PVAAS PENNSYLVANIA

PVAAS Methodologies

Measuring Growth & Projecting Achievement

Last Updated December 2023.

pennsylvania DEPARTMENT OF EDUCATION

education.pa.gov

Commonwealth of Pennsylvania
Josh Shapiro, Governor

Department of Education
Dr. Khalid N. Mumin, Secretary

Office of Elementary and Secondary Education
Dr. Carrie Rowe, Deputy Secretary

Bureau of Curriculum, Assessment, and Instruction
Brian Campbell, Director

Division of Assessment and Accountability
Brian Truesdale, Chief

For Inquiries Concerning Nondiscrimination in Employment:
Pennsylvania Department of Education
Equal Employment Opportunity Representative Bureau of Human Resources
Voice Telephone: (717) 783-5446

For Inquiries Concerning Nondiscrimination in All Other Pennsylvania Department of Education Programs and Activities:
Pennsylvania Department of Education | School Services Unit Director
333 Market Street, 5th Floor, Harrisburg, PA 17126-0333
Voice Telephone: (717) 783-3750, Fax: (717) 783-6802

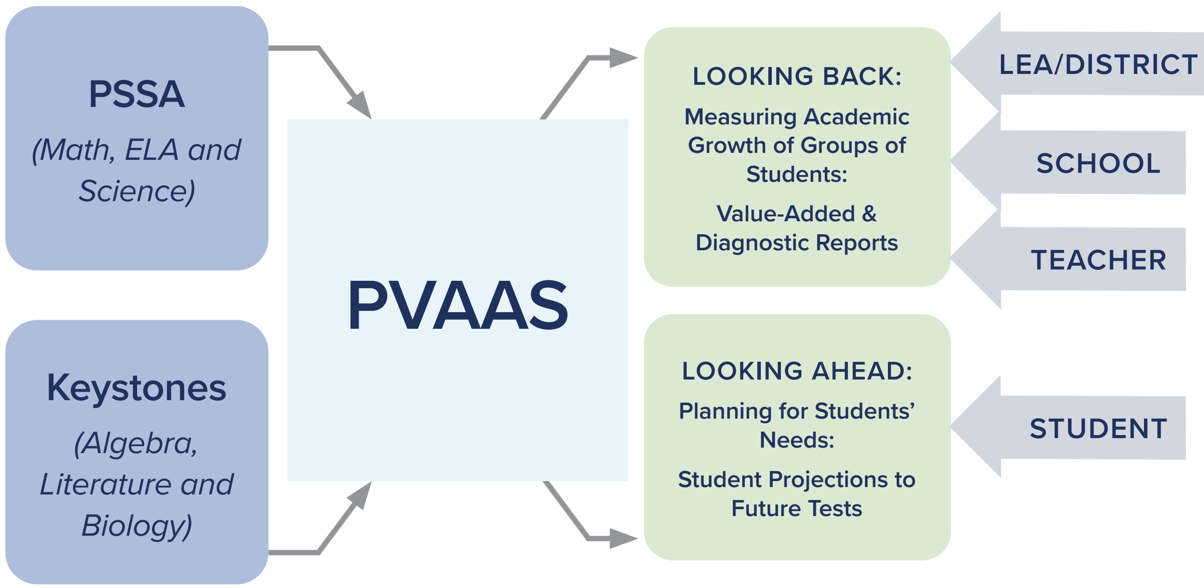
If you have any questions about this publication or for additional copies, contact:
Pennsylvania Department of Education | Bureau of Curriculum, Assessment, and Instruction
333 Market Street, 5th Floor, Harrisburg, PA 17126-0333
Voice: (717) 787-8913 | education.pa.gov

All Media Requests/Inquiries: Contact the Office of Press & Communications at (717) 783-9802.

Contents

What Does PVAAS Provide to Educators?

PVAAS provides a wealth of information to educators through a variety of reports within a web-based system. This information can be divided into two main categories – measuring growth and projecting achievement.



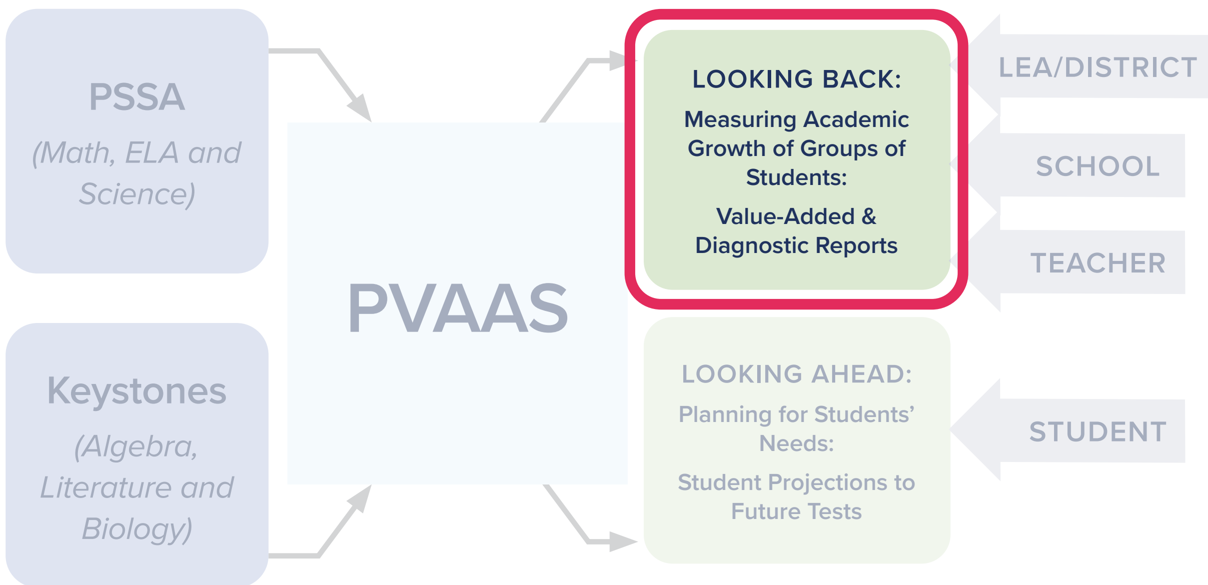
Note: All prior data are used.

First, PVAAS provides information on the academic growth students have made in the most recently tested school year. This is what we call "looking back" information as it helps districts, schools, and teachers assess how their educational programs and instructional strategies are impacting the academic growth, or progress, of groups of students. This "looking back" information is about the growth of groups of students, NOT individual students, and is available at the LEA/district, school, and teacher levels. In PVAAS, there are two ways of measuring growth depending on whether Pennsylvania's state assessments (PSSA and Keystones) are administered in consecutive years in the same subject. Both ways of measuring growth are discussed in detail in later sections.

Second, PVAAS provides information on students' possible academic performance, or achievement, on future assessments (PSSA, Keystones, Advanced Placement, ACT, ACCESS for ELLs, PSAT, and SAT) – called PVAAS Student Projections. This is what we call "looking ahead" information as it helps districts, schools, and teachers plan appropriately for the needs of its students – including decisions regarding intervention placement, enrichment opportunities, course selection, and differentiated instruction. This "looking ahead" information is available for educators at an individual student level, or for grade-level groups and demographic student groups. Details for projecting student achievement will be discussed in a later section.

Concept of Growth:
What is meant by academic growth?

When talking about the academic growth for groups of students, it is important to first understand what is meant conceptually about growth versus achievement before getting into the more specific details of how growth is measured in PVAAS.

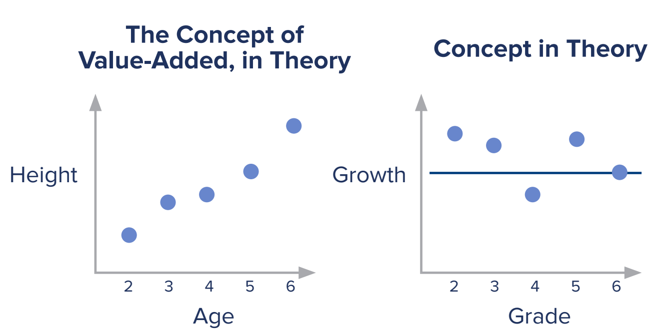


Note: All prior data are used.

While PVAAS does not provide measures of growth for individual students, it is helpful to understand the concept of academic growth by imagining a child's physical growth curve. Every year, a child stands up against the wall; the parent puts a ruler on his/her head and measures the child's height at ages 2, 3, 4, and so on.

From these data points, the parent can construct a graph to illustrate the height of the child and then use these points to construct a graph of the growth of the child, as you see at bottom left. Often, parents discover that this growth curve does not resemble the smooth line seen on a pediatrician's chart. Instead, there are "dimples" and "bubbles" in this constructed graph. Children may have growth spurts. In addition, errors of measurement are possible; the child may not have stood straight, or the parent did not hold the ruler 100% level.

Now apply the same process to education. The graph shown below measures growth for the same group of students for each grade. Imagine that a school has been testing students' annually in math and that the scores from these tests are used to construct each student group's math growth curve. The curve for any student will likely exhibit a pattern of dimples and bubbles similar to the physical growth curve seen for an individual child. However, if by aggregating the information on many students, we discover a dimple effect occurring in 4th grade math with a group of students, then the dimple is evidence that the "standards-aligned system" for 4th grade math may need to be examined.



