Rewarding Excellence in Instruction and Leadership

REIL Guidebook: 2012-2013

A practitioner's guide to understanding the REIL performance measures, calculating the REIL Score, and determining state performance designations

Employee Group 5

All K-12 teachers in REIL TNG districts

MCESA
Maricopa County Education Service Agency

PLEASE NOTE: The contents of this guidebook only apply to the 2012-13 school year. New guidebooks will be issued for 2013-14 to provide details on how performance will be assessed when all four components of the REIL performance measurement system are fully implemented.

Rewarding Excellence in Instruction and Leadership (REIL)

Rewarding Excellence in Instruction and Leadership (REIL), an initiative of the Maricopa County Education Service Agency, engages five Maricopa County school districts in implementing systemic change aimed at transforming how schools recruit, retain, support, and compensate effective teachers and principals. The ultimate goal is to build the capacity of educators to improve student learning. Rigorous, fair and transparent educator evaluations, targeted professional learning, tools for measuring student success, establishment of multiple career pathways, and sustainable, differential, performance-based compensation are critical elements of REIL's performance-based management system. The five-year initiative, which will culminate in 2014-15, is funded by a $51.5 million Teacher Incentive Fund grant from the U.S Department of Education.

Contents

Overview

Purpose of the Guidebooks

The purpose of the REIL Guidebook is to provide teachers with information on how their performance will be measured in the 2012-13 school year.[1] The guidebook provides details on the various instruments, methods, and processes that will be used to assess the quality of classroom instructional practice and measure student academic growth in your school. It also describes the procedure for calculating the REIL Score and outlines the criteria that will be used to determine teachers' annual state-required performance classifications (i.e., highly effective, effective, developing, ineffective) and REIL performance awards.

Differentiated Guidebooks for REIL Employee Groups

Teachers and principals are classified into eight different employee groups based on the measures and criteria that will be used to assess their performance in 2012-13. Separate guidebooks are provided for each employee group.

REIL Employee Groups in 2012-13

1. K-12 teachers in REIL Alliance districts
2. Preschool teachers in REIL Alliance districts
3. Special Education teachers in REIL Alliance districts
4. Principals in REIL Alliance districts
5. **K-12 teachers in REIL Extend districts** ★
6. Preschool teachers in REIL Extend districts
7. Special Education teachers in REIL Alliance districts
8. Principals in REIL Extend districts

This guidebook is for **Employee Group 5**: All K-12 teachers in the REIL Extend districts. Guidebooks for the other positions are available on the REIL Website.[2]

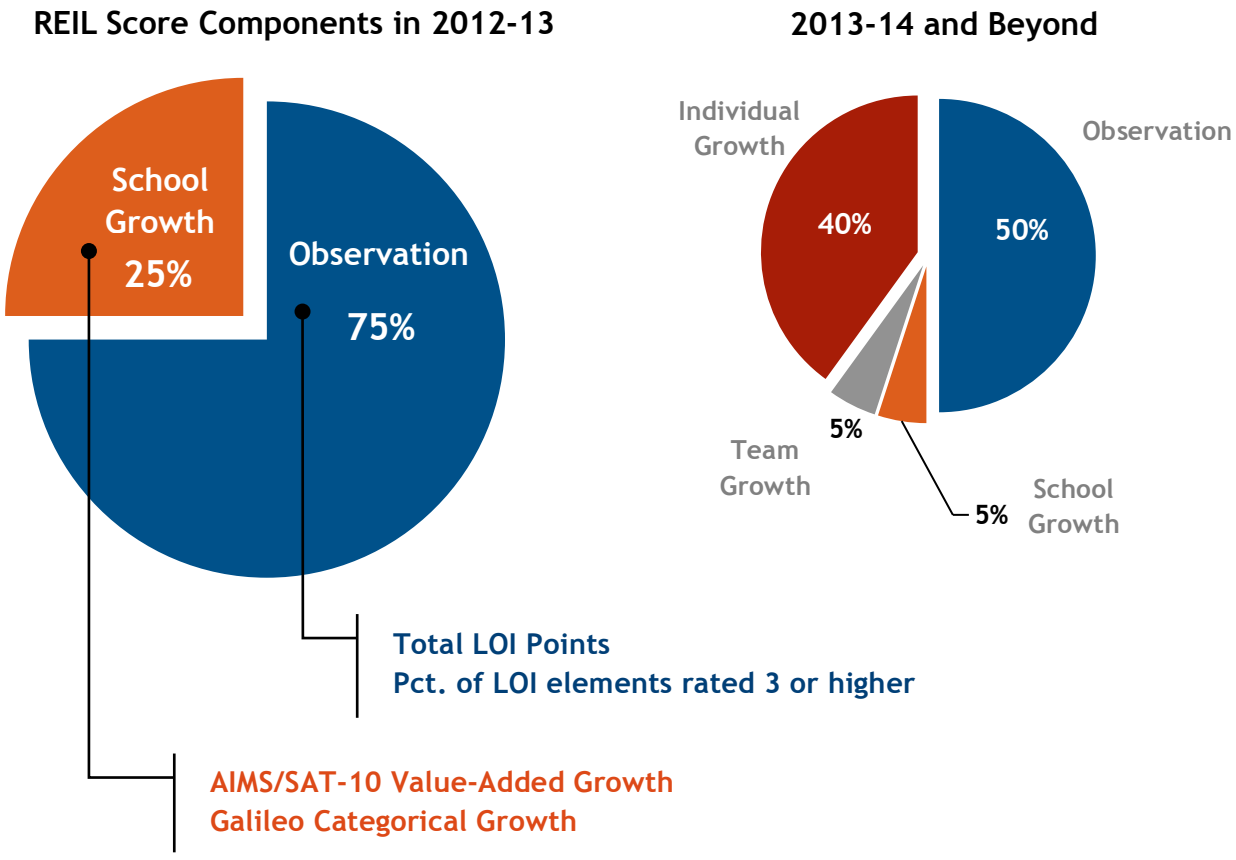
[1] The contents of this guidebook only apply to the 2012-13 school year. New guidebooks will be issued for 2013-14.

[2] REIL Website: http://www.maricopa.gov/schools/service-home.aspx?sid=1

Components of the REIL Performance Measurement System

The REIL performance measurement system will have four major components when it is fully implemented: Individual Growth, Team Growth, School Growth, and Learning Observation. In 2012-13, however, only two of the four components will be used to determine teachers' REIL Scores. Twenty-five percent of teachers' REIL scores will be based on the school growth component, which includes school-wide value-added measures in math and reading from the Arizona state assessments (AIMS and SAT-10) and school-wide growth measures in math, reading, science, and writing from the Galileo K-12 assessments. Teachers' cumulative results of the two observation cycles will account for the remaining 75 percent of the REIL Score.

