Shaun D. Hutchins, Ph.D. Paige Hartman DeBaylo, Ph.D. Holly Williams, Ph.D.

Publication 17.33 June 2018



Department of Research and Evaluation

Thinking Heuristically About Student Growth in Austin Independent School District

Abstract

On May 22, 2018 the Professional Pathways for Teachers (PPfT) oversight committee convened for the final meeting of the 2017–2018 school year. One point of discussion was the impact of Texas Education Agency's (TEA) State of Texas Assessment of Academic Readiness (STAAR) end-of-course (EOC) substitution policy on school-wide value-added Educational Value-Added Assessment System (EVAAS) scores. Specifically, the following question was put forth to the committee for discussion: How might the exclusion of Reagan High School's highest-achieving students affect the school-wide value-added EVAAS scores? Austin Independent School District's (AISD) Department of Research and Evaluation (DRE) staff responded by investigating the growth and achievement data of Reagan's students. Even though Reagan students demonstrated some of the lowest achievement in the district, they also demonstrated some of greatest growth in the district (due primarily to the growth of the lowestachieving students at the school). The question-driven exploration of growth and achievement data at Reagan provided a practical framework within which to organize some of the SAS EVAAS reports, and perhaps more importantly, to help demonstrate the relevance of growth data to principal decision making. The question-driven exploration of SAS EVAAS data also showed how easy it can be to conflate growth and achievement, rather than treating the two as independent measures.

The Main Takeaway

Students grow independently of how they achieve. High- and low-achieving students have equal opportunities to show year-to-year growth, and this equal opportunity for growth independent of achievement is realized in AISD's EVAAS data. To see growth as independent from achievement, we must resist our tendency to conflate growth with achievement (i.e., resist conceptualizing growth gains as dependent upon the inclusion of our highest-achieving students' data), remembering that we need to continuously work to grow our high-performing students, too, and embrace, maybe even celebrate, the idea that sometimes our lowest-achieving students show the greatest growth among all student achievement levels.

Purpose

The purpose of this exploration was to bring additional data, information, and ways of thinking about student growth to ongoing district conversations about the measurement of campus-level student growth.

Goals

- 1. To provide information regarding the TEA STAAR substitution policy and its implications for growth and achievement
- To address a growth model validity concern raised by the PPfT oversight committee regarding missing achievement data in the district's growth model due to TEA-approved substitute assessments for STAAR EOC assessments
- 3. Through use of a concrete and practical case, to demonstrate how SAS EVAAS web reports can be used to help staff deliberately and rationally think about the distinct dimensions of student achievement and student growth
- 4. To advance a working hypothesis about a common decision-making process, which we refer to as the achievement heuristic, that potentially explains how district staff process and respond to information about students' achievement and growth (In some instances, the achievement heuristic may lead staff to false conclusions, while in other instances, the decision processes associated with the achievement heuristic may be capitalized on and transferred to novel situations.)

TEA STAAR EOC Substitute Assessment Policy

The commissioner of TEA approves specific assessments (e.g., PSAT, SAT, ACT) that may be used as substitute assessments in place of a corresponding end-of-course (EOC) assessment. If a student achieves the equivalency standard on an approved EOC substitute assessment, then he or she may elect to not take the STAAR EOC assessment and substitute the satisfactory score on the substitute assessment to meet graduation requirements. If a student fails the STAAR EOC twice but achieves the equivalency standard on an approved EOC substitute assessment, then he or she may replace his or her failing score with the passing equivalency to meet graduation requirements.

The implications of substitute assessments are different for accountability and for growth purposes. For accountability, a score or equivalency is recorded in either case of substitution. However, with regard to growth modeling, STAAR EOC data are not included in growth models when the substitution happens for students electing not to take the STAAR EOC after achieving the equivalency standard on EOC substitute assessments. STAAR EOC data are included in growth models when the substitution happens for students failing the STAAR EOC and electing to replace their failing score with the passing equivalency on an EOC substitute assessment.