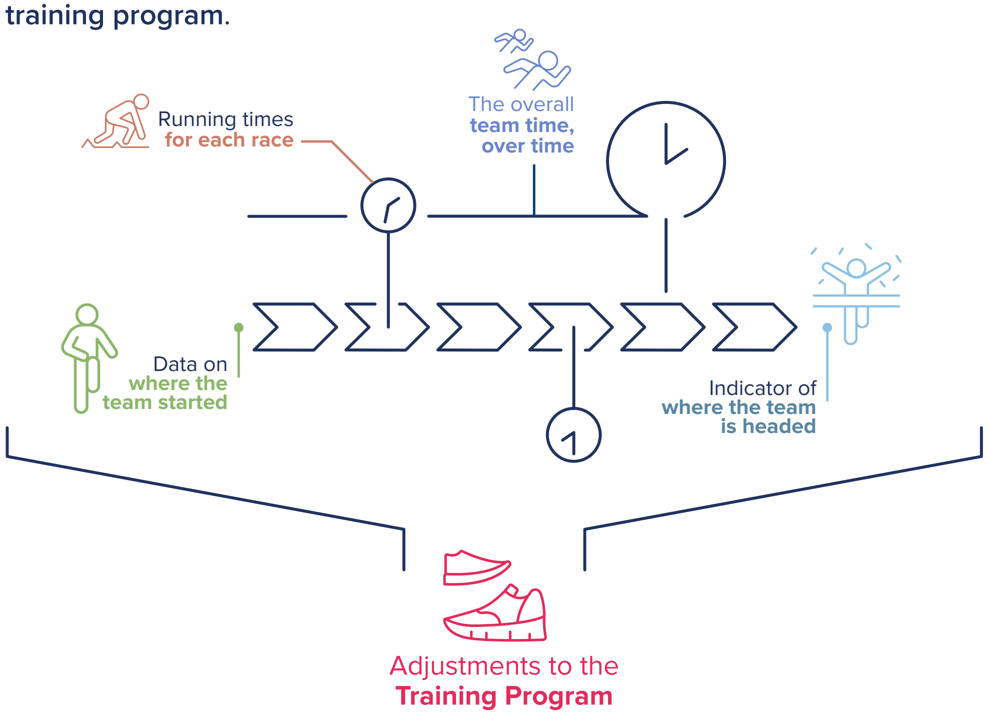
A Growth Analogy

Consider another analogy, measuring the athletic progress or growth of track relay teams. Let's consider two relay teams. We'll call them Team A and Team B. Each team is comprised of individual runners whose individual times contribute to the overall speed of the team. In terms of achievement:

1. Team A has never placed in the top half of all teams at a track meet.
2. Team B, however, has always placed in one of the top two spots.

Obviously, given their achievement level, Team A's coach wants his team to grow and improve. But Team B's coach also wants his team to grow and improve. So, how can both coaches track their team's progress or growth over time?

First, they'd need to know **their team's data** – that is, their running times – **for each race**. On any given day, an individual runner's time, and therefore the **overall team's time**, can vary. For an accurate picture of the impact of their techniques and training regimen, the coaches would need to use the times from all races over a period of time to know **where the team started** and **how the team has progressed**. The ongoing data can be used to **adjust the team's training program**.



A reasonable goal for many coaches would be to at least have their team maintain their speed over time. Regardless of their starting speed, no coach would want their team to lose ground or decrease their speed. Ideally, all coaches would be happy if their team was able to increase their speed. Perhaps the team can even make enough progress to break their team record. So, the coaches of both Team A – whose achievement is not where the coach wants it to be, and Team B – whose coach is proud of their current achievement, can assess whether their team has been maintaining its speed over time, increasing its speed over time, or losing speed over time.

The coaches need to track both achievement – how fast the team is, AND growth – how much the team has progressed, to know whether the team is improving, maintaining, or falling behind. This holds true whether the team is lower-achieving, or higher-achieving.

Connecting this analogy of track relay teams to how academic growth is measured for groups of students; we recognize that there are:

1. teachers, schools, and LEAs/districts in Pennsylvania whose students are more like relay Team B – their students tend to score in the higher ranges on the state assessments,

AND

2. teachers, schools, and LEAs/districts in Pennsylvania whose students are more like relay Team A – their students tend to not be where they need to be academically.

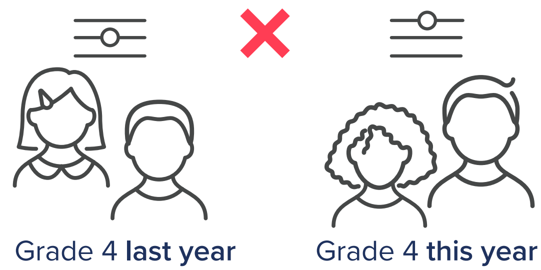
Just like the track relay teams, no teacher, school, or district wants their students to lose ground or show a decrease in their achievement.

Measuring Growth with PVAAS

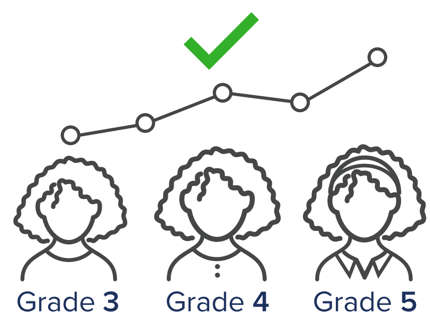
Each year across Pennsylvania, the academic performance of students is evaluated using Pennsylvania's state assessments. Each year, districts and schools receive information regarding the achievement results for their students in the grades and subjects in which they were assessed. This information includes the number and percentage of students who performed in each of the four academic performance ranges – Advanced, Proficient, Basic, and Below Basic. Achievement data from previous years is also included. But, because these data are based on different groups of students each year, it's sometimes challenging for teachers, schools, and districts to use the data to make fair comparisons. This is the purpose of PVAAS – to help educators make more fair comparisons.

PVAAS uses the state assessment data to measure the academic growth of groups of students from year to year at the district, school, and teacher levels. As illustrated to the right, PVAAS does NOT compare the achievement of one group of students in a particular subject and grade with the previous year's students in that same subject and grade.

For example, the achievement of the most recent group of students in 8th grade Math is NOT compared to the achievement of the previous year's group of 8th grade Math students.



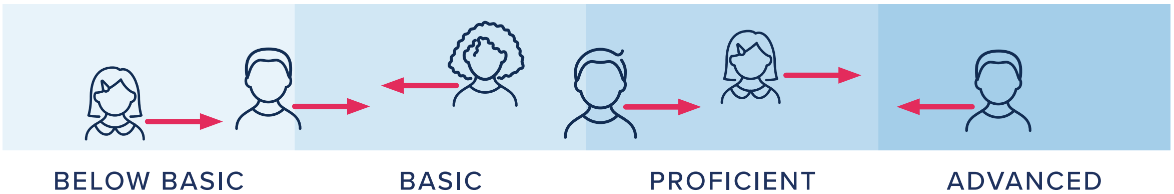
PVAAS does not compare the achievement of one group of students with the previous year's students in that same subject or grade.



PVAAS measures the growth of each distinct group of students as they progress through grade levels over time.

Rather, PVAAS measures the growth of each distinct group of students as they progress through grade levels over time. It compares the group's most recent academic achievement to their own prior academic achievement to assess their growth over time.

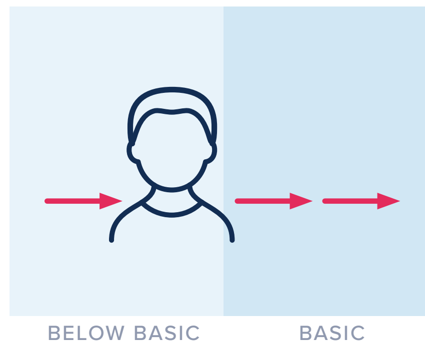
Additionally, PVAAS does NOT use the percentages of students at various academic performance levels to measure growth. Each performance level contains a range of scores, and students move around within these ranges, as well as between ranges.



PVAAS uses students' scores, rather than their academic performance level across grades and subjects to generate a reliable estimate of the true achievement level of a group of students. Then, these estimates of achievement are compared to estimate growth for a group of students.

Just as track coaches can track their team's progress, with PVAAS, educators can track what's happening with groups of students as they move through the educational system. PVAAS color-coding within the reporting lets educators easily see if the average achievement of a group increases, decreases, or remains approximately the same over time.

In Pennsylvania, growth is assessed against **the growth standard**, which is based on the philosophy that, regardless of the entering achievement of a group of students, they should not lose ground academically.



No matter the achievement level at which students enter, they should not lose ground academically.

YELLOW

RED

If educators see that the PVAAS growth measure for a group of students is color-coded yellow or red, it indicates that there is moderate or significant evidence that the group of students did not meet the growth standard (i.e., the group lost ground academically).

This should be cause for concern for educators; they would want to ask some questions about curriculum alignment, effective instructional practices, ongoing formative assessment, placement of students in appropriate courses, and the existence of appropriate academic opportunities for struggling students or students who may need enrichment.

GREEN

If educators see that the growth measure for a group of students is color-coded green, it indicates that the group of students met the growth standard – or, on average, the achievement of the group was maintained.

- In this case, educators would want to determine if green is good enough for that group of students.
- If the achievement of the group is high, then many teachers, schools, and districts may say "that green is good." However, even with a green, there are certainly

opportunities for students to increase their average achievement and for educators to support students in making academic growth.

- If the group of students is lower achieving, some educators might say that it's good that the group did not slip further behind. However, most educators would agree that green is not sufficient, or good enough, for a lower achieving group of students since this means that the group would simply be maintaining a lower level of achievement. For students with lower achievement, the goal of teachers, schools, and districts should be to raise the achievement of the student group.

LIGHT BLUE

DARK BLUE

If educators see that the growth measure for a group of students is color-coded with dark blue or light blue, it indicates that there is moderate or significant evidence that the group of students exceeded the growth standard (i.e., the group gained ground academically).

Even higher achieving groups of students can earn a light blue or dark blue. Remember, the achievement is an average for the group, so not every student in the group is achieving the highest possible score every year. Because of this, teachers, schools, and districts in Pennsylvania with high achieving groups of students can also gain ground. Teachers, schools, and LEAs/districts need both achievement data AND growth data to get the complete picture of student learning.

Teachers, schools, and LEAs/districts need **both achievement data AND growth data** to get the complete picture of student learning.

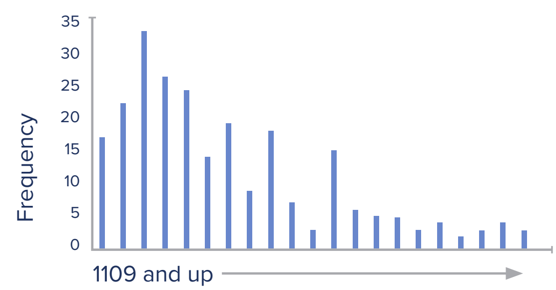
Measuring Growth with High-Achieving Groups of Students

It is important to remember that ANY group of students can meet or exceed the growth standard as measured by PVAAS value-added analyses.

- In PVAAS grades 4-8 Math and ELA reporting, meeting the growth standard is about maintaining students' achievement based on a specific group's prior academic performance.
- In Science and Keystone content areas, meeting the growth standard is about meeting predicted performance based on a specific group's prior academic performance.

