Los Angeles Unified School District

Today's Learners, Tomorrow's Leaders

Supporting All Employees (SAE) Strategy and Academic Growth over Time (AGT)
Frequently Asked Questions

Updated: October 2011

Table of Contents

POLICY ISSUES RELATED TO ACADEMIC GROWTH OVER TIME IN LAUSD

Why is LAUSD developing measures of Academic Growth over Time?

The Los Angeles Unified School District believes that success happens in the classroom. Therefore, the rigorous analysis and use of various measures of how our students are performing are core components of our work as educators. While useful and important, achievement data alone, such as proficiency rates in English and math, do not provide a complete picture of how our students are performing and ***how we are doing at improving student learning over time***. For that reason, the district is moving toward a new and comprehensive system of computing student gains (**Academic Growth over Time** - AGT), which helps us determine how much students have progressed on standardized tests from one year to the next, and most importantly, how we have done at taking students from point A to point B. AGT allows us to examine the impact that schools and educators have on student learning outcomes and uses a value-added method that controls for external factors which often influence student test results.

Measures of Academic Growth over Time inform several LAUSD strategies, including *Using Data to Drive Standards Based Instruction*, *Creating & Supporting Quality Schools,* and *Supporting All Employees.*

- AGT helps answer questions like:
  - Are students in a particular region, school or grade level growing faster than similar students from across the district?
  - Are specific groups of students in particular schools or classrooms growing faster or slower than the district average?
  - And with further observation, what instructional methods, programs and interventions are working to improve student outcomes?
  - What is the distribution of effective educators? Do we have the most effective educators working in the right places to achieve our goals?
  - What can we learn from places where we are achieving remarkable results?
- AGT is a robust statistical analysis that estimates the influence of schools and educators on student growth.
- AGT holds schools and educators accountable only for what they have direct control over, by accounting for such factors as prior achievement, a student's English Language Learner status, special education status and the like.
- AGT accounts for measurement error, which is inherent in all tests.

What results are included in the Fall 2011 release?

The Board of Education and district leadership have committed to a public release of Academic Growth over Time results in school year 2010–2011 and a confidential, no stakes release of results to educators, where the data and current modeling techniques support the provision of these results.

Our initial release of Academic Growth over Time results involves two phases.

In Phase 1, school level results were publicly released in mid-April, and educators received individual results confidentially in late May. Phase 1 results included the following grade levels and subjects:

- Grades 3 to 8
  - English Language Arts (ELA)
  - Mathematics
- Grade 9 (for first time 9th graders only)
  - English Language Arts (ELA)

Phase 2 results incorporated SY 2010–2011 CST results. Phase 2 models also involved modeling enhancements in order to expand the grade levels and subject matter addressed. We explored the following range of grade levels and subjects for Phase 2:

- Grades 3-11 ELA
- Grades 3-8 General Math
- Algebra I
- Geometry
- Algebra II
- Science Grade 5
- Science Grade 8
- Integrated Science
- Biology
- Chemistry
- Physics
- Grade 8 Social Science
- World History
- U.S. History`

NOTE: The following subjects are not being reported at the teacher-level.

Not included at the teacher-level due to multi-grade/multi-year curriculum:

- Grade 5 Science
- Grade 8 Social Science

Not included at the teacher-level due to not meeting LAUSD's criteria for inclusion (reliability, stability, predictive power, explained variance, teacher/student coverage):

- ELA Grade 10
- ELA Grade 11

Why is LAUSD releasing results for any subject and grade level before the district can include results for all subjects and grade levels?

We believe these data provide useful information for our schools, educators and many stakeholders, and so even though we cannot yet provide results for all subjects and grade levels, we are releasing those results where we have confidence in the rigor of the metrics as those results become available.

Is there a plan for including a wider set of subjects and grade levels? Is there a plan to consider other assessments in addition to the CST?

Yes. In the fall release we expanded the results to the wider set of grade levels and subjects that participate in the California Standards Tests (CSTs).With our partners at WestEd and the University of Wisconsin, we are also researching and analyzing the appropriateness of other assessments for the purpose of AGT analyses. This includes looking both at multiple choice-type assessments, as well as constructed response assessments, such as essays. New assessments will need to pass our criteria for reliability, stability, predictive power, explained variance and teacher/student coverage.

How will these results be used in school accountability?

We believe in "The Power of Two"—achievement and growth data. Currently, we emphasize achievement data in our various accountability systems. Indeed, the federally mandated requirements through No Child Left Behind regarding Adequate Yearly Progress focus entirely on whether students across an entire school and within subgroups (e.g., by racial subgroup) meet proficiency benchmarks. While our ultimate goal is proficiency for all, we recognize that students start at different places, and so we want to recognize and foster growth, as well. We will carefully consider the role of AGT in our school accountability systems for the district (e.g., when decision which schools should be part of the Public School Choice process) How will these results be used as part of local district and central office accountability?

Because we believe that success happens in our classrooms, we are developing a system of performance management around the LAUSD's five goals:

- Goal 1 – 100% Graduation
- Goal 2 – Proficiency+ for All
- Goal 3 – 100% Attendance
- Goal 4 – Parent and Community Engagement
- Goal 5 – School Safety

Our performance management system will ensure that from the individual student to the classroom to the school to the local district to the central office and to the Board room, we are all focused on and accountable to these goals. AGT will be one measure that, initially, helps us understand how we are doing to support Goal 2 – Proficiency+ for All.

How will these results be used as part of educator performance reviews?

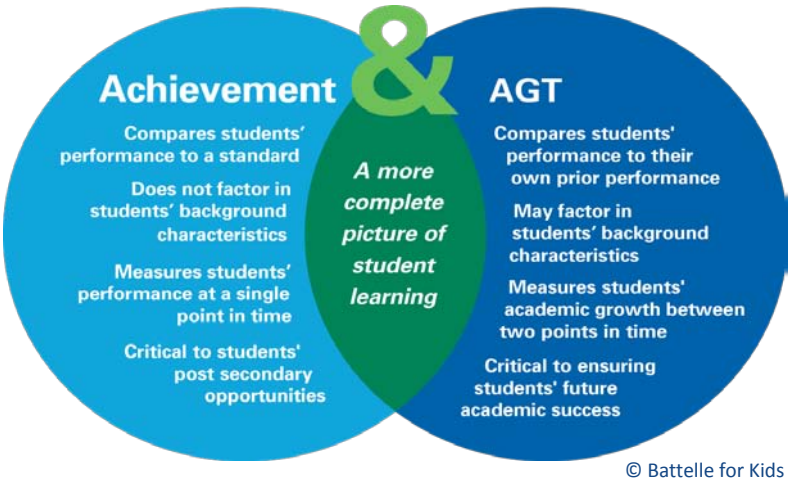
The Los Angeles Unified School District (LAUSD) recognizes that we face no more important task than to ensure that every classroom is led by an effective teacher, and that every school is led by an outstanding leader, who is surrounded by a team of excellent support personnel. To that end, the district is moving forward with plans to implement a ***multiple measures performance review system*** for all educators, where review of practice by trained observers will be the majority measure, and where, among other measures, we will include a balanced use of contributions to student outcomes.