Accessing Your School's Data

To access a school's SAS EVAAS web reports, log onto: https://evaas.sas.com/

Each AISD principal is provided with a login for their school's data. If you need help with access to your school's account, please contact the AISD Department of Research and Evaluation for assistance at 4-1724.

Aggregate school-level comparisons may be made with other AISD schools. However, each principal may only drill down into their students' EVAAS data.

An Example Case Examined With SAS EVAAS Web Reports

Like many high schools, Reagan High School sees some of its students fulfill the Algebra I EOC requirement with a substitute assessment. For example, in 2017–2018, 30 students at Reagan fulfilled the Algebra I EOC requirement with a substitute assessment, and 62 students fulfilled the English II EOC requirement with a substitute assessment. In the cases where students achieve the equivalency standard on an EOC substitute assessment, the students can opt out of taking that associated EOC assessment; therefore, no STAAR EOC data for those students are included in school-wide growth calculations. To understand the potential impact of omission of student data in the calculation of school-wide growth, SAS EVAAS web reports are used to create a data narrative about missing data at the highest and lowest student achievement quintiles at Reagan. To understand the potential impacts of data omission at other AISD schools, the same general procedure used in the present case should be followed using that school's data.

To start, let us first examine Reagan High School's overall composite school-wide value-added results (Figure 1). Looking only at the math and reading composite scores (the final school-wide value-added score used in PPfT appraisal is the average of the PPfT score for the reading composite and the PPfT score for the mathematics composite), a level of 5 is shown for math in 2017 and level 3 for reading.

Figure 1

Composite School Value-Added Report For Reagan High School

2016-2017 Composite Trends						
Composite Type	One-Year Trend*		Two-Year Trend*		Three-Year Trend*	
Composite Type	Index	Level	Index	Level	Index	Level
Overall	0.12	3	1.03	4	-2.34	1
Mathematics	4.78	5	6.95	5	1.85	4
Reading	-0.20	3	-0.38	3	-1.51	2
Science	-2.23	1	-1.54	2	-1.83	2
Social Studies	-1.97	2	-2.56	1	-3.11	1

Source. SAS EVAAS web reports: school composite value-added report

EVAAS composite scores are reported on a 1 to 5 scale, based on strength of evidence for meeting the growth standard, and are then converted to a 1 to 4 score that counts for 10% of the PPfT appraisal score (Table 1). Growth levels of 5 and 3 in math and reading, respectively, would result in a 3.5 school-wide value-added score, or 35 of 40 possible points, in a PPfT appraisal for the school-wide value-added component.

Table 1

School EVAAS index scores are converted into growth levels, which correspond to points teachers earn for PPfT appraisal.

EVAAS index score range	Growth level	Definition	PPfT appraisal points
Below -2.00	Level 1	Significant evidence growth was below growth standard	1
-2.00 to -1.01	Level 2	Moderate evidence growth was below growth standard	2
-1.00 to 0.99	Level 3	Met growth standard	3
1.00 to 1.99	Level 4	Moderate evidence growth exceeded growth standard	4
2.00 or above	Level 5	Significant evidence growth exceeded growth standard	4

Source . SAS EVAAS web reports; Schmitt and Hutchins, 2016

Reagan's school-wide value-added results suggest strong evidence that students are growing in math (Figure 2). Figure 2 shows a growth measure of 85.5, with blue coding to indicate significant evidence the school's students made more progress than the growth standard (i.e., dividing 85.5 by the standard error, 17.9, which is how the growth index was determined in Figure 1). The growth measure is an estimate of Reagan's school effect, based on a comparison of Reagan's students' progress on the STAAR EOC Algebra I assessment with the average school in the state of Texas.