Meeting the growth standard in PVAAS does NOT mean increasing students' achievement or increasing students' academic performance levels; rather it is about, at a minimum, maintaining achievement. For lower-achieving groups of students, meeting the growth standard (green in PVAAS) may not be sufficient or acceptable for students to meet long-term achievement goals of proficiency. However, for higher-achieving groups of students, meeting the growth standard (green in PVAAS) may be sufficient or acceptable.

Generally, there is a misconception that teachers, schools, and LEAs/districts with higher-achieving students cannot make academic growth or that it is harder for them to make academic growth. First, we need to consider how high achievement is defined. Some educators may use the total percentage of students who scored Proficient or higher; others may use the percentage of students scoring Advanced; and others may use the average achievement of all students.



Consider an uncommon example where 100% of a school's grade 7 students are performing at the Advanced level in Math. Remember, the Advanced performance level is a range of scores, and this range varies for each subject area and grade level. Using this hypothetical example, the range of possible scores in the Advanced range is 1109 and above in any given year, with many possible scores in between. So where do the students in our example school fall in the distribution of scores at the Advanced level? The graphic shown at right provides a look at the distribution of scores in the Advanced range for our sample school. As you can see, the distribution of scores is skewed towards the lower end of the Advanced range. There is certainly room for students, on average, to maintain or increase their achievement. Furthermore, even though all students are performing in the Advanced level, most students are performing at the lower end of the Advanced range; only a few students are performing at the higher end of the Advanced range.

When a school has a high percentage of students who are reaching proficiency, or even scoring at the Advanced level, we cannot necessarily assume that all students are scoring at the highest point within the Advanced range. In fact, of the approximately 720,000 students assessed on the PSSA in 2023:

- less than 0.09% (less than 690 students) scored at the highest point of the Advanced range in at least one subject;
- less than 0.01% (less than 30 students) scored at the highest point of the Advanced range in Math two years in a row; and
- 3 students scored at the highest point of the Advanced range in ELA two years in a row.

For the Keystone assessments that were administered in the 2022-2023 school year, approximately 158,000 to 215,000 test scores were included in the PVAAS model for each subject. Of these test scores:

- less than 0.01% (less than 10 students) scored at the highest point of the Advanced range in Algebra I;
- less than 0.01% (less than 10 students) scored at the highest point of the Advanced range in Biology; and
- less than 0.01% (less than 10 students) scored at the highest point of the Advanced range in Literature.

It is important to keep in mind that both the PSSA and Keystone exams meet the three requirements to be used in PVAAS value-added analyses.

1. They demonstrate sufficient stretch so that students with low and high achievement can both show growth.
2. They are aligned to state curriculum standards.
3. The scales are reliable from year to year.

Pennsylvania's state assessments do provide sufficient stretch to discriminate not only between proficiency and non-proficiency, but also to meaningfully discriminate between the academic performance levels of Below Basic, Basic, Proficient, and Advanced. Remember, regardless of the entering achievement of a group of students, they should NOT lose ground academically! A reasonable goal would be to at least maintain the achievement of the student group. Any group of students, even groups of students who are higher achieving, can make academic growth!

Any group of students, even groups of students who are higher achieving, **can make academic growth!**

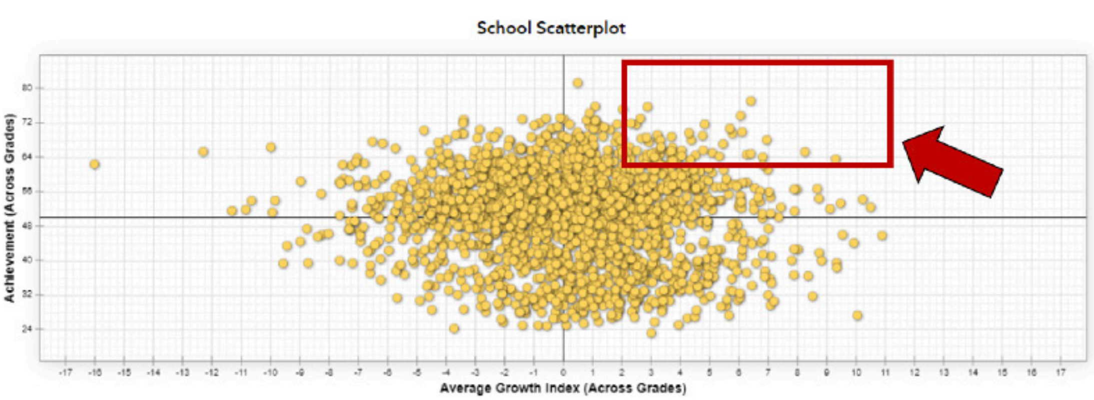
Statewide Evidence

There are groups of students with high achievement across Pennsylvania making high growth each year. The PVAAS scatterplots provide evidence of this. In these scatterplots, each dot represents a real Pennsylvania school, which has been placed on the graph based upon the average achievement and the PVAAS Average Growth Index. For each scatterplot:

- The vertical line set at 0 represents that the group of students has met the growth standard.
- A PVAAS Average Growth Index significantly above zero (to the right of the zero line) indicates academic growth greater than the growth standard.
- A PVAAS Average Growth Index significantly below zero (to the left of the zero line) indicates that the group of students did not meet the growth standard.

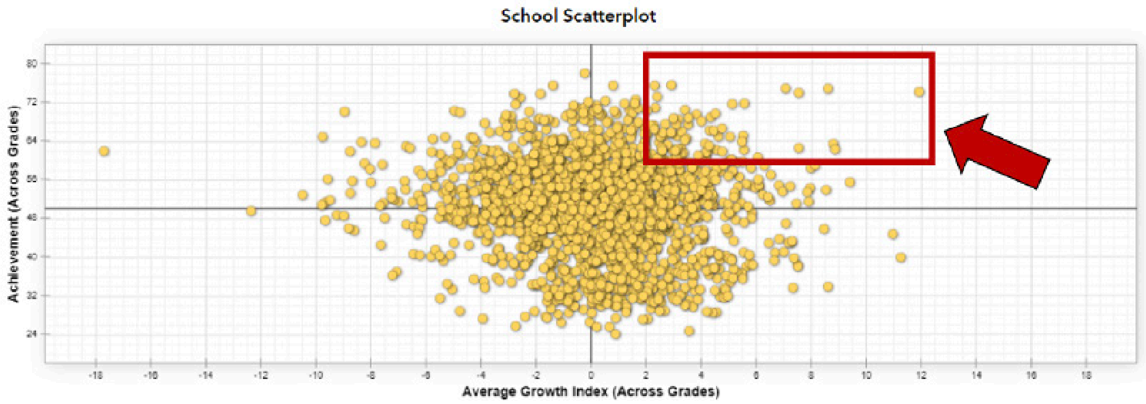
MATH

Rectangle showing schools with High Achievement & High Growth



ELA

Rectangle showing schools with High Achievement & High Growth



Measuring Academic Growth during a Transition of Pennsylvania's Assessment System

PDE's Executive Leadership, PDE's Bureau of Assessment and Accountability, along with the PVAAS Statewide Team, work with SAS EVAAS continuously as any changes are considered, or implemented, with Pennsylvania's statewide assessment system. This work is done to ensure continued PVAAS reporting for Pennsylvania's schools. The goal has always been, and will continue to be, to provide fair, accurate, and meaningful value-added measures for Pennsylvania LEAs/districts, schools, and educators.

Pennsylvania's state assessment data is assessed annually to ensure the quality needed to provide value-added reporting at all levels for all students, such as reliability and sufficient stretch to measure the growth of students with higher and lower achievement.

When a gap year occurs in state assessments or a change is made in Pennsylvania's state standards or state assessments, or even if the assessment is being given for the first time in a particular subject, academic growth as measured by PVAAS can be calculated – as long as sufficient evidence exists regarding the relationships between and among the assessments, as well as the strength of these relationships. Examples of these types of situations include, but are not limited to:

- A gap year in state assessments (which may occur with extended school closures)
- Tests change in rigor, or have a change in level of difficulty
- Tests are transitioned to be aligned to new academic standards
- Tests given for the first time statewide in a particular subject, grade level, or course
- Tests are shortened or lengthened

How can schools be expected to make growth if there is a gap year in state assessments, or if achievement changes significantly statewide?

When a gap occurs in state assessments, such as with extended school closures, it is still possible to measure the academic growth of groups of students in the year the assessments resume – when growth is measured using the PVAAS methodologies. Just like when other types of transitions or changes in testing occur, we may expect the achievement of student to change as a result. Even if the statewide achievement or performance changes significantly, PVAAS assesses whether a group of students exceeded, met, or fell short of the growth standard. For example, when measuring growth for PSSA Math and ELA, the PVAAS growth models analyze whether the group of students maintained their relative position on the statewide distribution of scores relative to themselves. If the group of students maintains the same relative position in the distribution of scores from one year to another (even if the overall statewide distribution of scores is lower), then the group of students will have met the growth standard.

SAS EVAAS does not use one, single model in its analyses. Rather, SAS EVAAS utilizes multiple robust, longitudinal statistical models that are flexible and can accommodate for:

- The use of both historical and current assessment data when tests change over time;
- The use of all longitudinal data for each student, even when the historical assessment data is on differing scales;

- Students with missing test scores without introducing major biases that come from either eliminating data from students or by using overly simplistic imputation procedures; and
- Data challenges associated with both student and teacher mobility.

PVAAS (which is based upon the SAS EVAAS models) utilizes two general types of value-added models to measure the academic growth students are making over time.

1. The growth standard methodology is used for tests given in consecutive grades, like the PSSA Math and ELA assessments implemented in grades 3 through 8.
2. The predictive methodology is used when a test is given in non-consecutive grades, or for other types of testing scenarios – like the PSSA Science or Keystone assessments.

Conceptually, growth compares the entering achievement of a group of students to their current achievement. Both methodologies used to measure growth in PVAAS compare the growth of a group of students to an expected amount of growth and provide information as to whether there is statistical evidence that the group of students exceeded, met, or did not meet that expectation of growth. No matter which type of methodology is used by SAS EVAAS to measure growth, PVAAS analyzes the relationship between and among Pennsylvania's state assessments when determining the entering achievement of students and estimating growth for a group of students.