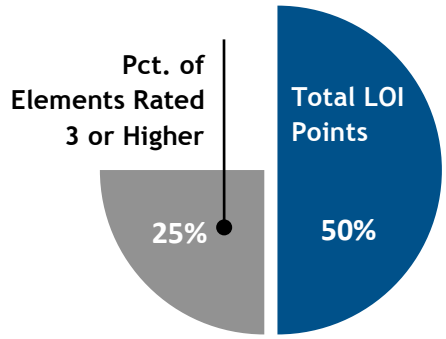
Figure 1. Components of the REIL Performance Measurement System



Observation Component

Classroom observation results account for 75 percent of the REIL Score in 2012-13. This section provides an overview of the observational tool and process used by REIL districts to measure the quality of teachers' instructional practice. It also describes how the results of the two observations that you will complete over the 2012-13 school year will be used to inform your REIL Score. More detailed information on the REIL observation tools and process is available on the REIL Website.[3]

Figure 2. LOI measures in 2012-13



The Learning Observation Instrument (LOI)

Classroom instructional practice in REIL districts is assessed using an observational tool called the **Learning Observation Instrument (LOI).** The LOI is the result of a collaborative effort to create a cross-district observational tool to be used as part of a performance-based evaluation system. It is designed to define effective teaching practices, encourage dialogue about instruction, and support differentiated professional growth.

The LOI is composed of six rubrics that align to the InTASC Model Core Teaching Standards. Five of the six rubrics are implemented consistently across REIL Alliance districts: *Content, Formative Assessment, Instructional Strategies, Learner Engagement, and Learning Community*. The sixth rubric, *Professional Responsibilities*, is scored at the discretion of each district and not used in the REIL Score.

[3] REIL Website: http://www.maricopa.gov/schools/service-home.aspx?sid=1

Each LOI rubric contains several elements (see Table 1). For example, the Formative Assessment Rubric contains three elements: Real Time Assessment, Student Progress and Correct Level of Difficulty. Across the five rubrics, a total of 21 unique elements are assessed in each observation cycle. Evaluators rate each element on a scale from 0 to 5 according to the criteria provided in the rubrics. One element –Task Analysis – is rated twice during each observation cycle (once during the pre-conference and once during the classroom observation), thus teachers receive 22 element ratings in each observation cycle.

Figure 3. Number of element ratings scored per observation cycle

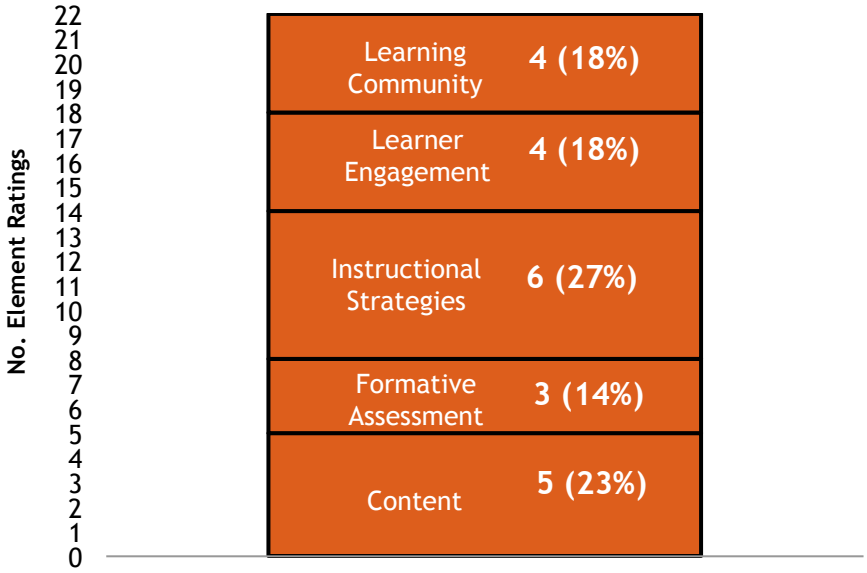


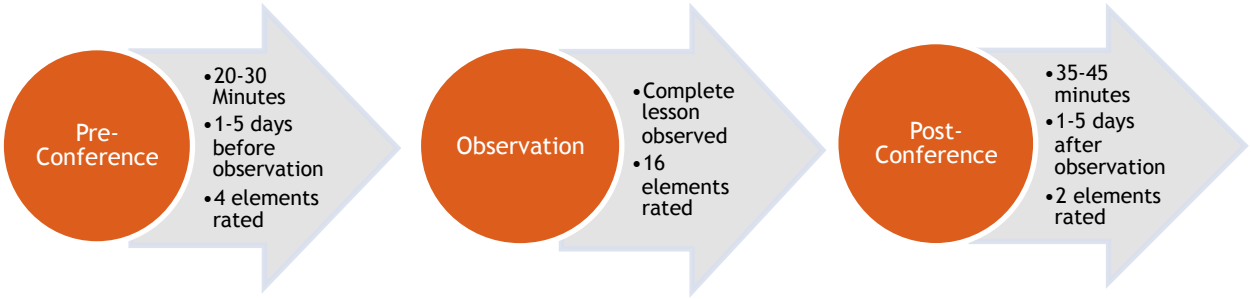
Table 1. LOI Rubrics and Elements

Rubric	Element	Setting		
		Pre Conference	Classroom Observation	Post Conference
Content	Conceptual Understanding		●	
	Task Analysis	●	●	
	Connections to Content		●	
	Content Accessibility	●		
Formative Assessment	Real-Time Assessment	●		
	Student Progress			●
	Correct level of Difficulty	●		
Instructional Strategies	Teacher Role		●	
	Instructional Approach		●	
	Practice/Aligned Activity		●	
	Feedback		●	
	Monitor and Adjust		●	
	Analysis of Instruction			●
Learner Engagement	Student-to-Student Interaction		●	
	Teacher-to-Student Interaction		●	
	Authentic Engagement		●	
	Critical Thinking		●	
Learning Community	Routines and Procedures		●	
	Responsibility for Learning		●	
	Monitoring Student Behavior		●	
	Relationships		●	
Total Element Ratings per Cycle:		**4**	**16**	**2**

The LOI Observation Cycle

Teachers in REIL Extend districts are expected to participate in 2 observation cycles in the 2012-13 school year. The evaluation cycle is an opportunity for teachers, administrators and REIL Peer Evaluators to have on-going communication regarding instruction. Each of the two cycles serves as a formative component to impact teacher effectiveness and student learning. There are two observation cycles per school year which will be facilitated by your site administrator or REIL Peer Evaluator. All cycles collectively contribute to your final REIL score. A single observation cycle, which includes a pre-conference, observation, and post-conference, is completed within ten school days.