Figure 2 School-Wide Value Added in Algebra I for Reagan, by Year

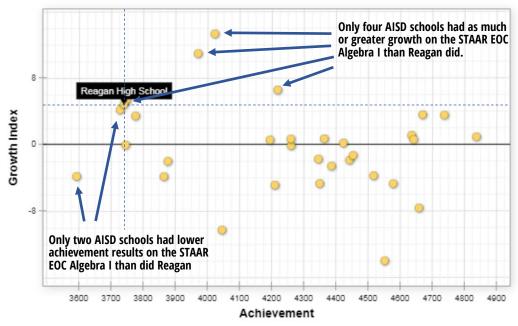
Subject	Year	Number of Students	Average Score	Average Percentile	Average Predicted Score	Average Predicted Percentile	Growth Measure	Standard Error
	2015	290	3546.9	26	3657.2	35	-109.0 R	16.4
Algebra I 2016	2016	200	3861.9	46	3749.8	36	112 2 DR	22.2
	2017	211	3829.8	32	3741.4	25	<u>85.5 DB</u>	17.9

Source. SAS EVAAS web reports; school value-added by subject report for the Algebra I EOC

In 2017, Reagan was one of the lowest-achieving schools in AISD, based on achievement results on the STAAR EOC Algebra I assessment (Figure 3). How is it possible for Reagan to show strong growth gains? Looking at Figure 3, only two schools had lower campus-level achievement results than Reagan did on the STAAR EOC Algebra I assessment. Yet, despite the achievement results for Reagan, only four schools in all of AISD had students who demonstrated as much growth as did the students at Reagan. On the 2017 STAAR EOC Algebra I assessment, Reagan exemplified the case of low achievement, high growth.

Figure 3





Understanding Your School's Data

The SAS EVAAS reports most relevant to principals are the school reports. There are four school-level reports available to principals:

- School composite value-added report
- School decision dashboard
- School value-added by subject report
- School diagnostics report

To get to your school reports from the SAS EVAAS home page, navigate to reports, school reports, and select one of the four school-level reports available for your school.

From the home page, you also have the option of understanding your school's data with the SAS EVAAS Guided Path for assessing schoolwide growth. The Guided Path offers a stepwise view of your school's data across three reports:

- School value-added
- School diagnostic
- Decision dashboard

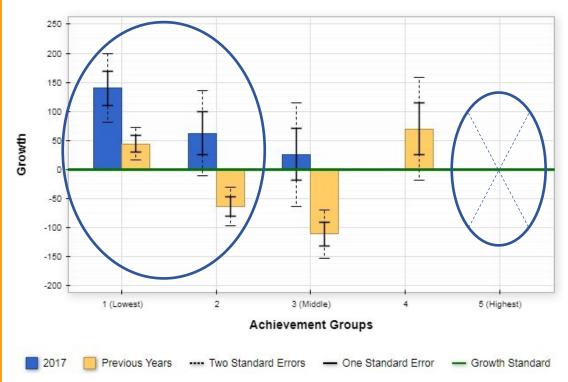
Each report in the Guided Path includes a series of reflection questions to guide your examination of the data.

Source. SAS EVAAS web reports; district summary scatterplot report in 2017 for the Algebra I EOC

Thinking Heuristically About Student Growth in Austin Independent School District

Given the overall strong school-level growth, it is important to understand which of Reagan's students made significant contributions to the positive result. In Figure 4, Reagan's students are divided into quintiles according to their STAAR EOC Algebra I achievement results (e.g., 1 groups Reagan's students based on their standing in the lowest 20% of all students in district distribution on the STAAR EOC Algebra I assessment, 5 groups Reagan's students based on their standing in the highest 20% based on all students in the district). The bars show growth relative to the growth standard; blue represents 2017, and tan represents average growth, with up to 3 years of prior data on the STAAR EOC Algebra I assessment. At least two findings should stand out. First, the majority of Reagan's strong growth in 2017 comes from the lowest-achieving students (i.e., those in the bottom 20% and 40% of the district distribution). Second, there is no student representation in the top 20%.