The use of PVAAS for assessing growth in all grades, subjects, and courses is predicated upon the strong relationships that do in fact exist between and among Pennsylvania's state assessments. The multiple correlation coefficients between and among Pennsylvania's state assessments are very strong correlations, meaning that the relationship between these assessments across subjects is strong. To illustrate using the correlations between 2022-23 and students' prior testing history, the average multiple correlation for:

- PSSA Math using all prior test scores ranges from 0.82 to 0.85;
- PSSA ELA using all prior test scores ranges from 0.81 to 0.83;
- PSSA Science using all prior test scores ranges from 0.77 to 0.79; and
- Keystones using all prior test scores ranges from 0.77 to 0.79.

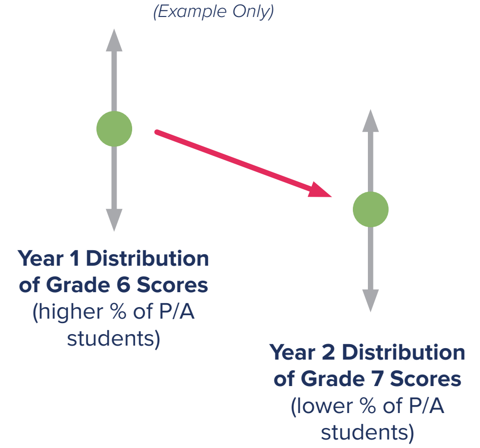
Remember that correlations range between -1.0 and +1.0, with values of -1.0 and +1.0 indicating a perfect relationship, which are virtually unheard of in the real world. In general, correlations above or below 0.6-0.7 are considered to indicate a strong relationship. (For more information, reference these resources on Interpreting Correlation Coefficients or What is R Value Correlation?)

Visual Example

The visuals provided here demonstrate how growth is measured, and more importantly, how a group of students can meet the growth standard if (1) the percentage of students reaching proficiency is lower due to an assessment change, or (2) a gap year in state assessments occurs.

In Figure 1, the first vertical line represents what could be a year 1 distribution of Grade 6 scores, with scores at the top representing higher achievement while those at the bottom represent lower achievement. The green dot represents where the student group of Grade 6 students scored relative to the statewide distribution.

Fig. 1. Change in Student Performance
(Example Only)

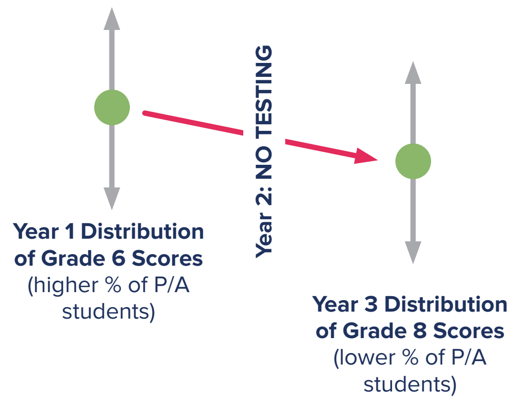


The second vertical line represents the year 2 distribution of Grade 7 scores. In year 2 of the example, fewer students scored proficient on the state assessment than in year 1, as represented by the green dot. Remember, though, that the distribution of achievement of all students statewide is lower in the example – meaning fewer students were proficient in year 2 as compared to year 1.

The question is, “Is the group of students at the same RELATIVE position in the distribution of statewide scores?” This is what PVAAS looks at to determine growth. Notice in this example that they are at the same relative position, which means this group of students met the growth standard, indicated by a green in PVAAS. In other words, these students “maintained their achievement.”

Figure 2 represents a similar example with one difference. Like the previous example in Figure 1, the first vertical line represents what could be a year 1 distribution of Grade 6 scores, with the green dot represents where the student group of Grade 6 students scored relative to the statewide distribution in that year. The second vertical line, however, now represents the year 3 distribution of Grade 8 scores as a gap year in testing occurred in year 2. In year 3 of this example, fewer students scored proficient on the state assessment than in year 1, as represented by the green dot. As in the prior example, the distribution of achievement of all students statewide is lower in our example – meaning fewer students were Proficient in year 3 as compared to year 1.

Fig. 2. Gap Year of Testing
(Example Only)

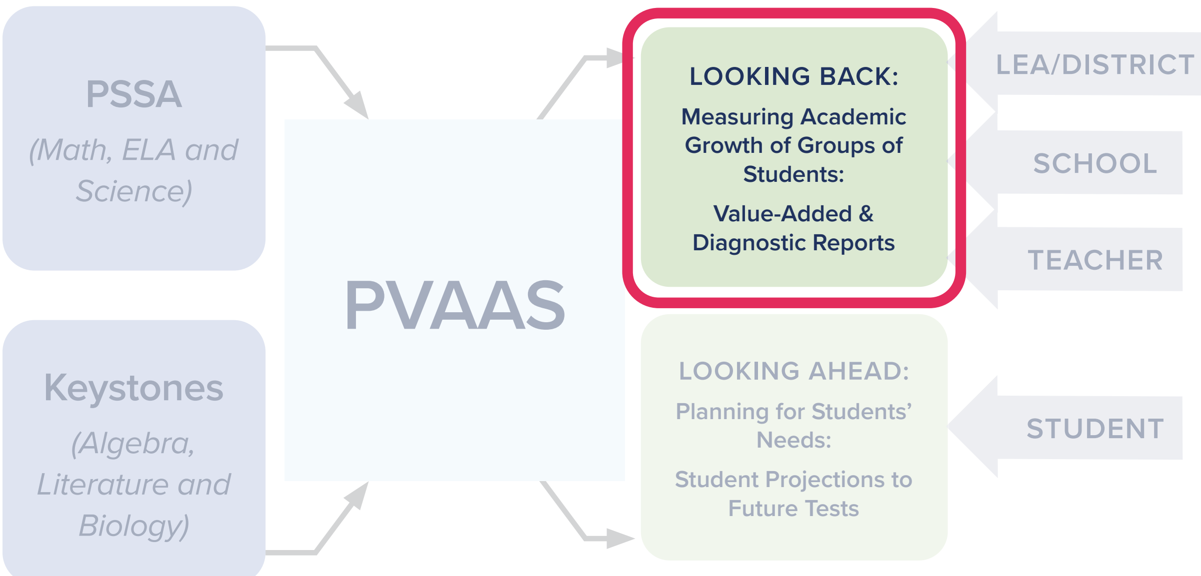


The question again is, “Is the group of students at the same RELATIVE position in the distribution of statewide scores?” They are in fact at the same relative position, which means this group of students met the growth standard, indicated by a green in PVAAS. In other words, these students “maintained their achievement.”

Growth Standard Methodology:
Measuring Growth in ELA and Math, Grades 4-8

As indicated in an earlier section, PVAAS provides information on the academic growth students have made in the last school year. This helps LEAs/districts, schools, and teachers assess how their educational programs and instructional strategies are impacting the academic growth, or progress, of groups of students.

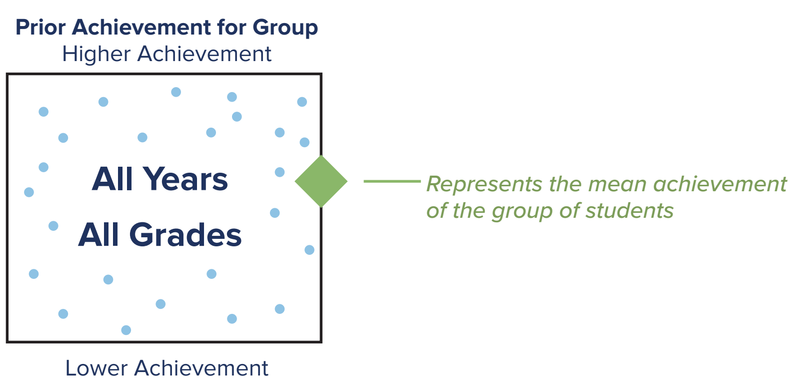
PVAAS utilizes two different methodologies for measuring growth and estimating growth measures for groups of students. These two ways of measuring growth are based on whether or not the state assessments are administered in consecutive grade levels in the same subject. This section details how growth is measured in ELA and Math in grades 4 through 8 where we assess students in consecutive grade levels each year.



Note: All prior data are used.

Conceptual Look at Measuring Growth in ELA and Math, Grades 4-8

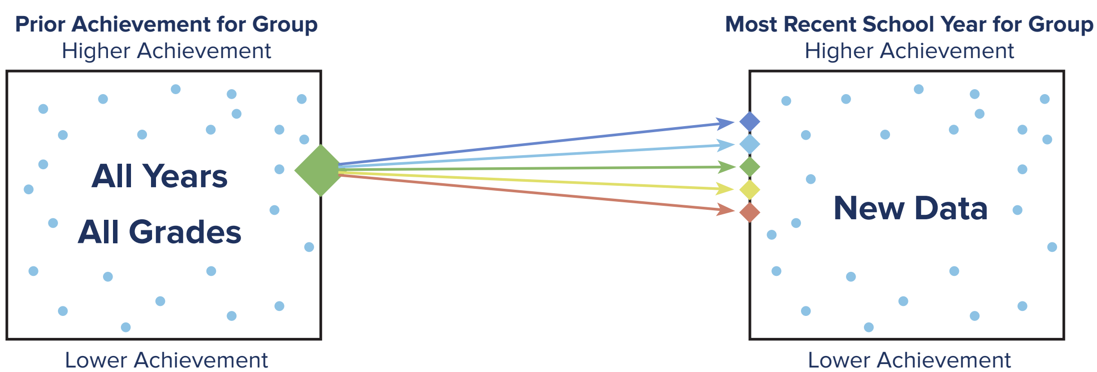
When assessments are administered in consecutive grade levels as they are in Math and ELA for grades 3 through 8, SAS EVAAS uses a Growth Standard Methodology to measure the growth of a group of students. Each year, these students are tested with the PSSA. In the graphic below, each dot represents a student in the group. Students in this group are going to score anywhere along the achievement distribution from low to high achievement. To estimate the achievement of this group of students, PVAAS uses all prior student level data from all grades in both ELA and Math to estimate the average achievement of the group.



After the next round of testing in the following school year, PVAAS again uses all the prior data for the group of students and adds the data from the most recent round of testing to get a new estimate of the average achievement of the group.

