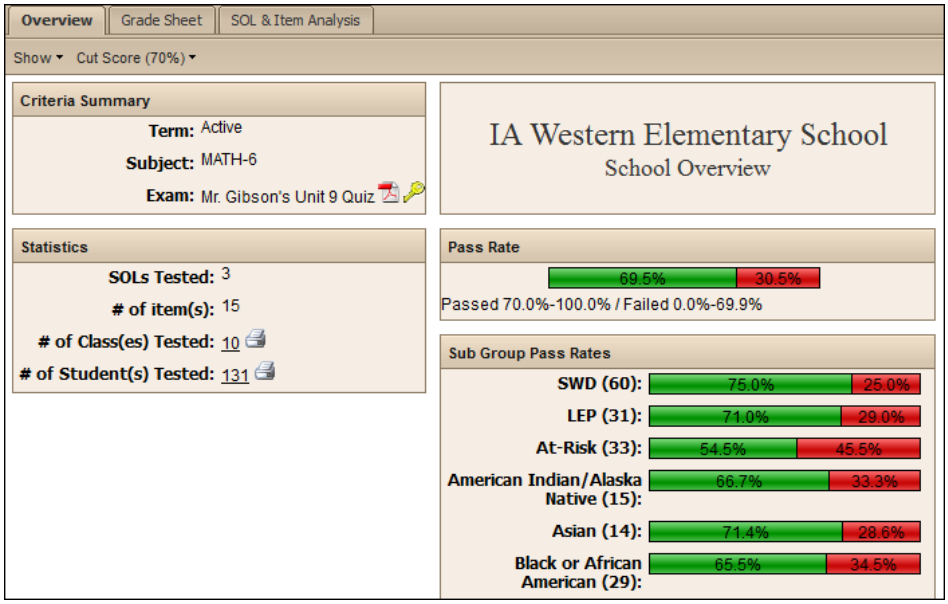
Interactive Achievement Reporting

The onTRAC™ Instructional Improvement System is designed by Virginia educators to provide precise reports for data analysis. All onTRAC reports provide detailed information appropriate for administrators, teachers, students, and parents. onTRAC™ reporting facilitates data-informed decisions and instruction at all levels, promoting a community based approach to increasing student achievement.

Report Bank (Pre-defined Reports)

Overview Report

The Overview Report in onTRAC™ provides a high level, quick-look view for assessment results at all levels (district, school, teacher, class, and student). The report also provides at-a-glance information regarding performance by NCLB AYP subgroups and SOL.



Grade Sheet Report

The Grade Sheet Report provides a question-by-question look at student performance as it relates to a single assessment administration. Administrators and teachers alike may analyze results and pinpoint areas of need for an entire class or individual students. This interactive report allows for disaggregation by SOL, Bloom's Taxonomy, Class and Student. Results of the Grade Sheet Report facilitate a data driven instruction model of data analysis.

Requirement:

Overview | **Grade Sheet** | SOL & Item Analysis

Show

Student	SCORE	Q:897 #1	Q:867 #2	Q:894 #3	Q:877 #4	Q:868 #5	Q:866 #6	Q:874 #7	Q:882 #8	Q:875 #9	Q:872 #10	Q:870 #11	Q:869 #12	Q:876 #13	Q:14216 #14	Q:880 #15
Class Totals	77.5%	62.5%	62.5%	87.5%	87.5%	100.0%	87.5%	100.0%	50.0%	100.0%	87.5%	87.5%	100.0%	50.0%	37.5%	62.5%
Armstrong, Louis	80.0%	√C	√H	√C	J[1]	√B	√F	√A	√G	√C	G[3]	√B	√J	√C	√F	D[1]
Hale, Billy A.	93.3%	√C	√H	D[3]	√G	√B	√F	√A	√G	√C	√J	√B	√J	√C	√F	√C
Kent, Clark	80.0%	√C	√H	√C	√G	√B	√F	√A	J[3]	√C	√J	√B	√J	A[2]	G[3]	√C
Martin, Miguel	60.0%	√C	F[2]	√C	√G	√B	G[1]	√A	H[2]	√C	√J	C[3]	√J	√C	G[3]	--
Mary, Rose	73.3%	A[1]	√H	√C	√G	√B	√F	√A	J[3]	√C	√J	√B	√J	A[2]	H[1]	√C
Moore, Mary T.	66.7%	A[1]	J[3]	√C	√G	√B	√F	√A	√G	√C	√J	√B	√J	A[2]	J[2]	D[1]
Strong, Tom	66.7%	A[1]	J[3]	√C	√G	√B	√F	√A	J[3]	√C	√J	√B	√J	A[2]	G[3]	√C
Tyree, Keisha	100.0%	√C	√H	√C	√G	√B	√F	√A	√G	√C	√J	√B	√J	√C	√F	√C

Item Analysis

The Item Analysis report in onTRAC™ is an investigative report that allows for easy aggregation and/or disaggregation of student results for the purpose of identifying areas for improvement. Administrators and teachers can “drill down” to the item level to display distracter analysis. Distracter analysis allows educators to identify areas for improvement while gaining a better understanding of student responses on an aggregate basis. Trends in student performance that are identified by onTRAC’s Item Analysis Report may be addressed quickly and efficiently so that student mastery may be positively affected.