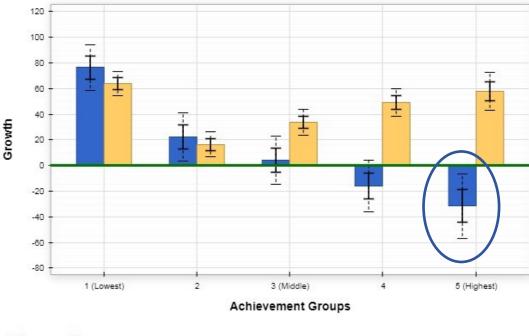
Figure 4 School Diagnostic for Reagan on the STAAR EOC Algebra I Assessment



Source. SAS EVAAS web reports; school diagnostic report for the Algebra I EOC

Given the absence of student representation in the highest-achievement quintile at Reagan in 2017, exploration of the possible impact of the omission of those students is warranted. What if the lack of student representation in the top 20% of achievement results was due to substitution for the STAAR EOC Algebra I assessment? To explore this, we turn to Figure 5. Figure 5 shows the overall district growth trend for all AISD students in 2017 on the STAAR EOC Algebra I assessment. If we assume that Reagan had student data for students in the top 20% and that they performed similarly to how the average AISD student in the top 20% performed, the growth falling below the growth standard likely would have washed out the overall positive results Reagan experienced on the STAAR EOC Algebra I assessment in 2017.





2017 Previous Years ---- Two Standard Errors — One Standard Error — Growth Standard

Source. SAS EVAAS web reports; district diagnostic report for the Algebra I EOC

Assuming Reagan's hypothetical group of high achievers would show growth like that of the average AISD high achiever might not be a fair assumption. Alternatively, an assumption could be made about Reagan's students' performance on the substitute assessments. Figure 6 shows the actual growth estimate for Reagan's test takers on the 10th-grade PSAT NMSQT math for each quintile, Figure 7 shows the actual growth estimate for Reagan's test takers on the 11th-grade PSAT NMSQT math for each quintile, and Figure 8 shows the actual growth estimate for Reagan's test takers on the SAT math for each quintile.

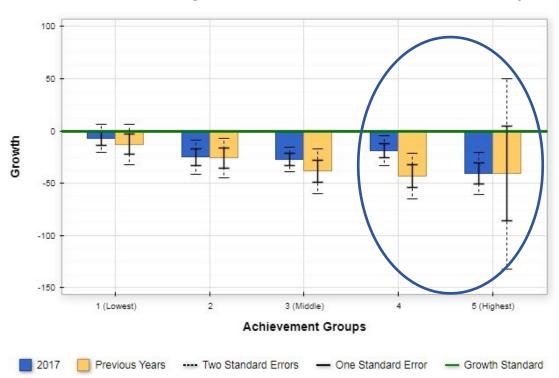
As is depicted in the growth estimates in Figures 6 through 8, Reagan's high-achieving students demonstrated growth below the growth standard on all three assessments. If the students who were high achievers on these substitute assessments had instead taken the STAAR EOC Algebra I assessment and demonstrated similar growth below the standard on the EOC, then the results below the growth standard likely would have washed out the overall positive results Reagan experienced on the STAAR EOC Algebra I assessment in 2017.

School Composite Value-Added Reports

The school composite value-added report presents an overview of all school-wide value added scores available for a school by subject.

The school composite value-added report provides an overall composite school level growth measure across all tested subjects, as well as, the composite school level growth measures for each tested subject. Each growth measure is assigned a growth level on a 5 point scale based on the strength of evidence of progress towards the growth standard (Table 1).

A description of the tests used in each composite is provided in a table at the bottom of the school composite value-added report. In some instances, multiple tests are combined in a subject's composite. For example, the high school reading composite typically includes test data from the English I EOC and English II EOC. Similarly, the middle school math composite typically includes STAAR math for grades 5 through 8 and the Algebra I EOC.