Represents the mean achievement of the group of students

Next, PVAAS compares the prior achievement of the student group to the new achievement of the student group. This comparison allows us to estimate the academic growth the student group has made in the past school year. Each Growth Measure is color-coded to ease in the interpretation of growth for each group of students.



Represents the mean achievement of the group of students

GREEN If the evidence indicates the achievement is roughly at the same point within the distribution of scores as the previous year, this indicates that the group of students maintained their achievement, which in PVAAS would be indicated as Green – evidence that the student group met the growth standard.

LIGHT BLUE If there is moderate evidence that the average achievement of this group of students was at a higher point within the distribution of scores, this would indicate that the group of students increased their average achievement indicated with a Light Blue in PVAAS – moderate evidence that the student group exceeded the growth standard.

DARK BLUE Similarly, if there is significant evidence that the average achievement of this group of students was at a higher point within the distribution of scores, this would indicate that the group of students increased their average achievement – indicated with a Dark Blue in PVAAS – significant evidence that the student group exceeded the growth standard.

YELLOW If there was moderate evidence that the average achievement of this group of students was at a lower point within the distribution of scores, this would indicate that the average achievement of the group of students slipped indicated with a Yellow – moderate evidence that the student group did NOT meet the growth standard.

RED Similarly, if there was significant evidence that the average achievement of this group of students was at a lower point within the distribution of scores, this would indicate that the average achievement of the group of students slipped – indicated with a Red – significant evidence that the student group did NOT meet the growth standard.

A More Detailed Look at Measuring Growth in ELA and Math, Grades 4-8

For the PVAAS Growth Standard Methodology:

1. Using all the available prior PSSA data in Math and ELA, PVAAS calculates a robust estimate of achievement for a group of students.
2. The Growth Measure (or gain) from the previous year to the current year can then be estimated.
3. The Growth Measure is then compared to the growth standard.

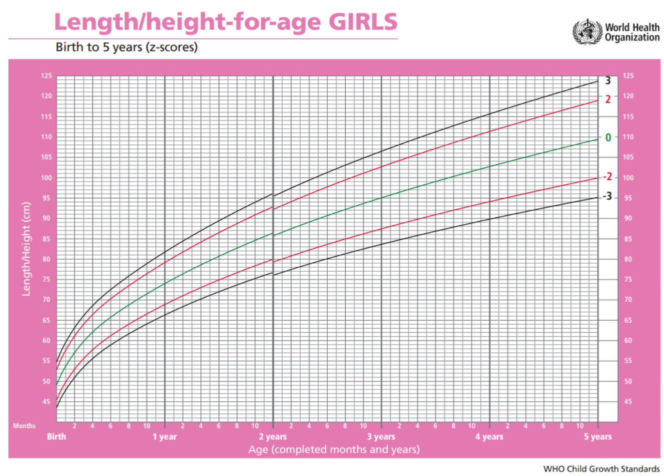
There are several important issues to consider when discussing how PVAAS calculates a measure of growth using the Growth Standard Methodology. First, when measuring the gain students make academically from one year to the next, we are measuring their growth by the difference in their academic achievement in consecutive years or grade levels. This simple idea becomes more complicated when we consider that the achievement scores from year to year must be comparable, or in other words, on the same scale. In Pennsylvania, the PSSA scale scores are NOT comparable from year to year as the scales are a bit different from year to year, grade to grade, and subject to subject. To make these scores comparable from year to year, SAS EVAAS converts all PSSA scale scores to NCEs for scores to be on the same scale and be comparable from year to year, including during any transition of a state assessment. The NCE distribution is similar to that for percentiles, with the important distinction that the NCE distribution is an equal-interval scale such that moving from 50 to 60 represents the same distance as moving from 70 to 80.

Second, if gain or academic growth is measured as the difference between consecutive years' achievement, then it is critical that the achievement reflects the best estimate of students' true level of understanding and competency. While any single year PSSA score is an estimate of achievement, it is still a single point in time. To obtain the best estimate of student achievement, SAS EVAAS uses all of the prior Math and ELA data rather than just a single year or single subject to obtain a better estimate of students' true achievement each year.

Let’s consider an analogy of why this is a best practice. Think about a teacher’s sample grade book like we have below.

	Quiz 1	Test 1	Quiz 2	Project 1	Test 2	Test 3	Final
Possible Max Score	25	100	30	75	70	100	**N/A**
Joe	20	97	23	72	65	55	**83**
Maria	14	76	27	65	56	56	**73.5**
Elizabeth	21	96	25	60	64	78	**86**

We clearly would not use Joe’s last grade of 55 as the only score to represent his achievement or understanding. For the same reason, we would not want to use only the last PSSA score solely to represent a student’s achievement. Rather, PVAAS uses all the Math and ELA data for the group of students to estimate the true level of understanding and achievement. Scores in both Math and ELA are used, as there is a strong relationship, or correlation, between students’ scores in these subjects statewide. The use of both subjects allows the Growth Measure to be more reliable and accurate.



World Health Organization (n.d.). Child Growth Standards: Length/height-for-age, Girls: birth to 5 years (z-scores). Source: https://www.who.int/tools/child-growth-standards/standards/length-height-for-age

Third, how do we know that the estimated gain or academic growth is what we should expect to see for a group of students? In Pennsylvania, academic growth is compared to the growth standard, which is based on the philosophy that despite the entering achievement of a group of students, they should not lose ground academically. The growth standard represents a group of students maintaining their average achievement as they move from grade to grade in consecutive years. When thinking about this standard, consider the physical growth chart for children (sample at right). A child is considered to meet the minimal expectation for physical growth if they maintain or improve their position in the distribution of length/height as they grow from year to year. Academically, this same concept applies. A group of students meets the growth standard if the group maintains its position in the statewide database of achievement scores as the group moves from grade to grade.

Once all achievement estimates are calculated, the gain (or Value-Added Growth Measure) from the previous year to the current year can be calculated and compared to the growth standard. If the current achievement of the student group is equal to the prior years’ achievement of the student group, then the difference between the two estimates of achievement would be 0.0; in other words, students’ achievement was maintained from one year to the next. This would be identified as a green in PVAAS indicating the group of students met the growth standard.

All Growth Measures reported on the PVAAS reports are estimates. There is natural error involved with any estimate, and this error, or variation, is expressed in terms of the Standard Error. The Standard Error allows you to determine the amount of evidence around the Growth Measure to determine if significant growth, or a lack of growth, is in fact evident for the group of students in question. One of the major functions of the Standard Error is that it allows us to evaluate the significance or level of evidence that the estimate provides that the indicated phenomenon is occurring. When we interpret a Growth Measure, we use the Standard Error to determine if the level of evidence supports that the group of students in question has exceeded or fallen short of the growth standard.

All five colors associated with the PVAAS Value-Added Growth Measure are explained below.

DARK BLUE, or DB, is an indication that the Growth Measure is more than 2 standard errors above 0. There is significant evidence of exceeding the growth standard.

LIGHT BLUE, or LB, is an indication that the Growth Measure is at least 1 but less than 2 standard errors above 0. There is moderate evidence of exceeding the growth standard.

GREEN, or G, is an indication that the Growth Measure is less than 1 standard error above 0 and no more than 1 standard error below 0. There is evidence of meeting the growth standard.

YELLOW, or Y, is an indication that the Growth Measure is more than 1 but no more than 2 standard errors below 0. There is moderate evidence of not meeting the growth standard.

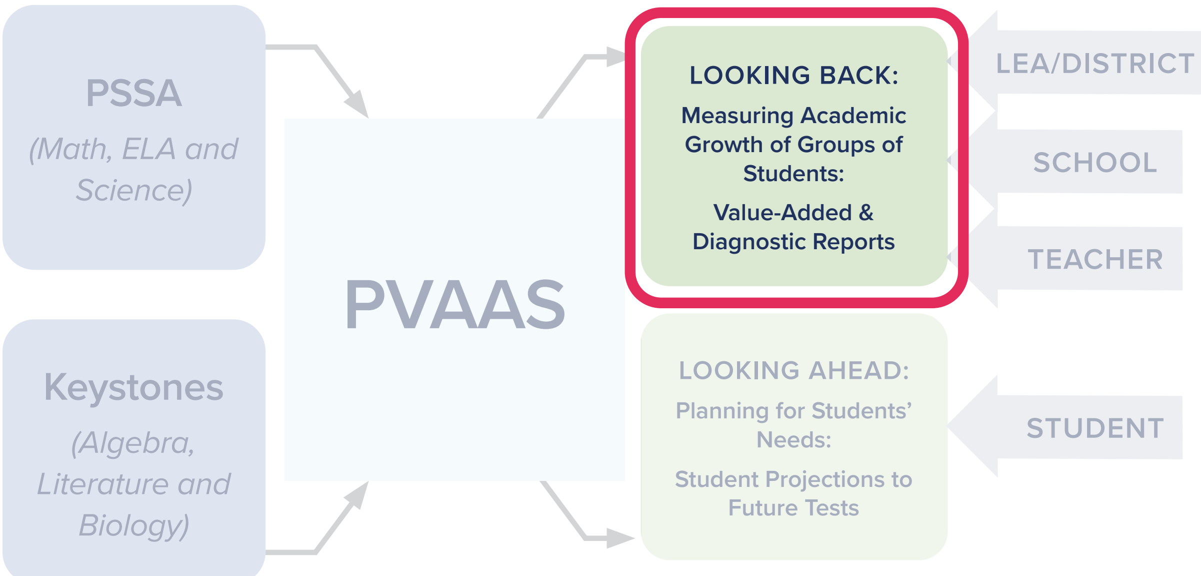
RED, or R, is an indication that the Growth Measure is more than 2 standard errors below 0. There is significant evidence of not meeting the growth standard.

For additional information on measuring growth using the SAS EVAAS statistical models on which PVAAS is based, please reference the following resources.

1. Sanders, W. L., Saxton, A. M., and Horn, S. P. (1997). "The Tennessee Value-Added Assessment System: A Quantitative, Outcomes-Based Approach to Educational Assessment." In Millman, J. (ed.) *Grading Teachers, Grading Schools*, Thousand Oaks, CA: Sage Publications, pp. 137-162.
2. Technical Documentation of PVAAS Analyses (2020). Available at: Statistical Models and Business Rules of PVAAS Analyses.

Predictive Methodology:
Measuring Growth in Science and Keystone Content Areas

Remember, as indicated in earlier sections, PVAAS provides information on the academic growth students have made in the last school year. This helps districts, schools, and teachers assess how their educational programs and instructional strategies are impacting the academic growth of groups of students.