Figure 4. The LOI Observation Cycle



Pre-Conference

The pre-conference is an opportunity to discuss the lesson that will be observed as well as your planning process. During this time, you will be scored on the four pre-conference elements of Task Analysis, Content Accessibility, Real-Time Assessment, and Correct Level of Difficulty.

Observation

The observation is an opportunity for you to present your complete lesson discussed in the pre-conference. The evaluator will observe the lesson within three school days. During your observation, the evaluator will script the lesson as well as students' responses and actions. Your lesson will be scored on the 16 observation elements of the LOI.

Post-Conference

The post-conference gives you an opportunity to analyze and reflect on your lesson as well as discuss your students' assessment results. The evaluator may ask questions for clarification, review any data and artifacts you may present, and script the discussion. The evaluator will score you on the two post-conference elements of Student Progress and Analysis of Instruction.

Performance Criteria for the Observation Component in 2012-13

Teachers who complete two observation cycles in 2012-13 will accumulate 44 element ratings (22 element ratings per cycle X 2 cycles). These 44 element ratings are the basis for 75 percent of the REIL score in 2012-13. Fifty percent of the REIL score will be based on the **Total LOI Points**, which is the sum of all 44 element ratings. Twenty-five percent of the REIL score will be based on the percentage of element ratings that are rated a 3 or higher, which is referred to as the **LOI 3+ Rate**.

22	x	2	=	44
Elements ratings per observation		Observation cycles		LOI element ratings in 2012-13

Total LOI Points (50 Percent of the REIL Score)

The Total LOI Points is the sum of the 44 element ratings that teachers receive from their evaluators over the two observation cycles. Evaluators rate each element on a scale from 0 to 5, so the maximum amount of LOI points one can accumulate over two observations is 220 (ratings of 5 on all 44 elements). The minimum amount of LOI points a teacher can earn is 0 (ratings of 0 on all 44 elements).

At the end of the school year, teachers' Total LOI Points will be converted to performance scores on a scale of 1 and 5. Note that all of the REIL performance measures are converted to this common 1 to 5 scale so they can be combined to create the REIL score. Table 2 displays the total number of LOI points necessary to achieve each performance score. For example, teachers who earn 189 or more LOI points over two observations will receive a performance score of 5. Teachers who earn between 0 and 66 points will receive performance scores of 1.

Table 2. Conversion of Total LOI Points Earned to Performance Scores

Performance Score	Total LOI Points Earned (2 Observation Cycles)
5	189-220
4	155-188
3	117-154
2	67-116
1	0 – 66

Pct. of LOI Elements Rated 3 or Higher (25 Percent of the REIL SCore)

Twenty-five percent of the REIL score in 2012-13 will be determined by the percentage of all element ratings received in 2012-13 that are rated a 3 or higher on the LOI's 0 to 5 scale. This measure is referred to as the LOI 3+ Rate and it is designed to reward teachers who demonstrate proficiency across all LOI elements. For instance, a teacher who receives ratings of 3s on all elements will have a higher LOI 3+ Rate than one who receives 1s on some elements and 5s on others, though the two may earn the same number of LOI points.

LOI 3+ Rates will range from 0 to 100 percent, where 0 percent indicates that no element ratings received in 2012-13 were greater than or equal to a 3 and 100 percent indicates that all 44 element ratings were at least a 3.

Teachers' LOI 3+ Rates will be converted to performance scores on the 1 to 5 scale using the criteria in table 3. In order for teachers to receive a score of 5, they must receive ratings of 3 or higher on 100 percent of the LOI elements. In other words, they must demonstrate proficiency on all LOI elements in both observation cycles.

Table 3. Conversion of LOI 3+ Rates to REIL Component Scores

Performance Score	Percent of Elements Rated 3 or Higher
5	100%
4	90% to 99%
3	60% to 89%
2	25% to 59%
1	0% to 24%

School Growth Component

School-wide student growth results make up 25 percent of the REIL Score in 2012-13. This section describes the assessments used to measure students' academic growth in 2012-13 and the statistical techniques used to estimate how much each school contributed to their students' average growth. It also describes how the school growth results from the various assessments will contribute to the REIL Score.

AIMS/SAT-10 Value-Added Measures

Value-added analysis is used to estimate schools' average student growth in reading and math on the Arizona statewide assessments: AIMS/SAT-10. The test scores used to measure growth in each grade are shown in table 4.

What are Value-Added Models?

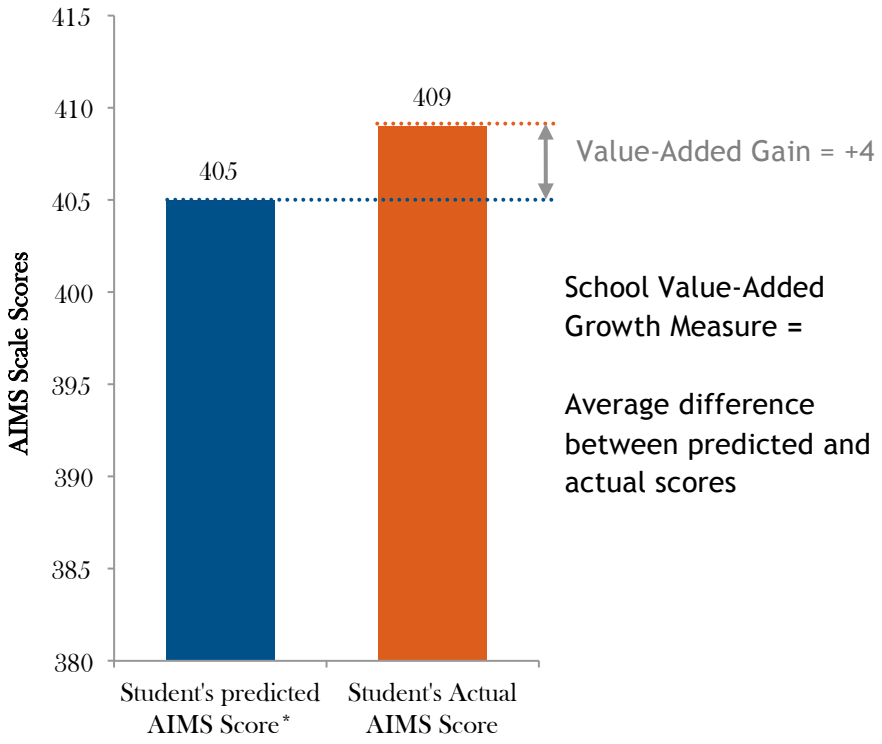
A variety of sophisticated statistical techniques that use one or more years of prior student test scores, as well as other data, to adjust for preexisting differences among students when calculating contributions to student test performance .