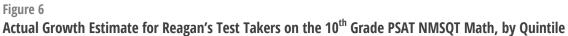
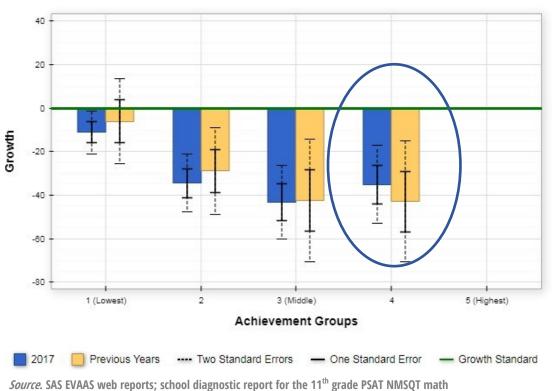


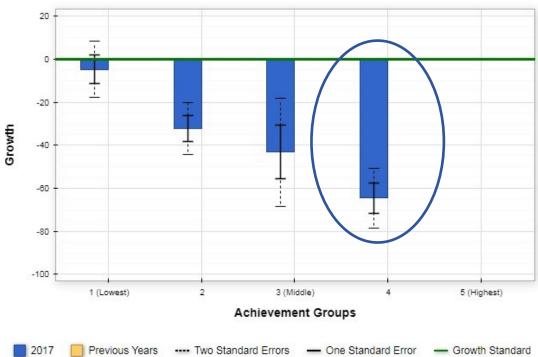
Figure 7



Actual Growth Estimate for Reagan's Test Takers on the 11th Grade PSAT NMSQT Math, by Quintile

Source. SAS EVAAS web reports; school diagnostic report for the 10th grade PSAT NMSQT math





Source. SAS EVAAS web reports; school diagnostic report for the SAT math

We may also wonder if high achievers can grow? With growth models, sometimes concerns are raised about the growth potential for high achievers; that is, can AISD's highest achievers also show growth? The growth standard requires evidence of year-toyear academic progress on pace with academically similar peers. Gaining ground indicates evidence of progress exceeding that pace, losing ground indicates evidence of progress falling short of that pace. Consequently, for high achievers to grow, they must show evidence of year-to-year academic progress on pace with, or exceeding, academically similar peers.

Here, we examine LASA to expand the example case to one of AISD's highest-achieving schools. For LASA, the answer is clear: Yes, high achievers can (and do) show significant evidence of growth exceeding the growth standard. Figures 9 through 11 show strong growth gains for all of LASA's high-achieving students on the 10th- and 11th-grade PSAT NMSQT math and SAT math assessments.

School Decision Dashboard

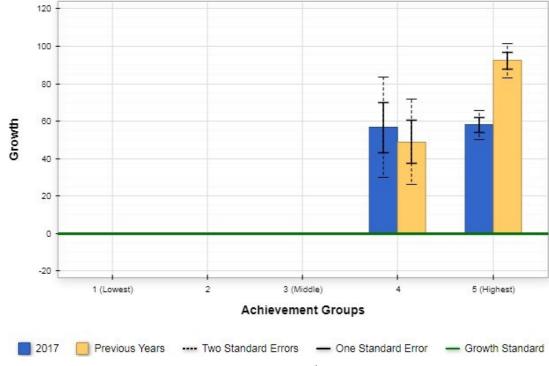
The school decision dashboard presents a consolidated view of information from the school valueadded and school diagnostic reports. Growth by assessment, grade, subject, and achievement level are integrated into a single graphical display using a table, pie charts, and color coding to summarize the various data.

An overall table organizes growth information by assessment, grade, and subject. Each column in the table represents a tested subject. Each row in the table represents an assessments, and where applicable, the assessment by grade combination. For example, EOC assessments are displayed in a single row without any grade information, but STAAR assessments receive a single row for each tested grade.