- AGT results will be a fractional measure in the overall effectiveness rating of teachers, where observation is the majority measure.
    - It would be inappropriate to judge a teacher's effectiveness only using student outcomes data.
    - It would be inappropriate not to include student outcomes data in a teacher's evaluation.

To learn more, visit the Supporting All Employees website at http://sae.lausd.net. There is also a section of this FAQ devoted to the connection between AGT and our Supporting All Employees strategy.

Does AGT tell us whether or not students have acquired the skills and abilities to be college prepared and career ready?

No. AGT tells us which schools, grade-level teams, educators and the like are having the most significant impact on progress toward these skill abilities. Our achievement measures, such as proficiency rates and college preparatory completion rates, tell us whether or not students are college prepared and career ready. This is why we emphasize "The Power of Two"—achievement and growth:



Does AGT tell us why some schools, grade-level teams and educators are having more success producing student outcomes?

No. AGT helps us identify classrooms that are achieving remarkable growth with particular student populations. To understand the 'why' behind these results, we must examine the practices and policies in these classrooms and schools. For instance, LAUSD's new Teaching & Learning Framework will provide a common language and understanding to discuss instructional practices and teacher actions. Using this Framework, we will be able to examine which specific instructional practices are connected to noteworthy AGT results. AGT tells us where to look for examples of excellence and shows us those most in need of greater support.

Will AGT results be available for the Magnet programs within schools?

At this stage, school-level results are school-wide and not broken down by programs within schools. At a later stage, we will look into breaking down the results in this and other manners. If the data set gives a unique ID to a Magnet School, then we calculate it separately. If a magnet program is within a school, we do not have a separate result for that program.

LAUSD ACADEMIC GROWTH OVER TIME METHODOLOGY:

What is Academic Growth over Time?

Academic Growth over Time is a statistical method used to identify the individual impact of a teacher (or school leader or entire school) on student learning. Academic Growth over Time compares the performance of each teacher's students to that of teachers with similar students.

Why does AGT control for certain factors and not others?

AGT is designed to fairly measure the contribution of teachers and schools on the academic growth of students. In order to accomplish this goal, the AGT model uses students' standardized test scores in combination with student demographics to create growth predictions. The predictions are customized to the students you are serving, which allows for a fair comparison of student growth for teachers and schools serving different student populations.

What allows for these fair comparisons is removing the effect of factors beyond the control of teachers and schools. In achievement models like percent proficient/advanced and API, high results are often more indicative of the type of student the school serves rather than the effectiveness of teaching at the school. To evaluate the effectiveness of instruction, we need to measure longitudinal student growth, which is the basis of AGT.

While academic growth is the right concept to focus on for measuring effectiveness of instruction, there are factors other than the teacher or school that may have an impact on the growth of students. Some examples of factors that may affect student learning include English Language Learner status of a student, cognitive disabilities, access to books/computers at home, parental support with homework, and a multitude of other factors. The test itself may also have properties that would cause a trend that higher achieving students would gain less raw points on the test than middle or lower achieving students. Knowing that there are many factors that can affect the academic growth of students, the AGT model attempts to remove the effect of non-school factors so educators have a level playing field for AGT estimates.

To remove the effect of non-school factors, AGT first analyzes the actual growth of all students across the LAUSD to determine the effect of various factors. This is accomplished with a multivariate regression model based on test scores (prior and current) as well as student characteristics. Using this model allows us to determine the impact of specific non-school factors so we can make fair comparisons of growth at different schools.

For example, after analyzing the scores of all 5th grade math students across the LAUSD, we might find that students that qualified for Free/Reduced Price Lunch tended to grow 3 fewer scale score points on the 5th grade math students when compared against Non-Free/Reduced Price Lunch students. If we did not control for Free/Reduced Price Lunch in the model, schools serving a higher portion of Free/Reduced Price Lunch students would be at a disadvantage in a simple growth model.

In order for a factor to be included in the model, it must:

- Increase predictive power in the model
- Be out of the control of schools and teachers
- Be collected for the vast majority of the LAUSD's students
- Pass data quality / reliability tests to ensure the data is consistently collected across the LAUSD

For the 2010-2011 AGT results, the data that met these criteria were:

- **Prior CST scores**
  - CST scale scores from prior year tests are used in the model.
    By including prior test scores, we can determine if students with higher or lower achievement were more likely to gain points on the test. AGT accounts for the relative difficulty of gaining points on the test and takes that into account when making predictions for your students.

    For example, if we found that on average across the LAUSD, students scoring very high on prior tests tended to gain very few points, the predicted outcome of high achieving students would be adjusted to take this into account. In this way, AGT can fairly compare the growth of students from across the achievement spectrum.
- **Grade Level**
  - There are a few ways in which grade level may play a role in student growth. AGT's method of comparing students to their observationally similar peers accounts for these issues to create a fair comparison of learning. For example:

    - Across the entire state of California, the average CST score in ELA changes significantly from grade to grade
      - 2nd Grade – 360
      - 3rd Grade – 345
      - 4th Grade – 372

    Using traditional achievement measures, we would come to the conclusion that 4th grade ELA teachers were much more effective than 3rd grade ELA teachers.

    Rather than achievement, AGT uses academic growth of students to evaluate teacher effectiveness. However, if we used a simple growth model, we might still come to an unfair comparison. Due to the relative difficulty of test standards at 3rd grade and the relative ease of test

standards at 4th grade, California's 3rd grade teachers would be at a disadvantage.

In order to provide a fair comparison of teachers, student growth is only compared to students who are taking the same test sequence from one year to the next. For example, 3rd grade growth predictions are based on the growth of other 3rd graders. 4th grade growth predictions are based on the growth of other 4th graders. By making predictions based on like students, differences in test scale no longer bias the results of teachers at certain grade levels.

- Students at different grade levels may grow at different rates.
  For example, even after controlling for prior test performance, we may find that 7th graders taking Algebra I may grow at a different rate than 9th graders taking Algebra I. By controlling for grade level of students, we create fairer comparisons of teacher effectiveness.

- At the high school level, it is more common to have students from multiple grade levels in a single class. For example, your Algebra II class may include 9th, 10th, and 11th graders. AGT uses prior tests from both Math and ELA to predict Algebra II performance. Since your Algebra II students are from different grade levels, the tests they took the prior year may have been very different. For example, one student may have taken 8th grade ELA and 8th grade general math last year. Another student may have taken 10th grade ELA and Geometry last year.

  In order to account for the different test sequences students may have taken, your students are compared against students who took the same test sequence the previous year. Other students' growth from the same prior tests is used to predict the growth of your students. By customizing the predicted growth of your students based on their test sequence, AGT can fairly evaluate the student learning of your classroom or school even if students are from different grade levels.