-Braun, Chudowsky, & Koenig (2010). *Getting Value out of Value-Added.* National Academies Press.

Value-added analysis uses statistical techniques to predict the test scores that each student is expected to receive on the AIMS/SAT-10 based on what is typical in the state for students in the same grade, with similar prior test scores and individual background characteristics, and enrolled in a school with similar characteristics. Separate scores are predicted for each student in reading and math.

After making these predictions, students' *actual* test scores are compared to their *predicted* test scores in order to determine their value-added achievement gain, which can be positive or negative (see figure 4). Your school's value-added growth measure on the AIMS and SAT-10 assessments is based on the average difference between students' actual and predicted test scores. If students systematically score above their predicted scores ("exceed expectations"), your school will have positive value-added results. Conversely, if students systematically fall short of their predicted scores your school will have negative value-added results.

Figure 4. Calculating Schools' Value-Added Growth Measures



Factors used to Determine Students' Growth Expectations (Predictions)

The REIL value-added model relies on the growth trends of all tested public school students in the state as the basis for its predictions of how students should perform on the AIMS and SAT-10. Students' growth expectations (predictions) for the school-level model are determined by the following factors (covariates):[4]

- Prior achievement in reading and math
- Gifted and Talented status
- Disability/special education status
- Free and reduced meal eligibility status
- English Language Learner (ELL status)
- Fluent English Proficient (FEP status)
- Migrant status
- Homeless status
- Student mobility status (new to the school in non-promotional year)

[4] Additional covariates will be included in the teacher-level value-added model that will be implemented in 2013-14.

Students Included in the School-wide Growth Measures

In order for a student's AIMS or SAT-10 results to contribute to the value-added growth measure of a particular school, the student must have AIMS or SAT-10 test scores from the Spring of the prior school year and have been enrolled in that school for the full academic year. According to the Arizona Department of Education, a student is considered enrolled for the full academic year if he/she is enrolled in the school during the first 10 school days and remains enrolled through the date of the AIMS/SAT-10 test.[5] Schools with fewer than 15 students who meet these criteria will not receive value-added growth measures on AIMS because growth measures based on such small samples are not reliable.

Resources on Value-Added Analysis

More details on the REIL value-added model used for AIMS and SAT-10 are available on the REIL Website.[6] An extensive list of research publications written for practitioners is provided on page 20, which includes a helpful tutorial on value-added measures from the Value-Added Research Center at the University of Wisconsin.

[5] *AZ School Accountability System 2009 Technical Manual: Volume II - Section 4 pg. 13 (2010)*

[6] REIL Website: http://www.maricopa.gov/schools/service-home.aspx?sid=1

Table 4. Assessments used to Measure Growth in 2012-13

	AIMS/SAT-10 Value-Added Measures				Galileo K-12 Growth Measures							
	Reading		Math		Reading		Math		Science		Writing	
	Outcome Measure	Baseline Measure	Outcome Measure	Baseline Measure	Outcome Measure	Baseline Measure	Outcome Measure	Baseline Measure	Outcome Measure	Baseline Measure	Outcome Measure	Baseline Measure
	Spring 2013	Spring 2012 (Prior Grade)	Spring 2013	Spring 2012 (Prior Grade)	Spring 2013	Fall 2012	Spring 2013	Fall 2012	Spring 2013	Fall 2012	Spring 2013	Fall 2012
K					●	●	●	●				
1					●	●	●	●				
2					●	●	●	●				
3	●	●*	●	●*	●	●	●	●				
4	●	●	●	●	●	●	●	●				
5	●	●	●	●	●	●	●	●	●	●	●	●
6	●	●	●	●	●	●	●	●	●	●	●	●
7	●	●	●	●	●	●	●	●	●	●	●	●
8	●	●	●	●	●	●	●	●	●	●	●	●
9	●*	●	●*	●								
10	●	●*	●	●*								

*Indicates the assessment is SAT-10

Performance Criteria for the AIMS/SAT-10 Value-Added Measures

In 2012-13, a school-wide composite value-added measure will be calculated in each subject (reading and math) based on the results of all students in grades 3-10 who are tested on AIMS or SAT-10 in the spring of 2013 and have AIMS or SAT-10 scores from the prior school year. The composite value-added measures are then converted to performance scores on the common 1 to 5 scale using the criteria shown in table 4. Teachers will receive separate school-wide value-added growth scores in reading and math, which will contribute equally to their REIL scores.

The performance criteria in table 4 are based on the number of standard deviations the school's value-added estimate differs from the statewide average. This approach is similar to those used in other performance evaluation systems developed by Teacher Incentive Fund grantees. [7]

Schools that fall within 0.5 standard deviation of the state average (above or below) are considered to have met expectations. Teachers in these schools will be assigned a performance score of 3. Teachers in schools that perform 0.5 and 1.5 standard deviations above the state average are assigned a performance score of 4, and those falling 0.5 to 1.5 below the state average are assigned scores of 2. Researchers have found that a difference of 0.5 standard deviations represents about one-fifth of a year's worth of academic growth. [8] Teachers in schools that fall more than 1.5 standard deviations above the state average will receive scores of 5, while those in schools falling 1.5 standard deviations below the state average will receive scores of 1. A difference of 1.5 standard deviations represents approximately three-fifths of a year.

---

[7] See Milanowski, A. (2011) *Resolving some issues in Using Value-Added Measures of Productivity for School and Teacher Incentives: Ideas from Technical Assistance and TIF Grantee Experience*. The Harvesting Project at the Center for Educator Compensation Reform. Madison, WI: Center for Educator Compensation Reform.