Within each subject by assessment cell of the table, a pie chart provides disaggregated growth by achievement quintile (i.e., 1 groups students based on their standing in the lowest 20% of all students in Texas on the assessment, 5 groups students based on their standing in the highest 20%). Each wedge of the pie charts is color coded to show strength of evidence for student growth relative to the growth standard:

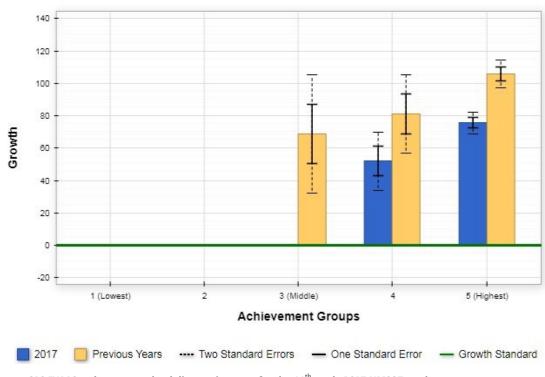
- Blue indicates that the school's students made more progress than the growth standard.
- Green indicates that the school's students made progress similar to the growth standard.
- Red indicates that the school's students made progress less than the growth standard.





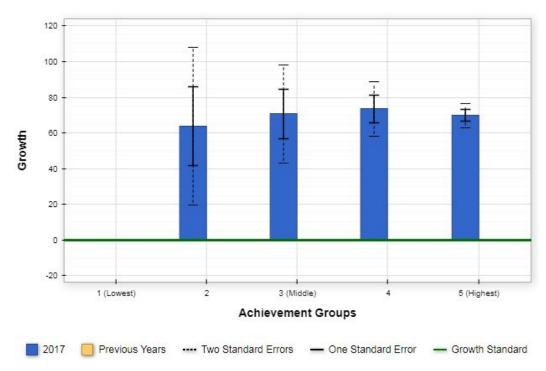
Source. SAS EVAAS web reports; school diagnostic report for the 10th grade PSAT NMSQT math

Figure 10 Actual Growth Estimate for LASA's Test Takers on the 10th-Grade PSAT NMSQT Math, by Quintile



Source. SAS EVAAS web reports; school diagnostic report for the 11th grade PSAT NMSQT math

Figure 11 Actual Growth Estimate for LASA's Test Takers on the SAT Math, by Quintile



Source. SAS EVAAS web reports; school diagnostic report for the SAT math

Achievement Heuristic

Achievement—in particular, the concept of student achievement in education conveys the idea of assessed knowledge and skills, often measured with standardized assessments. A heuristic may be thought of as a simple rule or mental shortcut, based on past experiences, that helps people quickly draw conclusions about complex matters, without deliberately or rationally processing all available information. Heuristics, although imperfect, are often adequate, despite their susceptibility to result in flawed conclusions and/or decisions (Kahneman, 2011; Tversky & Kahneman, 1974). Taken together, we advance the hypothesis of an achievement heuristic as a mental shortcut about the relationship between achievement and performance that is used to quickly draw conclusions about students' performance and decide or act upon those conclusions.

The simple rule underlying the achievement heuristic is that achievement and performance are positively associated such that, with high achievement, we can reasonably expect strong performance, and conversely, with low achievement, we can reasonably expect undesirable performance. Whether the achievement heuristic is a good fit for the information being assessed depends on what we map to performance. If we cognitively map, say, grade point average (GPA) to performance, then the heuristic's simple rule helps us quickly and adequately draw conclusions about high achievers and strong GPAs, or low achievers and undesirable GPAs. Here, the

School Value-Added by Subject Reports

The school value-added by subject report shows the average academic growth the school's students made in each tested subject and grade. For consecutively tested subjects, data by school year is also shown.

The growth measure for each grade and year shows the strength of evidence for student growth relative to the growth standard:

- Dark blue indicates significant evidence that the school's students made more progress than the growth standard.
- Light blue indicates moderate evidence that the school's students made more progress than the growth standard.
- Green indicates evidence that the school's students made progress similar to the growth standard.
- Yellow indicates moderate evidence that the school's students made less progress than the growth standard.
- Red indicates significant evidence that the school's students made less progress than the growth standard.