- **Gender**
  - Male/Female status of students it taken into account with this model.
    This is a good example of a student characteristic that a teacher cannot control. We might find that on 7th grade geometry, girls across the LAUSD tended to grow faster than boys.

    A teacher with a higher portion of boys in the classroom would be at a disadvantage if the AGT model did not control for gender of students in the

classroom. Since AGT does take gender into account, predictions for classrooms and schools serving different proportions of boys and girls are fairly evaluated.

- **Race/Ethnicity**
  - Students are classified into one of the following categories: African American, Asian, Latino, White.

    Teachers and schools do not control the race/ethnicity of students attending their school, so to fairly evaluate the effectiveness of instruction at these schools, we want to control for the race/ethnicity of students.

    Including race/ethnicity in the model does not mean a preconceived lower expectation for certain students. The AGT model determines whether students of different races/ethnicities grew at different rates across the LAUSD for certain subjects and grade levels. We do not believe that race/ethnicity is a causal factor for growth rates. When considering large groups of students, race/ethnicity tends to be correlated with factors that may have a causal relationship with student growth. (e.g. general socio-economic status)

    Including race/ethnicity does not mean that LAUSD has higher or lower expectations for students of different races/ethnicities. We recognize that students have different sets of circumstances and resources that they bring to the classroom.

    By using all the data we have available, we try to capture the most complete picture of the real situations of students. The more complete job AGT can do at controlling for external factors, the more accurate we can be about evaluating the effect of schools and teachers.
- **Low-Income Status (measured by Free/Reduced Price Lunch Status)**
  - Ideally, we would have data of household financial resources for each student. This data could help us predict the likelihood that a student had access to resources such as:
    - Computer at home
    - Number and quality of books at home
    - External help (like paid tutoring services)
    - Parental availability based on multiple jobs

    Data presenting a full continuum of household financial resources is not available for all students in the LAUSD. As a substitute for this ideal data, AGT controls for a student's eligibility in the Free/Reduced Price Lunch program.

- **English Language Learner Status**
  - English proficiency may play a role in student academic growth. AGT measures whether this has an impact by analyzing the growth of students with different

English proficiency categories.

Students in your classroom or school are compared against the average growth of students with similar English proficiency levels.

NOTE: only students who took the standard CST are included in AGT estimates. If a student took an alternative assessment, that student is dropped from the analysis.

- **Special Education Status**
  - In the AGT model, not only are students identified as SPED or Non-SPED, but the severity of SPED classification is also taken into account. The predicted growth for a group of mildly classified SPED students will be different than the predicted growth for severely classified SPED students. These predictions are based on the actual growth of students with the same SPED status across the LAUSD.

    Note: on the "student groups" page, this categorization based on severity is no longer used. In order to decrease the likelihood that you will receive "Insufficient Data or NA" for this category due to less than 11 students, AGT groups all SPED students into a single category for this calculation.

    NOTE: only students who took the standard CST are included in AGT estimates. If a student took an alternative assessment, that student is not included in the analysis.
- **Continuous Enrollment**
  - To be included in the model, a student must be continuously enrolled in your school. There are cases where a non-whole number would be displayed for the number of students included in the analysis.

    For example, you may see a result based on "19.5" students for example. This would indicate that enrollment in your class changed at the semester. In this case, a teacher may have had 20 students in the fall semester and 19 students in the spring semester.
    Once the LAUSD has consistent attendance data, the AGT model will be able to move to a "dosage" version, where more accurate student-teacher assignments are possible.
- **Homelessness**
  - AGT tests whether there is a detectable difference in growth rate between homeless and non-homeless students. If there is a difference, this is accounted for the in predicted outcome for students.

How is AGT calculated when there is no prior test for a certain content area in the previous year?

AGT results are based on comparing the actual test scores of your students to their predicted scores. These predictions are based on the average performance of similar students across the LAUSD. The AGT model analyzes test results of students across the LAUSD to determine which factors that have the best predictive power. Prior test performance consistently provides one of the best predictors of how students will achieve at the end of the year. By controlling for prior performance in the model, we can fairly compare the growth of your students to those of students with similar prior performance across the district.

In order to use as much data as possible to make the best predictions for the effect of prior knowledge on academic growth of students, the AGT model uses prior scores from multiple subjects.

For example, to predict Math performance at the end of 5th grade, the prior tests from these students' 4th grade ELA and Math performance are taken into account. Although the 4th grade Math test is better aligned to the content of the 5th grade Math test, using the 4th grade ELA test provides extra information that improves the predictions.

In each of the AGT estimates calculated in the LAUSD, multiple subjects are used to make predictions. The table below lists the subjects for which AGT estimates are provided (along the left side) and indicates which prior test subjects are used to make predictions for those AGT estimates (along the top).

		Prior tests used as predictors			
		Math	ELA	Science	Social Studies
AGT Estimates provided	ELA 3-8	X	X		
	ELA 9	X	X	X	X
	ELA 10	X	X	X	
	ELA 11	X	X	X	X
	Math 3-8	X	X		
	Algebra I	X	X	X	X
	Geometry	X	X	X	X
	Algebra II	X	X		
	Science Gr. 5	X	X		
	Science Gr. 8	X	X		
	Integrated Science	X	X	X	X
	Biology	X	X	X	X
	Chemistry	X	X	X	
	Physics	X	X	X	
	Social Science Gr. 8	X	X		

	World History	X	X	X	X
	U.S. History	X	X	X	X

By using multiple subjects to predict performance, AGT is able to provide estimates of teacher effectiveness for subjects even when a subject-specific prior test is not available.

For example, 5th grade Science does not have a 4th grade science available to measure prior science knowledge of students. In place of this content-specific test, each student's 4th grade Math and 4th grade ELA tests are used to predict their 5th grade Science performance. AGT analyzes the 5th grade Science performance of demographically similar students with the same 4th grade Math and ELA scores to calculated predicted outcomes for your students.

The Value-Added Research Center uses advanced statistical techniques to determine if using non-content specific prior tests produces reliable predictions for students in all subjects and grades. The models used to calculate estimates for subjects reported in the LAUSD have all passed a series of rigorous tests to ensure the reliability of estimates.

Why do I have "insufficient Data or NA" for a result?

There are several reasons why a result may not be reported.

- **Too few students to calculate a result**
  If there are fewer than 11 students with valid test data, the AGT model does not report a result for this group of students. This is due to several reasons, including privacy concerns for identifying individual students, and very larger confidence intervals due to less data to evaluate the growth of small student groups.
  For these reasons, the AGT model reports "Insufficient Data or NA" rather than producing an unreliable estimate.
  You may see fewer students reported in your result than you expected. For reasons why individual student results may not be included, please see the next FAQ below.
- **Too few students in the contrasting category to calculate a result**
  In the case of the "student subgroups" page, our AGT analysis determines if you were differentially effective at growing particular students based on certain characteristics. In the school-level and grade-level estimates, AGT measures your impact on student academic growth by comparing your students to similar students across the district.
  On the subgroups page, we categorize your students into particular groupings and then determine if you were more or less effective with one subgroup of students compared to the other.