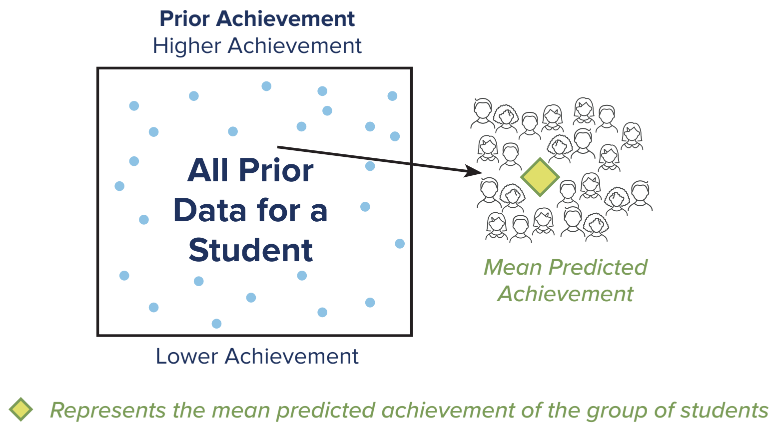


Note: All prior data are used.

PVAAS utilizes two different methodologies for measuring growth and estimating growth measures for groups of students. These two ways of measuring growth are based on whether or not the state assessments are administered in consecutive grade levels. This section details how growth is measured in Science and the Keystone content areas where students are not necessarily assessed in consecutive school years.

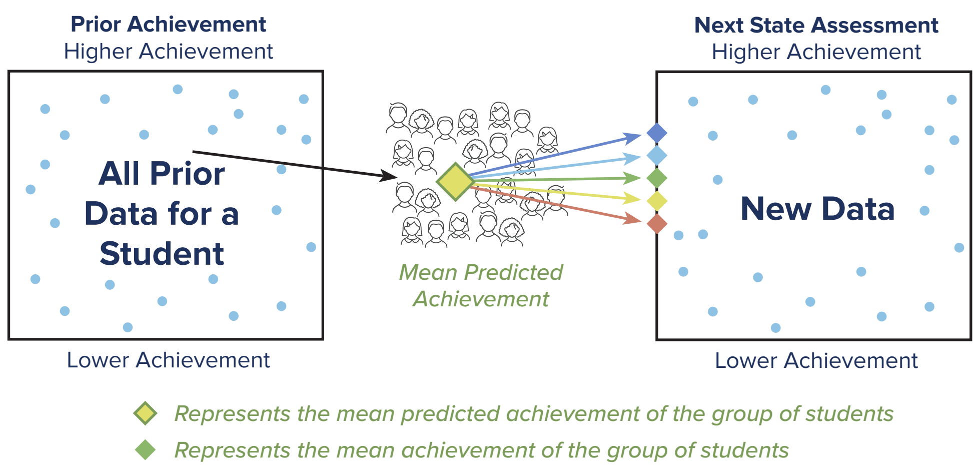
Conceptual Look at Measuring Growth in Science and the Keystone Content Areas

When the PSSA is not administered in consecutive grade levels as in Science or the Keystone content areas, SAS EVAAS uses a Predictive Methodology to measure the growth of a group of students. In these subjects, students are only assessed in specific grade levels or at the end of a specific course. In the graphic below, each dot represents a student in the group where prior state assessment data for that student are used. Those students are going to score along the achievement distribution from low to high achievement. PVAAS uses the student's prior data (as outlined in the table in the subsequent section) to predict the scale score or achievement for a student on their next state assessment in the subject or content area of interest. These predicted scale scores for all students are used to calculate the mean or average predicted achievement for the student group.



Next, these students are tested on the state assessment of interest. These students will again score along the achievement distribution from low to high achievement. The average actual achievement is calculated for the group. A comparison of the average actual achievement to the average predicted achievement is then made to estimate the Growth Measure for the student group.

.



Each Growth Measure is then color-coded to ease in the interpretation of growth.

GREEN If the evidence indicates the achievement is roughly the same as what was predicted, this indicates that the group of students maintained their achievement, which in PVAAS would be indicated as Green – evidence that the student group met the growth standard.

LIGHT BLUE If there is moderate evidence that the average achievement of this group of students was higher than predicted, this would indicate that the group of students increased their average achievement indicated with a Light Blue in PVAAS – moderate evidence that the student group exceeded the growth standard.

DARK BLUE Similarly, if there is significant evidence that the average achievement of this group of students was higher than predicted, this would indicate that the group of students increased their average achievement – indicated with a Dark Blue in PVAAS – significant evidence that the student group exceeded the growth standard.

YELLOW If there was moderate evidence that the average achievement of this group of students was lower than predicted, this would indicate that the average achievement of the group of students slipped indicated with a Yellow – moderate evidence that the student group did NOT meet the growth standard.

RED Similarly, if there was significant evidence that the average achievement of this group of students was lower than predicted, this would indicate that the average achievement of the group of students slipped – indicated with a Red – significant evidence that the student group did NOT meet the growth standard.

A More Detailed Look at Measuring Growth in Science and the Keystone Content Areas

The methodology used to estimate growth for Science and the Keystone content areas is a different methodology than what is used for estimating growth for ELA and Math in grades 4-8. The reason behind the different methodology is that in Science and Keystone content areas we do not necessarily test the same students in consecutive grade levels, whereas we do in the subjects of ELA and Math in earlier grade levels.

Students who take the Science exam will have PSSA data in ELA and Math from prior years. Additionally, students who take the Keystone exams will have PSSA data from prior years. The Predictive Methodology used for estimating growth in these subjects uses all prior data (across subjects) to calculate predicted scores for the student group. The predicted score for a student is calculated by observing how all students with a similar prior testing history performed on the test of interest, so we have an expectation of how that student should score. Growth is then a function of the difference between the average predicted score of the student group and the average actual score for the student group. Let's take a closer look at this.

Each year, Pennsylvania state assessment data (PSSA and Keystone) are sent directly from the test vendor to the SAS EVAAS team for PVAAS analyses. For the Predictive Methodology discussed in this section, the PVAAS value-added analysis begins at the student level with the collection of individual student data for a minimum of three prior data points (only exception is grade 4 Science analyses where two data prior points are used). Note that each scale score in each subject is considered a data point. By looking at the relationships between student scores, PVAAS is able to predict future achievement for students. In other words, PVAAS predicts students' future achievement based on students' past performances and the relationships between those subjects.

You may be asking, "Are scores across subjects truly related? Is there a correlation across subjects?" The answer to that is yes, students' prior test scores are in fact related to the current test we are investigating. In other words, the Keystone Algebra I test is highly related to PSSA Math scores in earlier grades. The Keystone Algebra I test is also related (perhaps to a lesser extent) to PSSA ELA and Science scores in earlier grades. Each student's prediction is based on the student's prior test scores, the relationship of those scores to the current year test, and the average schooling experience in the state for the current year. In more technical terms, the multiple correlation coefficient provides the correlation between the predicted score using all prior test scores and the actual score itself. The multiple correlation coefficients between Keystones and PSSA are very strong correlations, meaning that the PSSA provides good data for the student predictions. To illustrate using the correlations between 2022-23 and students' prior testing history, the average multiple correlation for:

- PSSA Math using all prior test scores ranges from 0.82 to 0.85;
- PSSA ELA using all prior test scores ranges from 0.81 to 0.83;
- PSSA Science using all prior test scores ranges from 0.77 to 0.79; and
- Keystones using all prior test scores ranges from 0.77 to 0.79.

Remember that correlations range between -1.0 and +1.0, with values of -1.0 and +1.0 indicating a perfect relationship, which are virtually unheard of in the real world. In general, correlations above or below 0.6-0.7 are considered to indicate a strong relationship. (For more information, reference these resources on [Interpreting Correlation Coefficients] or [What is R Value Correlation?])

When calculating the predicted scores for students, PVAAS specifically utilizes the historical state assessment data, when available, to predict future PSSA performances in Science and future Keystone performances in Algebra I, Biology, and Literature. The table provided below illustrates the data used for each subject.

Predicted Score IN...	Prior Years' Data Used to Calculate Predicted Score
PSSA Science	All prior scores in PSSA Math, ELA, and Science (if available)
Keystone Algebra I	All prior scores in PSSA Math, ELA, and Science (if available)
Keystone Biology	All prior scores in PSSA Math, ELA, and Science; and most recent prior score in Keystone Algebra I (if available)
Keystone Literature	All prior scores in PSSA Math, ELA, and Science; and most recent prior score in Keystone Algebra I and Biology (if available)

While the predicted scores for students are calculated, another analysis is run simultaneously to compute the average of these same students' actual scores. The Growth Measure is then estimated by comparing the average of actual scores to the average of the predicted scores. If students' average actual score is equal to their average predicted score, then the Growth Measure would be 0.0, or in other words, students scored as predicted given their previous testing histories. This would be identified as a green in PVAAS indicating the group of students met the growth standard.

All five colors associated with the PVAAS Value-Added Growth Measure are explained below.

- **DARK BLUE**, or DB, is an indication that the Growth Measure is more than 2 standard errors above 0. There is significant evidence of exceeding the growth standard
- **LIGHT BLUE**, or LB, is an indication that the Growth Measure is at least 1 but less than 2 standard errors above 0. There is moderate evidence of exceeding the growth standard.
- **GREEN**, or G, is an indication that the Growth Measure is less than 1 standard error above 0 and no more than 1 standard error below 0. There is evidence of meeting the growth standard.
- **YELLOW**, or Y, is an indication that the Growth Measure is more than 1 but no more than 2 standard errors below 0. There is moderate evidence of not meeting the growth standard.
- **RED**, or R, is an indication that the Growth Measure is more than 2 standard errors below 0. There is significant evidence of not meeting the growth standard.