The tabular format of the table facilitates comparisons of growth across grades, across years, and—by tracking cells diagonally—by cohort of students up through grades over consecutively tested years. Thinking Heuristically About Student Growth in Austin Independent School District

achievement heuristic is adequate, because achievement and GPA generally share a positive association. Consequently, when GPA is used as a proxy for performance, the mental shortcut is a good fit.

However, if we cognitively map student growth to performance, then the heuristic leads us to flawed thinking about growth potential for achievers of varying levels. In the case of student growth, the achievement heuristic is inadequate, because student growth is independent of student achievement. Achievers of any level can show growth gains, losses, or growth to the standard. Consequently, when student growth is used as a proxy for performance, the mental shortcut is not a good fit. It may seem a simple act to check assumptions and altogether avoid the flawed thinking about a positive association between achievement and growth. However, to check assumptions is to deliberately and rationally process (engaging what is referred to as system 2, or thinking slow), which is at odds with the quick, heuristic-based decision making we often find ourselves doing (engaging what is referred to as system 1, or thinking fast; Kahneman, 2011).

The achievement heuristic also bares heuristic value for associated decision making. That is, high/strong is good, and therefore students with high/strong scores are doing well, and we only need to work to maintain their strong performance. Conversely, low/weak is undesirable, and therefore these students with low/weak scores are not doing well and need intervention and extra attention. Although a simple decision-making strategy, when mapped to STAAR, there is utility for decision making without getting bogged down in the STAARs' complexity. AISD principals are expert consumers of and decision makers with STAAR data, despite the STAARs' complexity. They possess this expertise likely without collective understanding of how STAAR items are developed, scored, or scaled, or where the cut points come from. It is also likely that some proficient consumers of STAAR data do not understand precisely what a percentile is, how a percentile differs from a percentage, or the formula for computing a percentile.¹ However, heuristically, AISD principals understand that high student achievement scores are better than low student achievement scores; they understand that STAAR measures the extent to which their students have learned the Texas Essential Knowledge and Skills (TEKS) curriculum standards; they understand how to interpret whether students are performing differently on STAAR according to student group, grade, or subject; and perhaps most importantly, they know how to make actionable campus-level decisions based on their STAAR results (see Figure 12).

Figure 12

Conceptual Representation of Student Achievement

Focus resources here...

Students with low achievement (differences by student group, grade, and subject) Celebrate and recognize here...

Students with high achievement (differences by student group, grade, and subject)

Low ← →High

Achievement

¹ A percentile compares the performance of a student with that of other students who took the same assessment. Any given percentile indicates the percentage of test takers whose scores fell at or below a score out of the total number of test scores. Percentages are often computed for tests by examining the number of correct answers on the test out of the total number of questions on the test. A percentile is based on the number of test takers who answered fewer correct questions out of the total number of test takers (or sometimes the number of test takers who answered the same number or fewer of correct questions out of the total number of test takers). Two common formulas are used to determine percentiles: the cumulative frequency below a score divided by the total number of scores, or the cumulative frequency below a score plus one half of the frequency at a score divided by the total number of scores.

Student growth also has the heuristic potential for principal consumption and to drive campus-level decision making. Without any explanation of what student growth is or how it's computed, heuristically, AISD principals can understand that more student growth is better than less student growth; they can understand that students can demonstrate gains (and unfortunately losses) independent of their achievement; they can understand that student groups may show differing extents of growth gains or losses by grade and by subject; and again, perhaps most importantly, they likely know how to make actionable campus-level decisions based on their students' growth (See Figure 13).