  For example, you might have 15 ELL and 15 Non-ELL students in your class. Your overall AGT might be 4.0 for all students in Math. On the subgroups page, we would report your result for ELL and Non-ELL separately. Perhaps you have instructional strategies that are very effective with your ELL population, resulting in a 4.6 result for ELL students. Your result for Non-ELL students is still above the district average with an

estimate of 3.4.

In the above case, you had enough students (more than 11) to calculate a result for both the ELL and Non-ELL group.

Let's take another example where a different teacher has 28 ELL and 2 Non-ELL students in her class. In this case, we can still provide an overall AGT estimate for all her students, but when we try to calculate if she is differently effective at ELL and Non-ELL students, we don't have enough students in the Non-ELL category to determine if there is a true difference in growth. Since we cannot calculate a confident estimate of this teacher's effectiveness with Non-ELL students, we cannot report out if she is differently effective with her ELL and Non-ELL students. In this case, her ELL population would be 28 students, but her result would be "Insufficient Data or NA" for her ELL category on the student groups page.

- **A grade level or subject was not taught during the given time period**
  Since AGT is reported for prior year as well as a three year average; there are cases where we have results for one of these columns but not the other.
  For example, at the school level, a school that historically served 6-8th grade just switched to serving 7th-8th grade. In this case, the 6th grade previous year result would be "Insufficient Data or NA". There would still be a three year average for 6th grade that would reflect the two years that the school did have 6th graders.
  At the teacher level, teachers that have changed content areas or grade levels during the previous three years will also have "Insufficient Data or NA" for particular results for the same reason.
- **Only one teacher taught that grade or subject during the given period**
  A teacher's teacher-level results will only be viewable to that teacher and their principal. To respect this privacy and encourage use of AGT as a tool for improvement, we want to avoid publicly reporting a result that only reflects the teaching of a single teacher. School-level reports will have results for grade-level teams, but we suppress data where LAUSD's student information system shows only a single teacher of record for a particular grade or subject at your school.
  For example, if there is only one 7th grade ELA teacher at your building, the 7th grade ELA result will be "Insufficient Data or NA" even if there are more than 11 students in that group. This AGT estimate for 7th grade ELA will be visible on the teacher's teacher-level report.
- **No district-wide differential effect between student groups**
  On the student groups page, you may find that you have "Insufficient Data or NA" even if you have more than 11 students in both groups.
  For example, you may know that you have 40 ELL students and 50 Non-ELL students in

ELA, but you still have "Insufficient Data or NA" for both results on the student groups page.

In this case, a result is not shown because our analysis did not detect that schools were differentially effective at these student groups. There may be a district-wide growth gap between ELL and Non-ELL students, but this gap was consistent across schools. Since we already take into account this district-wide growth gap in all our results, the lack of variance between schools means we cannot reliably calculate whether the growth gap is larger or smaller at any given school. When we do not have sufficient information to make reliable estimates with confidence intervals, we suppress these results and display "Insufficient Data or NA".

- **Inconsistent prior tests used**
  On the student groups page, there are particular subjects and grades where "By Prior Student Achievement Level" is suppressed. This can be due to the different class sequence students take, especially in high school.

  "By Prior Student Achievement Level" is only reported for ELA 3rd-8th grade and Math 3rd-8th grade. In other subjects and grades, it is not clear how students should be classified based on prior achievement status, so the estimates are not reported on the student groups page.

  For example, across the LAUSD, Chemistry students may be 9th graders, 10th graders, 11th graders, or 12th graders. Although there is a sequence most students take, we include all students in the AGT calculation for Chemistry regardless of grade or prior class sequence. Since the AGT model uses prior test score to predict Chemistry performance, the classes each student took last year need to be taken into account. In order to be as fair as possible, growth of students at your school are compared to students who took the same class sequence across the district.

  For example, a particular 10th grader taking Chemistry at your high school may have taken 9th grade ELA, Algebra I, and Integrated Science during the previous year. An 11th grade student in the same Chemistry class may have taken 10th grade ELA, Algebra II, and Biology last year. These students have different predicted scores at the end of the year not only due to potential demographic differences, but also due to their different class sequences. In order to be as fair as possible, each individual student's predicted score is based on students who took the same class sequence from the previous year. This method produces the fairest comparison of student growth for your overall Chemistry result, but we run into problems when we try to provide results on the student groups page under "By Prior Student Achievement Level".

  In our Chemistry class, our two students took very different classes before they entered

Chemistry. Does being "Proficient/Advanced" in Algebra I for our first student represent the same level of knowledge as "Proficient/Advanced" in Algebra II for our second student? Does being "Basic" in Integrated Science for our first student represent the same level of knowledge as "Basic" in Biology? Since the "By Prior Student Achievement Level" categories are based on the starting achievement of students before they took Chemistry, it isn't clear how students who took different class sequences should be categorized. In these cases, the AGT estimates show "Insufficient Data or NA" for the "By Prior Student Achievement Level" results.

How does the color-coding system work?

The color or results is based on the statistical confidence of our estimates. The number inside the colored bubble represents the best estimate of your effect on the academic growth of your students (AGT). Average growth for similar students across the LAUSD is set to 3.0

The black line underneath your colored bubble is a 95% confidence interval around our best estimate of your effectiveness. We are 95% confident that your true AGT falls on this line. Because we have limited data, we report this confidence interval to prevent over interpretation based on too few results.

If the confidence interval crosses "3.0" on the 1-5 scale, the color of your estimate will be gray. This indicates that based on all available data, we cannot detect a result different than average. Our best estimate will still be reported in the gray bubble, but we cannot be confident if your true result is above or below average.

If the entire confidence interval is above "3.0", the color of your estimate will be green. This indicates that your effectiveness is significantly higher than average.
Green estimates indicate an area of success that should be celebrated at your school.

If the entire confidence interval is above "4.0", the color of your estimate will be blue. This indicates that your effectiveness is significantly higher than 4.0
Blue estimates are quite rare and indicate very high growth of students relative to their observationally similar peers.

If the entire confidence interval is below "3.0", the color of your estimate will be yellow. If the entire confidence interval is below "2.0", the color will be red.
Yellow and red estimates indicate areas for improvement. What can we learn from teachers and teams with green or blue estimates that we might be able to replicate in these classrooms?

NOTE: Since the black line representing the confidence interval is the basis for the color-coding, you may find numerical estimates that do not seem to make sense with the color-coding at first glance.

For example, you may find a green result of 3.8 and a gray result of 4.3. In this case, the black

line representing the 95% confidence interval for the green result will be shorter. Although our best guess for the gray estimate is higher than our best guess for the green estimate, the larger confidence interval means we are less sure of our gray estimate.