[8] Milanowski, A. (2011)

Table 4. Conversion of School Value-Added Measures to REIL Performance Scores

Performance Score	School Value-Added Estimates
5	Average student growth is well above expectations for schools with similar student characteristics: (More than 1.5 standard deviations above the state average; 1.6 years of growth)
4	Average student growth is above expectations for schools with similar student characteristics: (More than 0.5 standard deviations above the state average; + 1.2 years of growth)
3	The school's average growth meets expectations for schools with similar student characteristics: (Within 0.5 standard deviations above or below the state average; equates to approximately 1 year of growth)
2	The school's average growth is below expectations for schools with similar student characteristics: (More than 0.5 standard deviations below the state average; 0.8 years of growth)
1	Average student growth is well below expectations for schools with similar student characteristics: (More than 1.5 standard deviations below the state average; 0.4 years of growth)

In 2012-13, all K-12 teachers in REIL Alliance districts will receive reading and math scores on the 1 to 5 scale.[9] For teachers in schools that administer both AIMS/SAT-10 and Galileo K-12 (e.g., K-5, K-8), the average of their reading and math scores (rounded to the tenth decimal) will account for 15 percent of the REIL Score in 2012-13, with Galileo growth results making up the remaining 10 percent of their school growth component. For teachers in schools that do not administer the Galileo K-12 (e.g., 9-12), the AIMS/SAT-10 school value-added measures will comprise the full 25 percent of the school growth component of the REIL score (see Table 5).

[9] Until the roster verification system is in place and teachers have an opportunity to verify the grades, subjects, and students they teach, it will not be possible to base teachers' growth results on students' performance in specific subjects, grades, or classrooms.

Table 5. Weighting of the School Growth Component in 2012-13

	Teachers in Schools Administering AIMS/SAT-10 and Galileo K-12	Teachers in Schools Administering only AIMS/SAT-10
AIMS/SAT-10 Value-Added Measures	15%	25%
Galileo K-12 Categorical Growth Measures	10%	0%
Total	**25%**	**25%**

Resources on Value-Added Measures

Braun, H. (2005). Using Student Progress to Evaluate Teachers: A Primer on Value-Added Models. ETS. **http://www.ets.org/Media/Research/pdf/PICVAM.pdf**

Braun, H., Chudowsky, N., & Koenig, J. (2010). Getting Value Out of Value-Added: Report of a Workshop (2010). The National Academies. **http://216.78.200.159/Documents/RandD/Other/Getting%20Value%20out%20of%20Value-Added.pdf**

Brown Center on Education Policy at the Brookings Institute (2011). Passing Muster: Evaluating Teacher Evaluation Systems . Washington, DC: Author. **http://www.brookings.edu/~/media/Files/rc/reports/2011/0426_evaluating_teachers/0426_evaluating_teachers.pdf**

Center for Educator Compensation Reform (2012). Papers on Value-Added Measures. **http://www.cecr.ed.gov/development/valueAdded/papers.cfm**

Koretz, D. (2008). A Measured Approach. American Educator, Fall(2008). **http://www.aft.org/pdfs/americaneducator/fall2008/koretz.pdf**

Lipscomb, S. et al., (2010). Teacher and Principal Value-Added: Research Findings and Implementation Practices. Mathematica Policy Research. **http://www.mathematica-mpr.com/publications/PDFs/education/teacherprin_valueadded.pdf**

Milanowksi, A. (2011). Resolving some issues in using Value-Added Measures of Productivity for School and Teacher Performance Incentives. Center for Educator Compensation Reform. **http://www.cecr.ed.gov/pdfs/CECR_HP_ValueAdded.pdf**

Meyer, R. & Dokumaci, E .(2009). Value-Added Measures and the Next Generation of Assessments. Center for K-12 Assessment and Performance Management. **http://www.k12center.org/rsc/pdf/MeyerDokumaciPresenterSession4.pdf**.

RAND (2011). Using Student Performance to Evaluate Teachers (2011). **http://www.rand.org/content/dam/rand/pubs/research_briefs/2011/RAND_RB9569.pdf**

Value-Added Research Center (2012). Value-Added Tutorials. **http://varc.wceruw.org/tutorials/index.php**

Galileo Categorical Growth Measure

For teachers in REIL Alliance schools that serve students in grades K-8, 10 percent of their REIL scores in 2012-13 will be determined by their schools' overall growth results on the Galileo assessments. Galileo is an assessment system that includes summative and benchmark tests that align to Arizona state academic content standards. In 2012-13, REIL Alliance districts will administer fall (pretest) and spring (posttest) summative assessments in reading, math, writing, and science to students in the grades shown in table 6.

Table 6. Grades and subjects covered by Galileo pre and post tests in 2012-13

	K	1st	2nd	3rd	4th	5th	6th	7th	8th
Reading	●	●	●	●	●	●	●	●	●
Math		●	●	●	●	●	●	●	●
Science						●	●	●	●
Writing						●	●	●	●

Students' growth from the fall 2012 pretest to the spring 2013 posttest will be evaluated using a method called Categorical Growth Analysis (CGA), which was developed by the Galileo developers' Assessment Technology Incorporated (ATI). A summary of ATI's CGA approach is provided on page 23. CGA uses Chi-square analysis to estimate the proportion of students within each school that moved from scoring below the standard on the fall pretest to above the standard on the spring posttest (or vice-versa). More details on the Galileo pre and post tests and the CGA methods are available in ATI's white paper on Instructional Effectiveness Assessment.[10]