Figure 13

Conceptual Representation of Growth and Achievement

		subject)	subject) →High		
		(differences by student group, grade, and	(differences by student group, grade, and		
		Students showing low growth and low achievement	Students showing low growth and high achievement		
Gro	Less 🗲	Focus resources here	Recognize and question why these students are not being challenged here		
Growth	₹	(differences by student group, grade, and subject)	(differences by student group, grade, and subject)		
	→ More	Students showing high growth and low achievement	Students showing high growth and high achievement		
		Celebrate and understand what is working here	Celebrate and recognize here		

Achievement

If principals focus on both the growth and achievement of their students, then they will be better equipped to be the data-driven decision makers they already are. They only need to adapt and apply their existing decision heuristic from achievement data to growth and achievement data, while at the same time guarding against the false assumption that growth is dependent upon achievement. Growth and achievement data allow principals to go beyond asking how they raise the achievement level of their students. Growth and achievement data allow principals to learn from their highgrowth, low-achieving students; celebrate their gains; and make informed decisions about how to apply the successful best practices elsewhere. Growth and achievement data allow principals to see if they are still challenging and growing their highestperforming students. Growth and achievement data allow principals to understand if they are effectively growing their low-achieving English language learners (ELLs) proportionately to their low-achieving non-ELLs in English language arts assessments. Growth and achievement data better equip principals to make the decisions they are already making.

School Diagnostics Reports

The school diagnostic report provides information on the average academic growth of the school's students at different achievement groups based on student achievement in the selected subject and grade.

Achievement groups are formed by grouping students into one of five achievement groups, or quintiles, based on where their achievement falls in the district distribution of scores in the subject and grade.

- Achievement group 1 (i.e., the first quintile) groups students whose achievement was in the bottom 20% of all students tested in the district on the assessment.
- Achievement group 3 (i.e., the third or middle quintile) groups students whose achievement was in the middle 20% of all students tested in the district on the assessment.
- Achievement group 5 (i.e., the fifth quintile) groups students whose achievement was in the top 20% of all students tested in the district on the assessment.

The school diagnostics report allows principals to compare growth of the school's students across different levels of achievement. Understanding the statistical models underlying the SAS EVAAS growth measures is no more consequential to principal decision making than is understanding the psychometric properties of each STAAR assessment. The technical information for the growth models is available just as the technical information is available for STAAR, but the heuristicbased decision making is not dependent upon their mastery. Of greater practical importance than understanding the technicalities of student growth measurement is the take away that growth and achievement should be considered independent constructs that can and do operate independently of each other, while also providing different information about our student populations. SAS EVAAS is a way of measuring the year-to-year academic progress of students, defined here as the amount of growth students need to make to keep up with academically similar peers (i.e., the amount of growth achieved relative to the growth standard; Schmitt & Hutchins, 2016). STAAR achievement is a way of measuring the within-year, grade-level knowledge and skills of students relative to that of other students in the state. Used together to inform campus decision making and planning, principals are able to take action on both their students' current year academic achievement and their students' year-to-year academic progress.

References

Kahneman, D. (2011). Thinking, fast and slow. New York, NY: Farrar, Straus and Giroux.

- Schmitt, L. N., & Hutchins, S. D. (2016). SAS Educational Value-Added Assessment System (EVAAS®) scores for Austin Independent School District, 2015 (DRE Publication #15.07). Austin, TX: Austin Independent School District.
- Tversky, A., & Kahneman, D. (1974). Judgments under uncertainty: Heuristics and biases. Science, 185(4157), 1124– 1131.
- TEA. (2013). Chapter 101. Assessment, Subchapter DD. Commissioner's rules concerning substitute assessments for graduation. Retrieved from http://ritter.tea.state.tx.us/rules/tac/chapter101/ch101dd.html
- TEA. (n.d.). Substitute assessments standards. Retrieved from http://ritter.tea.state.tx.us/rules/tac/ chapter101/19_0101_4002-1.pdf

AUSTIN INDEPENDENT SCHOOL DISTRICT

Shaun D. Hutchins, Ph.D. Paige Hartman DeBaylo, Ph.D. Holly Williams, Ph.D.

Department of Research and Evaluation



1111 West 6th Street, Suite D-350 | Austin, TX 78703-5338 512.414.1724 | fax: 512.414.1707 www.austinisd.org/dre | Twitter: @AISD_DRE

June 2018 Publication 17.33

13