- Why are there fewer students counted…
  To be included in the analysis, a student needed to meet several criteria. If a student does not meet one of these criteria, this student is not included in the analysis.
  - Continuously enrolled in the same school from the date in October when we conduct our official census of students through the date in May when students take their California Standards Tests
  - Enrolled in a course linked to a CST and obtained a valid CST score for that course
    - Example: If a student's only math class is Algebra II, but this student takes the Geometry CST at the end of the year, this student is not included in the Algebra II or Geometry AGT estimate.
  - Has proper prior test coverage
    - For a list of prior tests used for AGT estimates, see FAQ "How is AGT calculated when there is no prior test for a certain content area in the previous school year?"
  - Assigned to a maximum of two teachers per course
    - Teacher-student assignments are divided by semester. If a student is assigned to more than two teachers for a course, the student is not included in the analysis.
  - Took the regular CST test (rather than an alternative assessment)

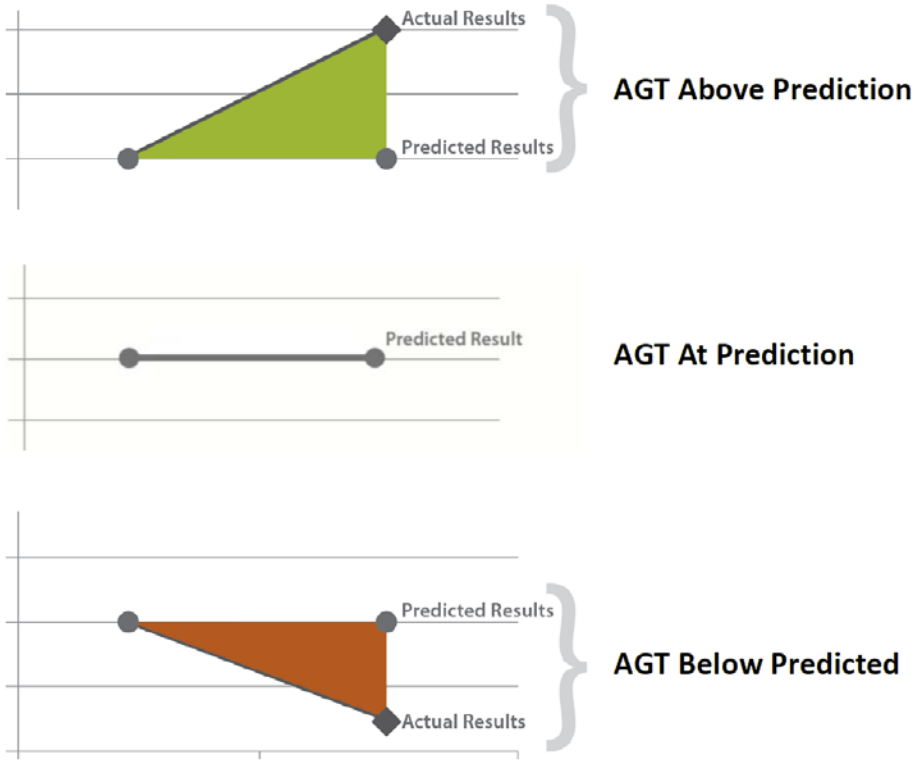
How are AGT results developed?

The basic steps are as follows:

- **Step 1 – Discard results for students where information is insufficient:** For some students it is not possible or appropriate to use an AGT result. This includes students who have not spent enough time in an individual teacher's classroom due to mobility (e.g., the student moves after only 20 days in a teacher's classroom), as well as students for whom there is not an appropriate prior year score to use for predicting purposes. For a list of reasons students may have not been included in your AGT estimate, please see the FAQ "Why are there fewer students counted in our results than the number of students we tested and/or taught?"

- **Step 2 – Predict student learning results:** Using prior achievement and other student factors related to learning outcomes (e.g., free or reduced priced lunch status, special education status), value-added models generate a prediction of each student's learning results for a given assessment (e.g., California Standards Test math results). This predicted outcome is based on the average growth of similar students across LAUSD.

- **Step 3 – Compare predicted results to actual results:** By comparing a group of students' actual results to their predicted results, one can calculate an AGT (value-added) estimate for the teacher or school serving those students.



- **Step 4 – Create the overall AGT result for each entity (e.g., school, grade-level team):** Individual student estimates can be aggregated for an overall AGT result for teachers, grade-level teams, schools or specific groups of students (e.g., English Language Learners in a school). If your group of students grew more than predicted, your AGT estimate will be above 3. If your group of students grew less than predicted, your AGT estimate will be below 3.
- **Step 5 – Discard results for entities where information is insufficient:** In LAUSD, we are only sharing results where there are at least 10 individual student estimates in the result. This both protects the anonymity of students and increases the likelihood we can have statistical confidence that results are accurate. For a complete list of why you may have "Insufficient Data or NA", please see FAQ "Why do I have "Insufficient Data or NA" for a result?
- **Step 6 – Report results:** We are reporting a variety of results at the school level for the public, and we are confidentially reporting individual teacher results.

Why are there fewer students counted in our results than the number of students we tested and/or taught?

To be included in the analysis, a student needed to meet several criteria, such as having been continuously enrolled in the same school from the date in October when we conduct our official census of students through the date in May when students take their California Standards Tests, having prior year test data, and having a complete demographic profile. For 9th grade, the student must have been a first time 9th grade student (and not a student repeating 9th grade) to be part of the analysis. AGT results are suppressed if there are less than 10 students in a particular grade or group.

Which factors has LAUSD incorporated into its AGT model?

LAUSD, with input from stakeholders and a Technical Advisory Group that includes national and regional experts on these methods, has decided to incorporate and 'control for' a variety of factors in our AGT model.

Individual Student Control Variables	Classroom Average Control Variables

• **Prior year CST achievement**
• **Ethnicity**
• **Gender**
• **Free or reduced priced lunch status**
• **Special Education status**
  – **Mild (SLDs and SLIs)**
  – **Severe (All others)**
• **Homelessness**
• **ELL status (multiple categories separated out)**
• **Continuous enrollment status (meets the enrollment standard to be included in the school's CST results – continuously enrolled from October to test day)**
	• Average prior year ELA achievement
	• Average ethnicity
	• Average gender
	• Average free or reduced priced lunch status
	• Average Special Education status
	  – Mild (SLDs and SLIs)
	  – Severe (All others)
	• Average homelessness
	• Average ELL status (multiple categories separated out)

How does the model actually account for all of these control factors? Is there a set percentage for each factor?

Control factors are calculated empirically, based on the actual test data of students across LAUSD. There is not a pre-determined percentage for each factor, since we base these adjustments on the performance of LAUSD students for each grade level and time period.

For example, based on analyzing actual test scores, we might find that across LAUSD, SPED students grew more slowly than non-SPED students during 4th grade for the 2009–2010 school year.