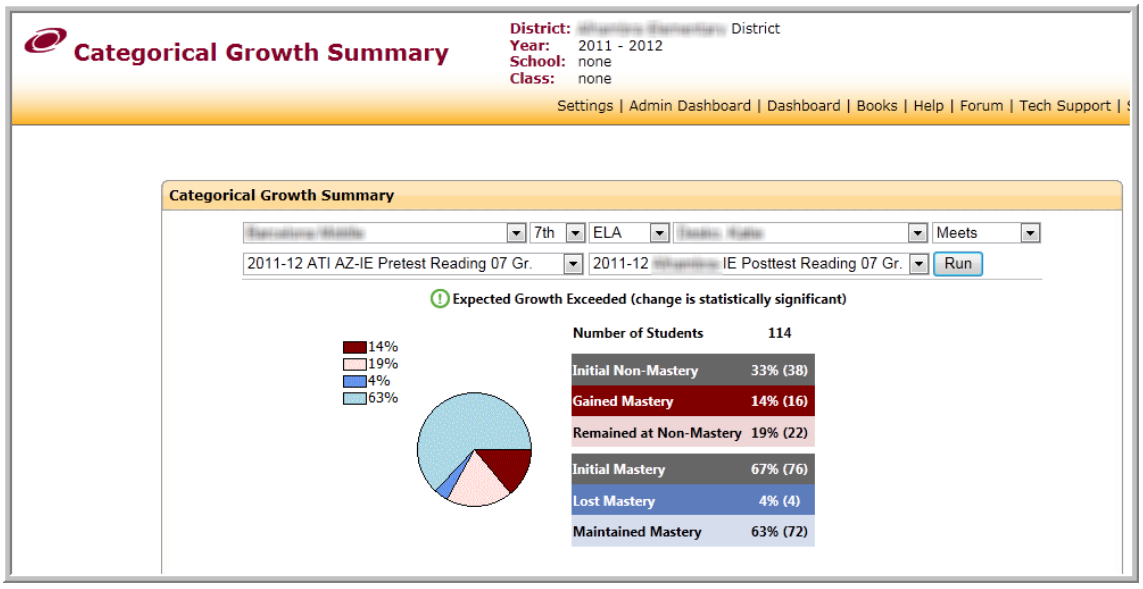
The CGA compliments the AIMS/SAT-10 value-added measures in a few notable ways. First, the CGA provides "criterion referenced" growth measures, which means that school growth is evaluated based on the extent to which students maintain or improve their mastery of the Arizona academic content standards. The AIMS/SAT-10 value-added measures, on the other hand, are purely "normative" in that schools are compared to statewide averages after adjusting for student and school characteristics. Second, the Galileo is administered in more grades and subjects than the AIMS/SAT-10, which allows for a more comprehensive analysis of schools' overall performance. The CGA growth measures will be strong leading indicators of schools' AIMS/SAT-10 value-added measures. The Galileo assessments are directly aligned to the Arizona state academic content standards, so schools that score high on the CGA measures can anticipate similarly strong results on the AIMS/SAT-10 value-added measures.

[10] *http://www.ati-online.com/pdfs/researchK12/InstructionalEffectivenessAssessment.pdf*

Overview of the Galileo Categorical Growth Analysis (CGA)

The basic approach to Categorical Growth Analysis (CGA) is as follows. First, cut scores for the pretest (fall) and posttest (spring) are established. The cut scores on the posttest are raised relative to those on the pretest by the amount of progress that is expected over the course of the year. The posttest cut score is based on the average expected growth across a large number of districts. After the cut scores have been established, students are classified at Time 1 (the fall pretest) with regard to whether they demonstrated mastery of the state standards. At Time 2 (the spring posttest), each student is categorized with regard to whether their mastery classification stayed the same as it was at Time 1 or changed. A finding of no significant change indicates that expected growth has been maintained. Two types of change are possible. One involves a significant increase in the relative proportion of students achieving standards mastery. The other reflects a significant decrease in the relative proportion of students achieving standards mastery.

The Galileo *Categorical Growth Summary* is currently available to all school districts (via ATI) that participated in the Galileo Instructional Effectiveness Pilot in 2011-12. The figure below illustrates the *Categorical Growth Summary*.



Assessment Technology Incorporated (2012). Instructional Effectiveness Assessment White Paper. Tucson, AZ: Author (p 7-8).

Performance Criteria for the Galileo Categorical Growth Measures

As described above, the CGA evaluates each school's growth according to how many students increase and decrease their mastery of the state academic content standards from fall to spring. ATI assigns a numerical score between 1 and 3 for the CGA results of each grade and subject tested in the school. A score of 1 indicates there was a significant decrease in the proportion of students achieving mastery, a 2 indicates that the majority of students maintained their mastery, and a 3 indicates a significant increase in mastery levels.

Table 7. Conversion of Galileo CGA Growth Measures to REIL Component Scores

Performance Score	Galileo CGA Criteria
5	**Substantial increase in mastery.** Large increase in the proportion of students achieving standards mastery. Average CGA score between 2.6 and 3.0.
4	**Increase in mastery.** Increase in the proportion of students achieving standards mastery. Average CGA score between 2.3 and 2.6
3	**Maintained mastery**: No change in the proportion of students achieving standards mastery. Average CGA score between 1.8 and 2.2.
2	**Decrease in mastery.** Reduction in the proportion of students achieving standards mastery. Average CGA score between 1.5 and 1.7.
1	**Substantial decrease in mastery.** Large reduction in the proportion of students achieving standards mastery. Average CGA score between 1.0 and 1.4.

The CGA scores for all tested grades within a subject will be averaged (rounding to the tenth decimal) in order to create a composite CGA score for that subject (math, reading, science, and writing).[11] These composite CGA scores, which will range from 1 to 3, will then be converted to the common 1 to 5 scale using the criteria in table 7. Schools where the

[11] The composite CGA score in each subject is constructed by averaging the results from all grades tested in the subject, weighting each grade's results by the number of students with valid pretest and posttest results in the subject.

majority of students increase their mastery of state standards will receive scores of 4 or 5, depending on the number of grade-levels that experienced increased mastery. Schools where the majority of students decreased their mastery levels will receive scores of 1 or 2, depending on the severity of the decline. Schools that maintain the proportion of students achieving mastery from fall to spring are assigned scores of 3.

Separate performance scores on the 1 to 5 scale will be calculated for each subject. These subject-level scores will be averaged to create a single performance score for each school, rounded to the tenth decimal, which will account for 10 percent of teachers' REIL Scores in 2012-13. In constructing the composite score, subject-level results will be weighted according to the number of students tested.

Calculating the REIL Score

The scores from the individual performance measures described in the previous sections will be combined to create a single measure of educator effectiveness called the REIL Score. The REIL Score will be used to determine teachers' annual state performance classifications and REIL performance awards.

Step 1: Convert the LOI and REIL Growth results to scores on a 1-5 scale

The first step to constructing the REIL score is to convert the results of all the individual LOI and school growth measures to the common 1 to 5 scale using the conversion chart in table X. Putting the individual measures on the same scale ensures the criteria used to evaluate performance are equivalent in rigor across all measures. This allows teachers and principals to make meaningful comparisons of their performance on different measures. The criteria in the conversion chart can be used by teachers and principals to set annual performance goals and communicate performance expectations.

Step 2: Combine the LOI and REIL Growth scores from step 1 using weights

After the LOI and REIL Growth measures are converted to the 1 to 5 scale, they can be combined to create the REIL score. This is done by multiplying the score for each individual performance measure by its weight and then adding the weighted scores together. The scorecards used to make these simple calculations are presented on pages 28 and 29. Teachers in schools that do not administer the Galileo assessments have a separate scorecard to reflect differences in the weighting of the school growth component of their REIL Score. These calculations result in the REIL Score, which will range from 100 to 500.