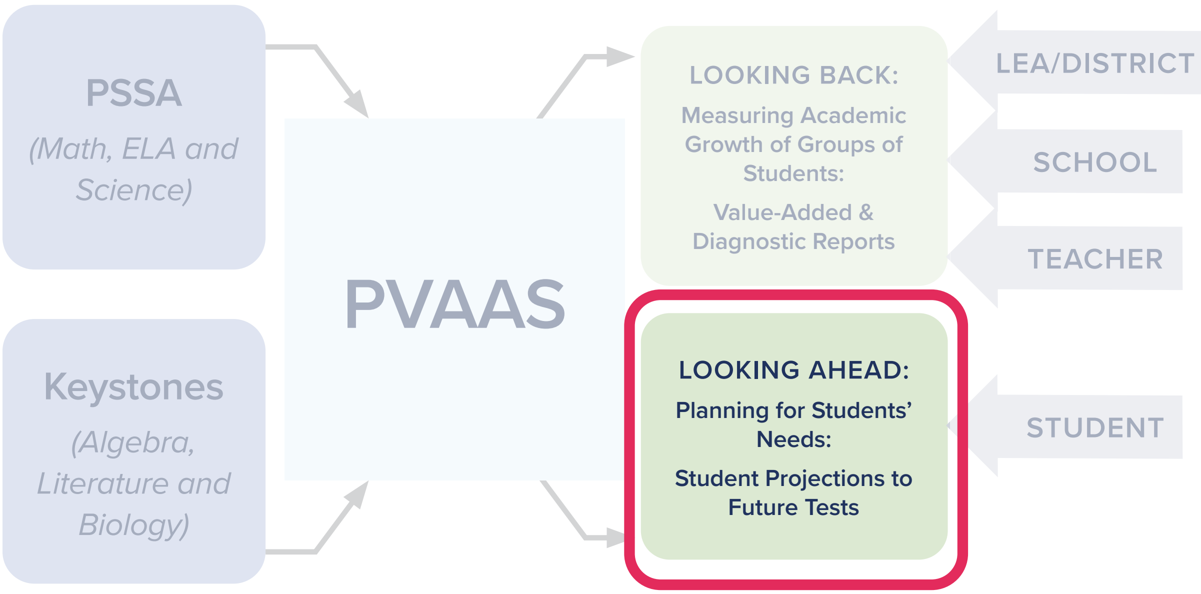
For additional information on measuring growth using the SAS EVAAS statistical models on which PVAAS is based, please reference the following resources.

1. Sanders, W. L., Saxton, A. M., and Horn, S. P. (1997). "The Tennessee Value-Added Assessment System: A Quantitative, Outcomes-Based Approach to Educational Assessment." In Millman, J. (ed.) *Grading Teachers, Grading Schools*, Thousand Oaks, CA: Sage Publications, pp. 137-162.
2. Technical Documentation of PVAAS Analyses (2020). Available at: [Statistical Models and Business Rules of PVAAS Analyses].

Projection Methodology

Estimating Projections to Students' Achievement on Future Assessments

In addition to providing measures of academic growth, PVAAS provides projected scores for individual students on future assessments the students have NOT yet taken, as well as for students not yet Proficient on a Keystone exam. Projections are available to PSSA, Keystone, PSAT, SAT, ACT, and Advanced Placement (AP) exams, as well as the ACCESS for ELLs assessment given to English Learners. These projections can be used to predict a student's current academic trajectory and may be used to guide counseling and intervention to increase students' likelihood of future success.



Note: All prior data are used.

It can be useful to think about the PVAAS projections as being very similar to expectations we set in our everyday lives. Each day, whether we realize it or not, we generate expectations – about our work, our personal lives, and the world in general. These expectations are based on two things: information and experience. The individual student projections that are provided in PVAAS also rely on information and experience. In this case, the information is a student's testing history, across grades and subjects. The experience is determined by students with similar prior achievement who have already taken the test of interest for the projection, so that students' prior test scores across grades and subjects serve as predictors to their projection to a future assessment. Using this data, the projection model quantifies the projection in a precise and reliable way. In PVAAS, this means that projections indicate what is most likely to happen for a student academically, or in other words, how a student will most likely score on a future assessment if they continue on the same academic path.

Most important to consider is that the projection methodology uses historical assessment data from the most recent five years, when available, to project future PSSA performances in Math, ELA, Science, and future Keystone performances in Algebra I, Biology, and Literature. Projections to future college readiness exams are also available for PSAT, AP, ACT, and SAT. The table below illustrates the specific data used for each subject and exam.

Projection TO...	Prior Years' Data Used to Calculate Predicted Score
PSSA Math	All prior scores in PSSA Math and ELA (if available)
PSSA ELA	All prior scores in PSSA Math and ELA (if available)
PSSA Science	All prior scores in PSSA Math, ELA, and Science (if available)
Keystone Algebra I	All prior scores in PSSA Math, ELA, and Science (if available)
Keystone Biology	All prior scores in PSSA Math, ELA, and Science; and most recent prior score in Keystone Algebra I (if available)
Keystone Literature	All prior scores in PSSA Math, ELA, and Science; and most recent prior score in Keystone Algebra I and Biology (if available)
PSAT, SAT, ACT, and AP	All prior scores in PSSA Math, ELA, and Science; and most recent prior score in Keystone Algebra I, Biology, and Literature (if available)
ACCESS for ELLs	All prior ACCESS for ELLs domain scores in Listening, Speaking, Reading, and Writing

The reason the predictors listed above can be used in calculating projections is that students' prior test scores are in fact related to the current test we are investigating. For example, the Keystone Algebra I test is highly related to PSSA Math scores in earlier grades. The Keystone Algebra I test is also related (perhaps to a lesser extent) to PSSA ELA and Science scores in earlier grades. Each student's projection is based on the student's prior test scores, the relationship of those scores to the current year test, and the average schooling experience in the state for the current year. In more technical terms, the multiple correlation coefficient provides the correlation between the predicted score using all prior test scores and the actual score itself. The multiple correlation coefficients between Keystones and PSSA are very strong correlations, meaning that the PSSA provides good data for the student predictions. To illustrate using the correlations between 2022-23 and students' prior testing history, the average multiple correlation for:

- PSSA Math using all prior test scores ranges from 0.82 to 0.85;
- PSSA ELA using all prior test scores ranges from 0.81 to 0.83;
- PSSA Science using all prior test scores ranges from 0.77 to 0.79; and
- Keystones using all prior test scores ranges from 0.77 to 0.79.

Remember that correlations range between -1.0 and +1.0, with values of -1.0 and +1.0 indicating a perfect relationship, which are virtually unheard of in the real world. In general, correlations above or below 0.6-0.7 are considered to indicate a strong relationship. (For more information, reference these resources on Interpreting Correlation Coefficients or What is R Value Correlation?).

Projections can be made for any students with any set of available predictor scores defined in the projection model. However, to protect against bias due to measurement error in the predictors, projections are made only for students who have at least three available predictor scores (except for projections to grade 4 assessments in which two available predictor scores are used). In addition to the projected score itself, the standard error of the projection is calculated. Given a projected score and its standard error, it is possible to calculate the probability that a student will reach a specified benchmark of interest. In the case of Pennsylvania, probabilities of performing at a Basic level or higher, at a Proficient level or higher, or at an Advanced level on a future PSSA or Keystone exam are calculated. Probabilities of performing at or above specified college benchmarks on the PSAT, SAT, and ACT are also calculated, as well as probabilities of scoring a 3 or higher, a 4 or higher, or a 5 on an AP exam. These probabilities are calculated as the area above the benchmark cut score using a normal distribution with its mean being equal to the projected score and its standard deviation being equal to the standard error of the projected score.

A More Detailed Look at the Projection Methodology

The PVAAS Projection methodology is in fact quite complex and involves the use of many data points across multiple subjects and multiple years. Let's look at a more detailed perspective of this methodology to provide further insight into this modeling process.

To do this, let's consider an example. Suppose we wish to project the 8th grade Math score for a student named Tyler. This student was new to Pennsylvania at the beginning of his 5th grade year. To date, we have the following state assessment data available for Tyler from grades 5-7:

1. Grade 5 Math and grade 5 ELA
2. Grade 6 Math and grade 6 ELA
3. Grade 7 Math and grade 7 ELA

First, we would create a projection model for 8th grade Math based on all of the students in the state who have already taken 8th grade Math and have the same prior scores as Tyler. The next step is to construct a formula, which can take all of those prior test scores and make the best possible prediction for all the students in the state who have already taken 8th grade Math.

Grade	5th	6th	7th	8th
Math	42	54	47	53
ELA	47	51	53	54

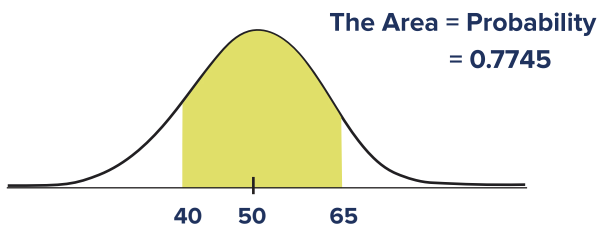
In this example:

> Projected Score (Grade 8 Math) = a (grade 5 Math) + b (grade 5 ELA) + c (grade 6 Math) + d (grade 6 ELA) + e (grade 7 Math) + f (grade 7 ELA) + g,
>
> *where a, b, c, d, e, f, and g are estimated numbers based on correlations across subjects, and grade 5 Math, grade 5 ELA, etc. are test scores.*

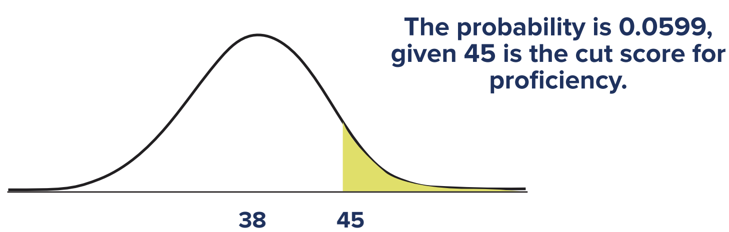
To then calculate the projected 8th grade Math score for Tyler, we need to (1) substitute the values of the test scores for Tyler into the formula, and (2) calculate the estimate of his 8th grade Math score. The estimate from this model would then be the projected score for Tyler. Since the projected score is an estimate, we would also calculate the standard error for that estimate. In our example, once Tyler's scores are inputted into the model, we get a projected score of 38 with a standard error of 4.5.

Now, what does the standard error really tell us? What can we do with the standard error? We can use the estimated projected score and the standard error to calculate the probability that the student will score in a selected performance level or higher. This probability is the PVAAS projection probability. The key to calculating the PVAAS projection probability is to consider the estimated projected score as the average, or mean, of all possible estimates, and the standard error as the standard deviation of all the possible estimates. With this consideration in mind, it is reasonable to think that all possible estimates would form an approximately normal distribution with the estimated projected score as the mean and the standard error of the estimated projected score as the standard deviation. We can then use statistical normal distribution techniques, and the area under the normal curve to determine probabilities.