VARC would calculate this difference in growth and use it to make predictions about how SPED and non-SPED students would score at any given school during this same grade and time period.

If a school in LAUSD served a disproportionally high number of SPED students compared to the district-wide percentage, this school would receive a positive adjustment to their predicted output based on the difference we observed comparing the growth of SPED to non-SPED students across LAUSD. Since we observed that SPED students grew more slowly than non-SPED students, we do not want to unfairly penalize this school for serving a high number of SPED students. By making a prediction for each school tailored to the students it serves, schools are placed on a level playing field so results can be fairly compared across the district regardless of student population.

What about classrooms or schools with many high achievers? Is there really any room for growth? What about classrooms where all of the students have perfect or near perfect scores?

Unlike API where credit is given for students based on proficiency category (Advanced, Proficient, Basic, Below Basic, Far Below Basic), AGT is based on the actual scale score on standardized tests. Even if students are already in the Advanced category, they can still have room to grow on the test unless they scored the maximum possible score on the test. Since AGT is based on growth from one year to the next, a student would have to obtain the highest possible score two years successively to truly have no room to grow.

As a general feature of the AGT model, prior performance on the CST is accounted for in the predictions. In general, VARC usually finds that higher achieving students tend to gain fewer points on the test. By analyzing the actual growth results of students from all achievement levels in LAUSD, VARC is able to make growth predictions for students at any starting point along the test scale. By customizing predictions for students based on their starting point, we can fairly compare classrooms with low-achieving students or high-achieving students.

What about classrooms or schools where there are many far below basic students who may just be guessing when they take the CST?

Just as is the case above, where VARC takes into account the starting point for the highest achievers in LAUSD, we also take into account the starting point for the lowest achievers in LAUSD. While there may be many Far Below Basic students in a school, there is variation in actual scale scores for those students. By using scale scores rather than raw scores, our calculations already account for guessing on the CST.

By analyzing actual outcomes of students with the same low achievement level on the CST, we make predictions for these Far Below Basic students compared to their peers.

Why are we using a complex statistical tool? Why not use something simple like comparing students' overall gains from one year to the next?

As one of our research partners tells us, "Simpler is better, unless it's wrong." Because student assessments in one year do not test the same things as the assessments the next year, it is not appropriate to do a simple comparison of student results. We also know that factors external to our schools and classrooms impact student learning rates, and AGT allows us to incorporate those factors. Finally, no assessment is perfect, and AGT incorporates these errors in measurement.

What are some limitations of AGT?

AGT models represent a narrow way to talk about the effectiveness of educators and schools. We must consider these results within the context of other metrics—both qualitative and quantitative. Like all measures related to student learning, this approach is subject to limitations, which is why (a) it is important to look at multiple indicators and results and (b) this approach should be used as one amongst multiple measures. Further, the power of this analysis is lost if it is not linked to ways to improve practice or to learn from those with clear positive effects on student learning.

What are the benefits of AGT?

When assessing the effectiveness of educators or schools, AGT is useful in that it tracks the progress or growth of individual students over time, rather than the percentage of students that meet an absolute target or standard. This allows for true 'apples to apples' comparisons of effectiveness. Indeed, AGT will allow us to identify, study and share the practices of schools and teachers who are achieving remarkable results. These remarkable results are based upon a common standardized measure of growth, so that when we compare growth in one school to another, we know that we are using a term with one definition. It will not be that School A has a different understanding/definition of growth than School B.

Does this system apply to norm-referenced measures only? If not, how does it apply when looking at criterion referenced tests?

AGT methodology can be applied to student data from both norm-referenced or criterion referenced tests. In LAUSD, AGT is based on the California Standards Tests (CSTs) for grades 3–8 in general math and grades 3–11 in English Language Arts. Algebra I, Geometry, Algebra II, Science Grade 5, Science Grade 8, Integrated Science, Biology, Chemistry, Physics, Grade 8 Social Science, World History and U.S. History are also included in the results. Grade 5 Science and Grade 8 Social Science are not included in the teacher-level reports due to multi-grade/multi-year curriculum. ELA Grade 10 and ELA Grade 11 are not included in the teacher-level reports due to not meeting LAUSD's criteria for inclusion (reliability, stability, predictive power, explained variance, teacher/student coverage).

The actual AGT estimates for schools and teachers are norm-referenced. An AGT estimate compares the academic growth of students in a particular school or classroom to the growth of observationally similar students across LAUSD. Since properties of the CST are not consistent across grades and years, creating a criterion referenced or benchmarked standard for AGT is not technically possible. For each grade level and time period, AGT compares a school or teacher to the rest of LAUSD for that same grade level and time period.

How does the model account for changes to curriculum?

The goal of AGT is to remove the effects of non-school factors from student growth in order to give schools and teachers a fair comparison of their impact on the growth of students. After these non-school factors are accounted for, we are left with the effect of the school or classroom. Factors related to the school (like an individual school or teacher's curriculum) are not removed from the calculation, so the AGT result includes the effect of curriculum on student growth.

However, if there was a district-wide change in curriculum, all schools will experience this change. Since AGT results compare the growth of a school to the growth of students across LAUSD, a change in curriculum that affects all students will not affect AGT estimates.

For example, if a new district-wide curriculum improves student learning in ELA across the whole district, then the predicted growth of all students increases accordingly based on the real growth of students that year. Since schools are compared to the district as a whole, district-wide changes in curriculum will not break the AGT model. If one school is more effective at implementing this new curriculum than the district average (and their students learn more), that WILL show up as a higher AGT.

How does the model account for class size?

Class size is not factored into the model. Control factors like English Language Learner Status and Free/Reduced Price Lunch Status are included in the AGT model to remove the effect of these non-school factors on student growth. The prior knowledge and demographic characteristics of students being served by schools are not something schools can affect. AGT controls for these student characteristics in order to create fair predictions for schools. By eliminating these non-school factors, AGT can fairly measure the growth of students at schools serving different student populations. In the case of class size, this is a factor that the district or school can change through policy decisions and resource allocation. By not controlling for class size, schools can determine if strategies involving different class sizes are still producing student learning.

For example, a school may want to attempt a strategy where high achieving students are placed in larger class sizes with the school's teachers who are most effective at growing high achieving students. This school could then place low achieving students in smaller class sizes (less than 17) with teachers who are most effective at growing low achieving students.

Since AGT does not control for class size, AGT results would be accurately measuring whether students at this school were growing above the district average for similar students in both of these ability-grouped classrooms.

If class size *was* controlled for in the model, the district-wide effect for class size could be reflecting several different effects. It might reflect less student learning occurring district-wide in larger classes, but it would potentially be reflecting the opposite effect in the strategy attempted above. If highly effective teachers were strategically assigned more students than less effective teachers, this would confuse what the model was truly measuring in a district-wide class size effect.