Overview | Grade Sheet | **SOL & Item Analysis**

View By ▾ Expand To ▾

Exam	Average
Mr. Gibson's Unit 9 Quiz (15 Questions)	77.5%
Category	**Average**
[MATH-6.9] Measurement (15 Questions)	77.5%
Sub-Category	**Average**
[MATH-6.9a] Length (7 Questions)	75.0%

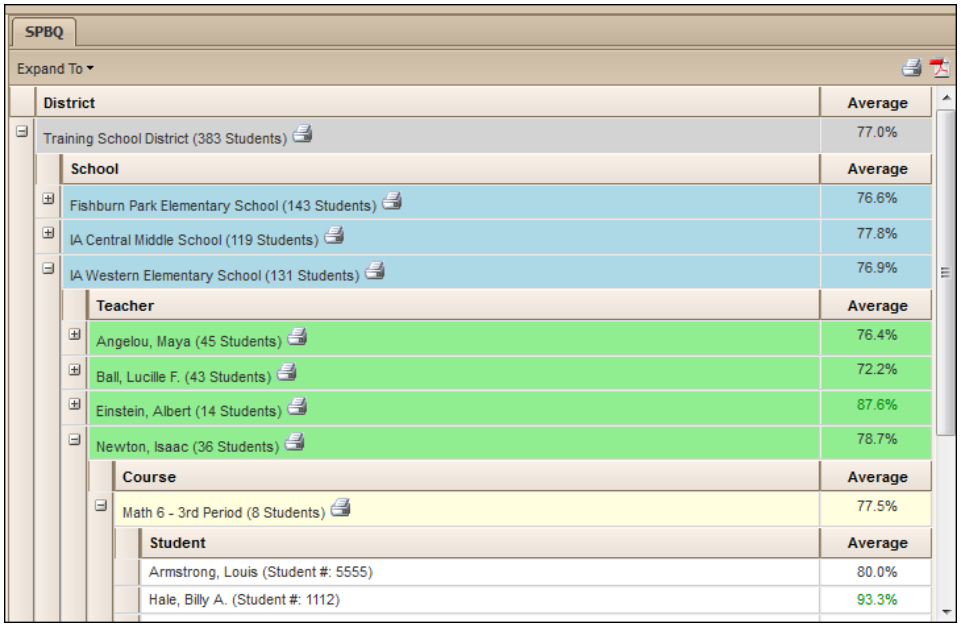
Question	√ Correct	× Dist. 1	× Dist. 2	× Dist. 3	Not Ans.
#1 Q:897	62.5%	37.5%	0.0%	0.0%	0.0%
#3 Q:894	87.5%	0.0%	0.0%	12.5%	0.0%
#4 Q:877	87.5%	12.5%	0.0%	0.0%	0.0%
#6 Q:866	87.5%	12.5%	0.0%	0.0%	0.0%
#8 Q:882	50.0%	0.0%	12.5%	37.5%	0.0%
#10 Q:872	87.5%	0.0%	0.0%	12.5%	0.0%
#15 Q:880	62.5%	25.0%	0.0%	0.0%	12.5%

[MATH-6.9b] Weight/Mass (4 Questions)	90.6%
[MATH-6.9c] Liquid Volume (4 Questions)	68.8%

Requirement:

Student Performance by Question (SPBQ)

The Student Performance by Question (SPBQ) report in onTRAC identifies student performance by standard, Bloom's Taxonomy, and question. By analyzing the results of onTRAC's SBPQ report, educators target areas of need at the district, school, teacher, class and student levels.



SAS EVAAS Reporting

SAS has a wealth of experience providing reporting across groups of districts and schools, each with different resources, goals, staff, and student populations. To ensure that each educational entity receives useful and insightful reporting, SAS provides over 30 reports with interactive features and drill down capability so that users can customize queries to meet their specific educational goals. With this SAS-hosted solution, LEAs can gain significant savings through quick deployment, low overhead and minimal hardware and staffing requirements. Data is sent to SAS to be extracted, transformed and loaded into the system. Reporting tools are then delivered through a secure Web browser in as little as four to six weeks upon receipt of clean and usable data. The following is a summary of some of the main report types that educators will be provided with SAS EVAAS for K-12:

- **Value-Added metric**- reflectively identifies strengths and opportunities for improvement, grade to grade and subject to subject, for districts, schools and teachers.
- **Individual student-level projections**- proactively predict student success probabilities at numerous academic milestones including upcoming grade level proficiency, high school graduation requirements, and various college success indicators, such as advanced placement exams, the PSAT, SAT and the ACT.
- **Custom Scatterplots**- interactive scatterplots allow users to visualize the impact of various progress metrics and the students who are served. For example, student

Requirement:
achievement or growth can be compared to a variety of socioeconomic and demographic variables to ensure that all subgroups experience optimum educational outcomes. This powerful feature allows users to select the x and y axes variables to plot, drill down to detailed performance information at the district, school, and teacher level, and zoom into any area within the scatterplot for further analysis. • **Diagnostic reports**- identify which students are at risk for underachievement by viewing the incoming levels of academic preparedness of students and the growth patterns within classrooms. • **Customized reports-** flexibility to display information in a variety of forms: District & School Dashboards, College Readiness Dashboards, and Custom Student List Reports. • **A complete sample reporting document is included in Appendix B** Upon release of the web-based reporting, educators can learn how to fully utilize the reporting for continuous instructional improvement in a variety of ways. As a hosted solution, each report has a customized help page for instantaneous support in interpreting and applying the information. SAS also provides self-paced video learning modules and instructor-led training for additional professional development.

Requirement:
5.3 Describe all data tools available to school division staff for the analysis of data and the creation of customized reports.
Refer to 5.2 and the sample reports in Appendix B.