Step 3: Determine performance classifications

The REIL score is used to determine teachers' annual state performance classifications (see Scorecards). The state requires all classroom teachers to receive one of the four classifications: Highly Effective, Effective, Developing, and Ineffective. The REIL system partitions the "Effective" classification into two sub-classifications: "Effective 1" and "Effective 2". Cut points are placed on the REIL Score continuum to identify range of REIL scores within each performance classifications. For example, in order for teachers to receive a designation of "Effective 1", their REIL score must fall between 300 and 349.

REIL Component Score Conversion Chart: 2012-13

Measure	Score Criteria				
	5	4	3	2	1
Total LOI Points Earned	Total LOI points earned over 2 observation cycles is between 189 and 220.	Total LOI points earned over 2 observation cycles is between 155 and 188.	Total LOI points earned over 2 observation cycles is between 117 and 154	Total LOI points earned over 2 observation cycles is between 67 and 116.	Total LOI points earned over 2 observation cycles is between 0 and 66.
Pct. LOI Elements Rated 3 or Higher	100% of LOI element ratings are greater than or equal to 3	90% to 99% of LOI element ratings are greater than or equal to 3	60% to 89% of LOI element ratings are greater than or equal to 3	25% to 59% of LOI element ratings are greater than or equal to 3	0% to 24% of LOI element ratings are greater than or equal to 3
AIMS/SAT-10 School Value-Added Measure (Reading & Math)	Average student growth is well above expectations (More than 1.5 standard deviations above the state average)	Average student growth is above expectations (More than 0.5 standard deviations above the state average)	The school's average growth meets expectations (Within 0.5 standard deviations (plus or minus) of the state average)	The school's average growth is below expectations (More than 0.5 standard deviations below the state average)	Average student growth is well below expectations (More than 1.5 standard deviations below the state average)
Galileo K-12 Categorical Growth (Composite – All Subjects)	Large increase in the proportion of students achieving standards mastery. Average CGA score between 2.6 and 3.0.	Increase in the proportion of students achieving standards mastery. Average CGA score between 2.3 and 2.5	No change in the proportion of students achieving standards mastery (mastery maintained). Average CGA score between 1.8 and 2.2.	Reduction in the proportion of students achieving standards mastery. Average CGA score between 1.5 and 1.7.	Large reduction in the proportion of students achieving standards mastery. Average CGA score between 1.0 and 1.4.

The REIL Scorecard

REIL SCORECARD 2012-13

K-12 teachers in schools with both AIMS/SAT-10 and Galileo Growth measures

Teacher: _____

District: _____

School: _____

Assignment: _____

Experience: _____

REIL Score Range	Performance Classification
100 to 199	Ineffective
200 to 299	Developing
300 to 349	Effective 1
350 to 399	Effective 2
400 to 500	Highly Effective

Measure	Performance Score (1-5)		Weight		REIL Points Earned	Total Possible REIL Points
Total REIL Points Earned	_____	X	50	=	_____	250
Pct. LOI Elements Rated 3 or Higher	_____	X	25	=	_____	125
AIMS/SAT-10 School Value-Added Measure	_____	X	15	=	_____	75
Galileo Categorical Growth Measure	_____	X	10	=	_____	50
				REIL Score:	_____	**500**

REIL SCORECARD 2012-13

K-12 teachers in schools with only AIMS/SAT-10 growth measures in 2012-13

Teacher: _____

District: _____

School: _____

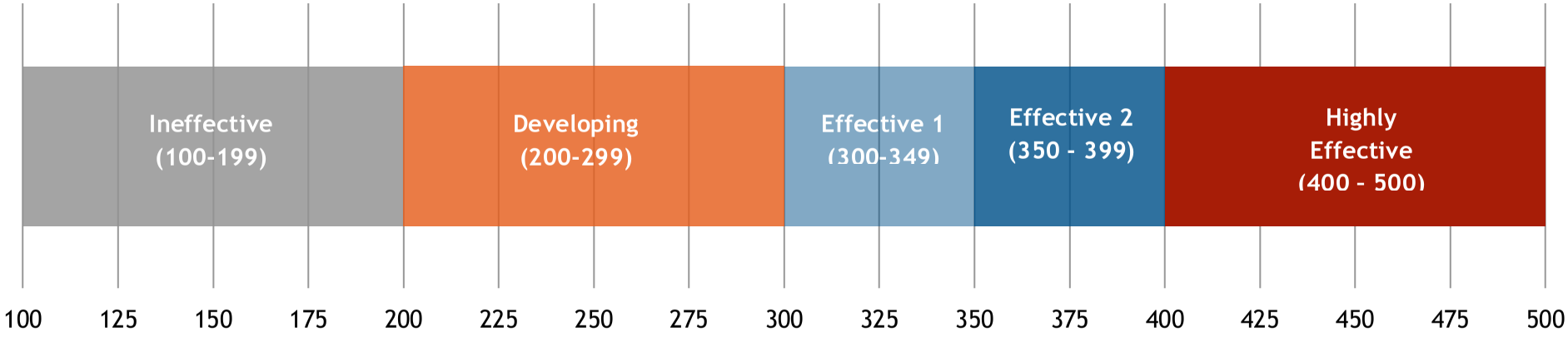
Assignment: _____

Experience: _____

REIL Score Range	Performance Classification
100 to 199	Ineffective
200 to 299	Developing
300 to 349	Effective 1
350 to 399	Effective 2
400 to 500	Highly Effective

Measure	Performance Score (1-5)	Weight		REIL Points Earned	Total Possible REIL Points
Total REIL Points Earned	_____	X 50	=	_____	250
Pct. LOI Elements Rated 3 or Higher	_____	X 25	=	_____	125
AIMS/SAT-10 School Value-Added Measure	_____	X 25	=	_____	125
		REIL Score:		_____	**500**

The REIL Score Continuum



Key Terms in the Guidebook

AIMS/SAT-10 Value-Added Measure

An estimate of a school's contribution to student growth in reading and math on AIMS and SAT-10. A school's value-added estimate represents the difference between its students' actual and predicted test scores, where the predicted test scores are determined based on students' prior achievement, background characteristics, and school characteristics. Schools' value-added estimates are compared to the average value-added estimate in the state to determine REIL performance scores.

Common 1 to 5 Scale

The scale on which results of all individual performance measures are placed before they are combined into the REIL Score. Conversion charts are used to place the Total LOI Points, LOI 3+ Rate, AIMS/SAT-10 value-added measures, and Galileo Categorical Growth on the common 1 to 5 scale.