Although larger class sizes certainly mean more grading and potential classroom management challenges, current research suggests that in class sizes above 17 students there is not a detectable difference in student learning based on higher class size. Since schools may choose to alter class sizes based on the effectiveness of teachers (based on AGT), it is not clear

whether the model would be measuring a true difference in learning due to class size or would be measuring the result of class size assignment strategies of schools based on teacher strengths.

How does the model account for student attendance?

Attendance is not included in the AGT model. LAUSD's data systems do not currently have enough information to produce a reliable AGT dosage model based on attendance data. LAUSD is working to improve the data system so future analysis years will take this data into account.

When reliable data is available, the AGT model will account for prior attendance rather than attendance for the year being measured by AGT.

Why would we account for prior attendance rather than current attendance?
As with other variables controlled for in the AGT model, we want to remove the effect of factors beyond the school's or teacher's control. When a student is assigned to a teacher, there may be strategies the teacher can use to get a previously low-attending student to come to class more often (strong parent interactions, engaging classroom activities, encouraging classroom community building, etc.) On the other hand, this particular teacher does not have any control over the attendance history of that student from previous years (before they entered this teacher's classroom).

Let's assume that across the district, our analysis found that students with lower attendance grew less than students with higher attendance.

In the following example, we will consider that there are two 7th grade math teachers. These two teachers have been assigned two groups of students who have exactly the same characteristics, including a low average attendance rate (75%) from last year.

During the past year, teacher A used strategies in her classroom that resulted in her students having a much higher attendance rate than these students had last year. The average attendance rate for her students was 90% instead of 75%.

In teacher B's classroom, there were no strategies used to increase student attendance, and her students actually had a much lower attendance rate than they had last year. The average attendance rater for her students was 60% instead of 75%.

With the situation above, let's explore the results we would get with an AGT model that controlled for current attendance vs. prior attendance.

In an AGT model, student growth for a given teacher's classroom is compared to the predicted growth of those students. This prediction is based on the average growth of similar students (accounting for prior test performance and student characteristics) across LAUSD during the same time period.

If AGT controlled for current attendance in the above scenario:
Teacher A's prediction for growth of her students would be based on students across the district who also had 90% attendance rate. Since we assumed the probable outcome that students with higher attendance rates learn more during the school year, teacher A's predicted growth would be higher since her students came to school more often than they did in the past.

In the case of teacher B, her prediction for the growth of her students would be based on students across the district that also had 60% attendance rate. Teacher B's predicted growth would be lower since her students came to the school less often than they did in the past.

In this case, teacher A (who managed to get incredible attendance for her students) is held to a higher prediction than teacher B (who ended up with much lower attendance for her students). This would set up an incentive structure where teachers would be more likely to get high AGT if their students had lower attendance. We want higher attendance for our students, so AGT should not be set up this way.

On the other hand, if AGT controlled for prior attendance in the above scenario:
Teacher A and teacher B would have exactly the same predicted growth for their students. Their predicted performance would be based on the growth of similar students across the district that had the same 75% attendance rate for the previous year.

Teacher A, with her higher attendance rate, would be more likely to have a higher AGT result with that increased instruction time in the classroom. Teacher B, with her low attendance rate, would be more likely to have a lower AGT result with less instructional time in the classroom.

This matches our intended incentive structure, where teacher who use strategies to get historically low attending students to come to school get recognition for this results in their student growth.

How does the model account for student mobility and students moving from school to school?

A preferred method for accounting for the high mobility rates of our students is to use what is known as a 'dosage' model, where student AGT results are ascribed based on the percentage of time the student spent in a teacher's classroom. Our current attendance data, though, does not support a full dosage model. This may be possible for future releases.

For the school reports released in Fall 2011, the students included in the analysis are those that have been continuously enrolled in school from the date in October when we conduct our official census of LAUSD students through the date in May when students take their California Standards Tests.

How will you ensure that students are accurately matched to teachers? What about teachers providing support to small groups of students? What about team teaching, itinerate teachers and special education teachers who work with students in several classrooms?

Student-teacher assignments are pulled from the LAUSD SIS (student information system). In order to appear on the list, a student must be continuously enrolled in your class and have a result for the regular CST associated with the course (students taking an alternative assessment are not on this list) and have valid prior year test scores (CST scores from the 2009–2010 school year.

For several reasons (e.g., team teaching, teacher leaves of absence), the teacher of record may not be the only or best teacher to associate with a student's results. Indeed, in many scenarios, two or more educators may be appropriately connected with a student's outcomes. Studying this matter involves both a careful analysis of our Human Resources data as well as a roster verification study. In a roster verification study, teachers and administrators go through a process whereby they review and identify inaccuracies in the way students have been linked to teachers. The roster verification process also involves connecting students with intervention, itinerant, and resource specialist teachers that have worked with them.

In Spring 2011, we conducted a mini-pilot roster verification study to inform plans for future analysis so that we can address this matter in advance of utilizing these results in a formal teacher performance review. With the results of this study and through analyzing best practices in this area, we exploring roster verification options for 2011–2012 that would be incorporated into Fall 2012 AGT results.

For this release, you will be able to view a student roster file that identifies which students were used in your AGT calculations. This roster file is available through the same portal you used to download this PDF report.

In the long term, we are looking at ways to incorporate this kind of information into the student information system, so that we are tracking interventions, team teaching and the like on an ongoing basis.

In the short term, Fall 2011 teacher-level results will be accompanied by a document listing which students were included in each teacher's results.

If you notice an extra student on your roster that you did not teach or are missing a student who should have been included, please contact SIS staff in your school's main office to determine if the student in question was enrolled in your class during this period of time. Any student

discrepancies would then be followed up by the school's office staff and SIS support to clarify the situation.

Do these results incorporate special education students?

These results incorporate those special education students who take the California Standards Test, and not those who take alternative state assessments. We are examining the appropriateness of these alternative state assessments for future AGT analyses.

Will student-level AGT results be available?

AGT analyses aggregate student-level results for a classroom, a group of classrooms (e.g., a teacher, a grade-level team, a school) or a specific group of students (e.g., English Language Learners within a school), and not for individual students. Another key point here is what is being measured. AGT is measuring the degree of impact of the school, grade-level team or teacher upon the student's learning.

How will the Common Core Standards and the plans for new statewide assessments in 2014 impact AGT?

AGT estimates will still be possible with a change in assessments. AGT uses prior test performance as a predictor of current performance, but the predictions do not have to be based on the same assessment. Currently in LAUSD, a student's ELA and math CST scores are used as predictors for ELA performance. The same can be said for Math performance. AGT estimates are test agnostic, and LAUSD's model of using multiple pre-tests as predictors are at the cutting edge of value-added analysis. Even if the entire test as a whole changes, we can still use scores to predict outcomes.

As a real world example, Wisconsin's standardized test changed from the WKCE to the WKCE-CRT while VARC was producing value-added calculations in the state. By analyzing performance of students from one test to the next, Value-Added Estimates were still produced during this time period.