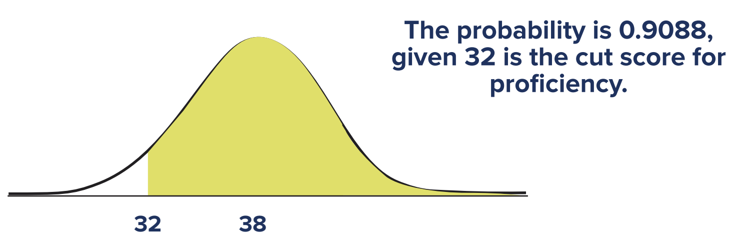
Going back to a "Stats 101" course, recall that the area under a normal curve is always equal to 1, or a probability of 100%. In a general example with a normal distribution with a mean of 50 and a standard deviation of 10, we could calculate the probability that a value would be between 40 and 65. To do that, we would calculate the area of the region under the curve that is between 40 and 65. Using statistical tables or software, we would be able to determine that the area under the curve is 0.7745, and therefore the probability that a value is between 40 and 65 is 77.45%.



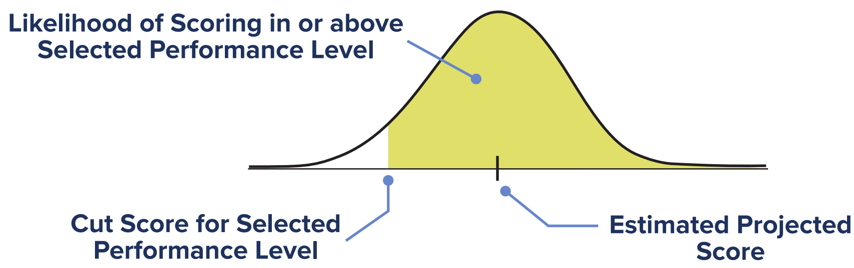
Let's revisit our example of calculating a projected score and projection probability for Tyler for 8th grade Math. Remember that we calculated Tyler's estimated projected score to be 38 with a standard error of 4.5. So, how likely is it that Tyler will score a 45 or more, knowing that the cut score for proficiency on the 8th grade PSSA Math assessment is 45? In this case, we would use the normal curve technique described above to find the area under the curve to find Tyler's projection probability of reaching a Proficient level or higher on the 8th grade Math assessment. With a cut score of 45 for proficiency on the 8th grade Math assessment, we see that Tyler has a low likelihood (only 5.99%, or 6%, probability) of reaching proficiency or higher on the grade 8 Math test given his own testing history and his projected score.



However, what if the cut score for proficiency is 32, not 45? How likely is it then that Tyler will score a 32 or more, knowing that the cut score for proficiency on the 8th grade PSSA Math assessment is 32? Again, we would use the normal curve technique to find the area under the curve to find his projection probability. With a cut point of 32 for proficiency on the 8th grade Math assessment, we see that Tyler has a high likelihood (90.88%, or 91%, probability) of reaching proficiency or higher on the grade 8 Math test given his own testing history and his projected score.



In summary, PVAAS calculates the likelihood of a student scoring in a selected performance level or higher (or selected benchmark or higher) by calculating the probability that the actual score will be above the associated cut score. PVAAS uses the estimated projected score and the standard error from the PVAAS Projection Model to calculate this projection probability.



Statistical Considerations

PVAAS analyses make use of scores from Pennsylvania's state assessments, specifically the PSSA and Keystone exams. The scales from these assessments must meet three criteria to be used in value-added and projection analyses by SAS EVAAS.

1. They demonstrate sufficient stretch so that students with a high achievement history and those with a low achievement history can show growth.
2. They are aligned to state curriculum standards.
3. The scales are reliable from year to year.

Dealing with Missing Data

A common problem with using test scores is missing data. There are many reasons why test data are missing. For example, a student could move into the district this year from out of state, a test record could be lost, or a student could be sick on the series of test days including the make-up days. The importance of dealing properly with missing data is illustrated in the following example.

Assume that ten students are given a test in two different years with the results shown below in Table 1. The goal is to measure the academic growth of this group of students from one year to the next. The right side of the table shows what happens when some of the scores are missing. Two simple approaches to take when data are missing are to calculate the average of the differences or to calculate the differences of the averages. When there are no missing data, these two simple methods provide the same answer (5.80 on the left in Table 1). However, when there are missing data, each method provides a different answer (9.57 versus 3.97 on the right in Table 1). Hence, a more sophisticated model is needed to address this problem of missing data.

Table 1. LEFT – Scores without missing data; RIGHT – Scores with missing data

Student	Grade 4	Grade 5	Gain
1	51.9	74.8	22.9
2	37.9	46.5	8.6
3	55.9	61.3	5.4
4	52.7	47.0	-5.7
5	53.6	50.4	-3.2
6	23.0	35.9	12.9
7	78.6	77.8	-0.8
8	61.2	64.7	3.5
9	47.3	40.6	-6.7
10	37.8	58.9	21.1
Average	49.99	55.79	**5.80**
Difference	**5.80**		

Student	Grade 4	Grade 5	Gain
1	51.9		
2	37.9		
3	55.9	61.3	5.4
4	52.7	47.0	-5.7
5	53.6	50.4	-3.2
6	23.0	35.9	12.9
7		77.8	
8		64.7	
9	47.3	40.6	-6.7
10	37.8	58.9	21.1
Average	45.01	54.58	**3.97**
Difference	**9.57**		

The more sophisticated statistical model used by SAS EVAAS for value-added analyses in PSSA Math and ELA uses the correlation, or relationship, between students' current and prior years' scores in the non-missing data to estimate an average for the previous and current score as if there were not missing data. The model does this without explicitly imputing values (or making up) values for the missing scores. The difference between these two estimated averages is an estimate of the average gain for this group of students.

In this small example, the estimated difference is 5.7. This method of dealing with the missing data is closer to the "no-missing data" average of 5.8 than the 3.97 answer obtained by the average of the differences and the

9.57 answer obtained by the difference of the averages. This method of dealing with missing data has been shown, on average, to outperform both simpler methods (Wright, 2004). Much larger data sets, such as those used in SAS EVAAS analyses, provide better correlation estimates, which in turn provide better estimates of averages and gains.

Accounting for Student Demographics

Due to the relationship of demographics and achievement, many assume there is the same relationship between growth and demographics. However, a review of the literature indicates that demographic variables (such as socioeconomic status and racial/ethnic background) typically have little to no significant relationship with student growth measures as long as you can sufficiently account for the prior achievement of students.

The SAS EVAAS team has found that the only way to adequately account for prior achievement is to use all available test data for each student to dampen the effects of measurement error. Hence, value-added analyses in PVAAS measure the change in students' academic achievement levels from one point in time to another using ALL prior data to sufficiently account for the prior achievement of students. Factors that remain relatively constant over time have been shown to have little or no impact on students' progress or growth. This approach has been confirmed through a variety of robust statistical analyses. In 2004, a SAS, Inc. and Vanderbilt team published a study that closely examined SES and demographic adjustments and concluded that "SES and demographic [variables] add little information beyond that contained in…test scores."

As a result, additional adjustments for the demographic status of students are not included in the analyses as it accounts for students' prior achievement. Evidence from PVAAS reporting in Pennsylvania has yielded results to show that there are many districts, schools and teachers in Pennsylvania making significant growth with EL students, minority students, economically disadvantaged students, and students with IEPs.

The following are literature reviews on the topic of growth and student demographics:

1. Ballou, D., W. Sanders, and P. Wright. 2004. "Controlling for Student Background in Value-Added Assessment." Journal of Education and Behavioral Statistics, 29(1), pp. 37-65.

2. An economist-based perspective by UCLA researchers Pete Goldschmidt, Kilchan Choi and Kyo Yamashiro provided a similar finding in their study comparing value-added models: "First, adding in an adjustment for student SES (as measured by eligibility for free- or reduced-price lunch) adds very little once a student's initial status is controlled… This indicates that student initial status captures many of the effects that SES is attempting to measure. In other words, by controlling for initial status, the model already captures the preceding effects that SES might have on students." Choi, K., P. Goldschmidt, and K. Yamashiro (2006). Exploring Models of School Performance: From Theory to Practice (CSE Report 673). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST), p. 24.

3. A single measure of student achievement has inherent limitations due to the fact that achievement is correlated to a student's socioeconomic status and past performance (Hershberg, et al.; Olson, 2007; Sanders, 2000).

4. Fallon (2003) reports that the importance of value-added assessment is it being based on the experimental design that removes virtually all influence of genetics and socio-economic factors. The design provides a measure of the direct effect of the effectiveness of schools.

5. [Value-added assessment systems] can remove the effects of factors not under the control of the school, such as prior performance and socioeconomic status, and thereby provides a more accurate indicator of school or teacher influence than is possible when these factors are not controlled (McCaffrey, Lockwood, Koretz & Hamilton, 2003; Ross, Wang, Sanders, Wright & Stringfield, 1999a; Wright, Horn & Sanders, 1997).

Resources

The SAS EVAAS methodologies and algorithms have been published in open literature since 1997. For those interested in learning more about the business rules and technical aspects of the statistical models used in all applications in Pennsylvania, please see the following two resources.

1. Sanders, W. L., Saxton, A. M., and Horn, S. P. (1997). "The Tennessee Value-Added Assessment System: A Quantitative, Outcomes-Based Approach to Educational Assessment." In Millman, J. (ed.) Grading Teachers, Grading Schools, Thousand Oaks, CA: Sage Publications, pp. 137-162.
2. Chapters 12 through 16 focus on the Education Value-Added Assessment System (EVAAS) upon which PVAAS is based.
3. Technical Documentation of PVAAS Analyses (2020). Available at: Statistical Models and Business Rules of PVAAS Analyses

There are e-Learning resources and online help available on the PVAAS website. The e-Learning resources can be found by clicking on the e-Learning button on the right side of the blue menu bar. Online help can be found for any report by clicking on the Help button on the right side of the blue menu bar when on any web-based report.

Visit **education.pa.gov/pvaas** for additional resources on this topic.