Conversion Charts

Charts that explain the criteria for translating results from the individual performance measure (LOI, AIMS value-added, and Galileo Categorical Growth) to the 1 to 5 scale.

Galileo Categorical Growth Measure

A criterion-referenced growth measure developed by ATI. This measure tracks the proportion of students who maintain or advance their mastery of Arizona state content standards from the fall pretest to spring posttest.

The Learning Observation Instrument (LOI)

The instrument used to measure the quality of teachers' instructional practice through classroom observations. The LOI is made up of five rubrics: Content, Formative Assessment, Instructional Strategies, Learner Engagement, and Learning Community.

LOI 3+ Rate

The percentage of LOI element ratings that teachers receive from all observation cycles that are rated a 3 or higher on the LOI rubrics. The LOI 3+ Rate accounts for 25 percent of the REIL score in 2012-13.

LOI Observation Cycle

The three stages to each evaluation of classroom instruction: Pre-Conference, Observation, Post-Conference. Teachers in REIL extend districts will complete two observation cycles in 2012-13. A single observation cycle should take ten school days to complete.

REIL Components

The four types of REIL performance measures: School Growth, Individual Growth, Team Growth, Learning Observation

REIL Score

The combined measure of educator effectiveness used to determine state performance classifications and REIL performance awards. REIL Scores range from 100 to 500.

State Performance Classifications

The four evaluation designations that districts must assign to Arizona classroom teachers: Ineffective, Developing, Effective, Highly Effective. In the REIL system, the "Effective" category is partitioned into "Effective 1" and "Effective 2".

Total LOI Points Earned

The total number of points that teachers accumulate from their evaluators' ratings of LOI elements on the 0 to 5 scale.

Weights

The proportion of the REIL Score that is based on an individual performance measure. Weights are applied to each individual performance measure before they are added together to create the REIL Score.

Frequently Asked Questions

Why are two observation cycles required per year?

Research shows that in order to reliably measure the quality of your instruction, the teaching and learning in your classroom must be observed on multiple occasions.[12] A single lesson provides just a snapshot of your instructional practice. Taking multiple snapshots will yield a more complete picture of the teaching and learning that occurs in your classroom and in turn provides you with more reliable feedback on how you can improve. All teachers know that some lessons are more successful than others. Each additional observation reduces the risk that your overall learning observation results for the school year will be affected by a single lesson that does not go as planned.

What should I do if I notice a reporting error in my LOI results?

Evaluators are responsible for entering the results of your observation cycle into the REILize Decision Support system. If you notice an error in your LOI results you should talk to your evaluator as soon as possible so that the error can be corrected. When the REILize Decision Support System is in place, teachers will be able verify their scores online and report any errors in observation results.

What steps are taken to ensure that all evaluators rate the LOI rubric criteria with the same levels of rigor and objectivity?

In order for the LOI results to be fair and accurate for all teachers, it is essential for all principals and peer evaluators to use the same standards when rating the LOI criteria. To ensure this is the case, all evaluators must be trained and certified before conducting classroom observations. To receive certification evaluators must receive a passing score on an assessment of their knowledge of the LOI rubric criteria and ability to accurately score teachers. In addition, evaluators' LOI ratings are monitored by MCESA staff and any systematic discrepancies or abnormalities in their scoring will be flagged and investigated.

---

[12] Kane, T.J., & Staiger, D.O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: Bill & Melinda Gates Foundation.

Why doesn't the professional responsibilities rubric count toward my REIL Score?

The REIL Score is currently informed by elements that can be observed through classroom observation. The professional responsibilities rubric cannot be scored during a classroom lesson. In addition, each district handles professional responsibilities outside of the formal evaluation process according to their established policies and procedures.

Where can I learn more about the Learning Observation Instrument?

An informational webcast on the LOI is hosted at: REIL Learning Observation Instrument Webcast. Detailed information on the LOI is also available in the Learning Observation Handbook as well as in the REIL 2012-13 Guidebook.

What assessments are used to measure the school growth component in 2012-13?

In 2012-13, AIMS and SAT-10 reading and math scores for students in grades 2-10 will be used to create school-level growth measures. Elementary and middle schools will also be administering Galileo K-12 assessments in reading, writing, math, and science. Students' growth from fall 2012 to spring 2013 on these assessments will also be incorporated into the school growth component of the REIL score. For more information on the school growth measures please refer to the School Growth Component section of the guidebook.

What factors are controlled for in the value-added model?

Value-added models evaluate teacher or school effectiveness by tracking student test score growth from one grade to the next using statistical techniques accounting for other factors that also impact student learning but are beyond the control of teachers and schools. The school value-added model used in 2012-13 controls for the following characteristics at both the student and school-level:

- Prior achievement in reading and math
- Gifted and Talented status
- Disability/special education status
- Free and reduced meal eligibility status
- English Language Learner (ELL status)
- Fluent English Proficient (FEP status)
- Migrant status
- Homeless status
- Student mobility status (new to the school in non-promotional year)

Our student population is highly mobile and students often transfer in and out of the school throughout the year. Will these students be included in our school growth results?

Only students who are enrolled in the school for the full academic year will contribute to your school's growth results. According to the Arizona Department of Education, a student is considered enrolled for the full academic year if he/she is enrolled in the school during the first 10 school days and remains enrolled through the date of the AIMS/SAT-10 test. Additionally, the value-added model controls for the effects of across-year student mobility on student growth by including an indicator of students who are new to the school at the beginning of the school year because of a non-promotional school transfer. This variable equals 1 if a student is new to this school and not enrolled in the first-grade offered by the school.[13]

How do the growth measures account for student attendance?

Students who are chronically absent are likely to make slower achievement growth than their peers. The value added model attempts to control for the the influence of student attendance on student achievement growth by factoring absentee and tardiness rates from the prior year into the equation used to predict students' expected test scores in the current year. Attendance from the prior year is used instead of attendance in the current year because student attendance in a given teacher's class will be explained to some extent by the quality of instruction that the teacher is providing. Since teachers can influence their students' attendance in the current year, it violates one of the criteria for selecting control variables (it is not outside the control of teachers). Therefore, we use attendance from the prior year as a control variable.[14]

[13] This variable is designed to capture mobility that is not due to normal promotions from elementary to middle school and middle school to high school. Prior studies have found this type of mobility associates with lower academic achievement (Ballou, Sanders, & Wright, 2004).

[14] Similar approached are used in the value-added evaluation systems in Washington D.C. (by Mathematica) and New York City (by Value Added Research Center) public school districts.