ACADEMIC GROWTH OVER TIME AND THE LAUSD SUPPORTING ALL EMPLOYEES STRATEGY:

Why did LAUSD decide to revamp its current evaluation procedure?

The April 28, 2009 Board motion (Quality Leadership and Teaching to Ensure a World Class Education For All) brought forward by Board member Yolie Flores, Board President Monica Garcia, and Board member Dr. Richard Vladovic charged Superintendent Ramon Cortines to create a task force to develop recommendations for enhancing the ways in which the district ensures that the most effective teachers, administrators and support personnel work with our students every day.

The Teacher Effectiveness Task Force (TETF), chaired by Dr. Ted Mitchell, focused on employee evaluation, support mechanisms, tenure, compensation and legislation. This group was charged with reviewing current practices; studying relevant research; and, developing recommendations and a plan for action to achieve meaningful changes to the Education Code, state rules and regulations, and district policies and practices related to its focus areas.

The Los Angeles Unified School District recognizes that the most important task we face is to ensure that every classroom is led by an effective teacher, and that every school is led by an outstanding leader, who is surrounded by a team of excellent support personnel.

National research supports this and so does our common experience. The recommendations of the LAUSD Teacher Effectiveness Task Force, a multi-stakeholder body including parents, students, teachers, school leaders, district leaders, union members and community partners set the course for a new, more comprehensive approach to this paramount issue for LAUSD.

How will the district reach its goals for helping educators become more effective?

We have developed a three phase plan:

- Phase I – Research and Development (SY 2010–2011): During this school year, we will put in place a series of efforts, including input by all stakeholders, to analyze data and current practices and build a new evaluation tool which looks at many measures, not just student test data. This model will also look at ways to provide additional compensation for employees who are meeting the needs of students.

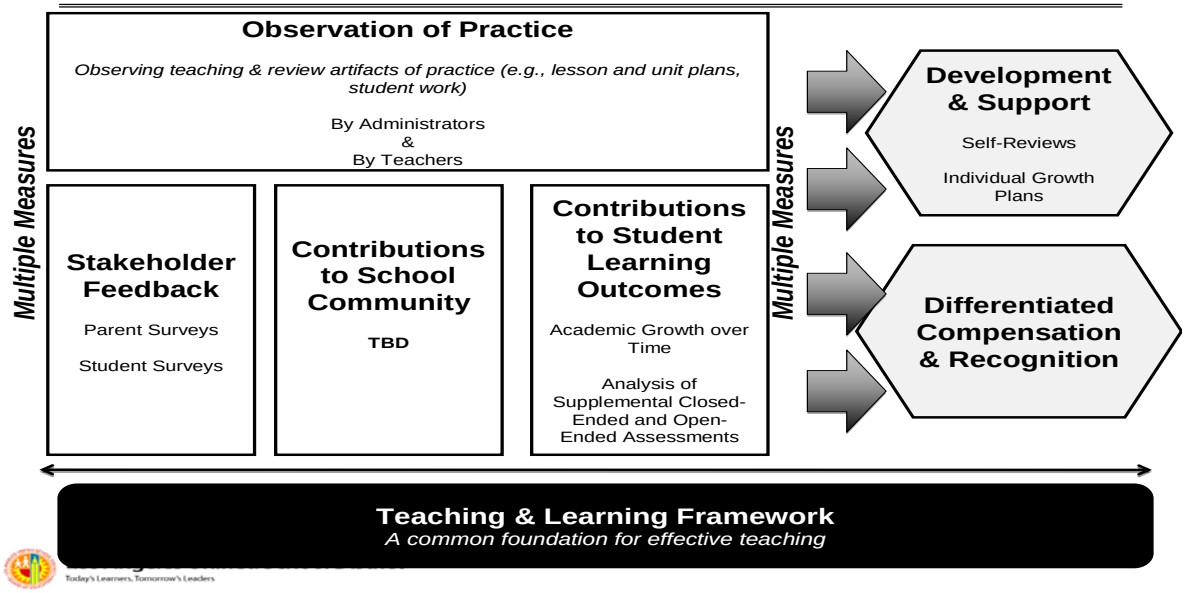
- Phase II – Initial Implementation (SY 2011–2012): Next year, the district will test these new measures with a sample set of schools and identify ways to implement this system across the District.

- Phase III – Scale (SY 2012–2013): The third year of our plan involves district-wide implementation and working on improving quality for all schools.

What will the new evaluation process look like?

The system will consist of multiple measure reviews with aligned professional development and support structures. Performance measures for teachers will include: Observation of Teacher Practice, Contributions to Student Outcomes, Stakeholder Feedback, and Contributions to School Community.

The diagram below outlines these plans. Similarly, we will develop a School Leadership Framework with corresponding measures of effectiveness, aligned Individual Growth Planning, and voluntary opportunities for Differentiated Compensation and Recognition.



Is this going to link in any way to value-added (what LAUSD calls Academic Growth over Time)?

Our plans call for utilizing measures of Academic Growth over Time (AGT) that provide information regarding how teachers and school leaders contribute to student learning outcomes as one of the multiple components in the overall effectiveness rating of an educator (where observation of practice is the majority measure). AGT is a robust statistical analysis that estimates the influence of schools and educators on student growth, while controlling for the non-school factors that also influence student growth.

AGT results will be a fractional measure in the overall effectiveness rating of teachers, where observation is the majority measure. It would be inappropriate to judge a teacher's effectiveness only using student outcomes data. It would be inappropriate not to include student outcomes data in a teacher's evaluation.

How will teachers, administrators, parents/guardians and other stakeholders be involved?

At each stage of the process, we have, and will continue to engage the input of all key stakeholders, including parents, students, teachers, school leaders, district leaders, union members and community partners. Each of these groups has participated in focus groups, providing on-going input, submitting surveys or as members of the task forces.

How can I learn more about the Supporting All Employees Strategy?

All our material can be found at our Supporting All Employees website at http://sae.lausd.net

How can I find out more about the AGT results for my child's entire school?

You can find each school report at agt.lausd.net. In addition, this data is included in the School Report Card that each parent in the District receives. As research points out, the validity of

value-added scores grows stronger when derived from groups of teachers, such as a whole school. We feel this is the appropriate and transparent role for public use of these results.

LAUSD ACADEMIC GROWTH OVER TIME AND THE *LA TIMES* VALUE-ADDED WORK:

How similar is LAUSD's AGT model to what the *LA Times* does in its own value-added analysis?

Both LAUSD and the *LA Times* have used a value-added methodology. LAUSD has incorporated more data into our model and has provided a wider set of results. While we cannot speak to the model development process performed by the *LA Times*, LAUSD has engaged in a rigorous analysis of our data and worked with a broad set of national and regional experts on the development of our model.

Why would some of the LAUSD AGT results be different than the *LA Times* results?

While both models are value-added models, there are several reasons for differences including modeling choices, which results are reported, and measures of precision.