



EVAAS “Myth Busters”

Myth: EVAAS is too complicated to understand.

Fact: Basically, EVAAS compares the change in achievement of a group of students from one year to the next *with* an expected amount of change based on the students' prior achievement history. Simply put, “added value” refers to how much the student learned during the school year. While the *concept* of value-added is simple, the *application* is more complicated. Any value-added model should take into account variables such as errors in measurement; missing test scores; how tests are scaled; and educators serving different types of students. To use student test data in a *reliable* way, value-added models must be statistically sophisticated.

Myth: I can't understand how EVAAS scores are calculated, so I can't use it to improve my teaching.

Fact: You may not know how to build an iPhone, but you can still use it to do amazing things. While it's true that EVAAS calculations use complex statistics and multiple years of data, the reports on the EVAAS website (<https://hisd.sas.com/>) do provide usable information.

- Diagnostic reports on students based on their incoming achievement level provide data-driven support to ensure all students reach their potential
- Student-projection reports show teachers and administrators which students are on a positive or negative growth trajectory and need interventions or advanced coursework
- Teacher reports help administrators augment the strengths of effective teachers and implement professional development of teachers who have room for improvement
- District-level training helps teachers use EVAAS reports to become more effective

Myth: HISD's EVAAS system is unreliable "because the error bars [graphed data that indicates error] are so large."

Fact: Standard error is a measure of uncertainty around an estimate. Larger standard errors mean that there is more uncertainty about a teacher's influence on his or her students' academic progress. Factors that may influence standard error include:

- Range of scores in student performance (i.e. A larger range of scores will result in a larger standard error.)
- The number of students used in the analysis (i.e. Teachers with fewer students will have larger standard errors.)
- Range of time in the analysis (i.e. Single-year estimates are less stable than multi-year estimates and therefore have a larger standard error.)

Value-added models such as EVAAS can provide *fair, valid, and reliable* information to educators on students for a variety of issues.

- Fair because they do not depend on the types of students a teacher receives each year
- Valid because various challenges in student testing are taken into account and students serve as their own control
- Reliable because they are repeatable and consistent; multiple-year estimates protect teachers from misclassification



EVAAS “Myth Busters”

Fact: The term “error bars” is somewhat misleading in that the bars do not imply any mistake in the analysis, but has a very specific meaning in statistics. The bars actually serve to *protect* educators from misclassification (i.e. identifying teachers as ineffective when they are truly effective). The range of “error bars” for EVAAS teacher reports is consistent with what you would see in any model with similar sample sizes (e.g. class sizes).

In general, the range of NCE gains at the teacher level is around -15 to 15 in Houston. The standard errors range from 1 to 4 in HISD. This does not appear to be 30-50% (as purported by opponents to EVAAS) from what the data are showing.

We are able to differentiate a good number of teachers with the measures that we are using. Compared to the district, there are approximately 26% of teachers that either fall into well above or well below categories compared to the district average in any given year. Accounting for standard error (and using multiple years of data) keeps us from misclassifying teachers as well above or well below.

Myth: A teacher can have a low value-added score one year and a high value-added score the next, which means it is basically an unreliable measure of performance.

Fact: This can be true with single-year estimates because a teacher may be more effective with one year’s group of students over the next. Less effective teachers may also improve over time. For teachers identified as ineffective or highly effective based on three-year estimates, the results will be more consistent. If a teacher is ineffective for three years, the likelihood is higher that they will remain ineffective. Likewise, a highly effective teacher will continue to be effective.

Fact: Evaluation of EVAAS data (from another state) over a 14-year period provided valuable insight into how teachers may change in effectiveness over time.

- Highly effective teachers are very likely to remain effective.
- Teachers identified as highly effective after their first three years of teaching were extremely likely to remain effective three years into the future.
- For teachers identified as ineffective based on three-year estimates, approximately 50 percent will be ineffective three years later.
- Some year-to-year variation among individual teachers is to be expected, depending on the variation in student groups from year to year.

Myth: Teachers identified as “highly effective” based on value-added data do not necessarily produce more learning for their students.

Fact: Students with very effective teachers three years in a row were able to improve their performance on standardized tests by more than 50 percent in comparison to students who had ineffective teachers three years in a row (Sanders and Rivers 1996). A similar study conducted in Dallas ISD found similar results. A teacher’s impact on student learning lasts up to four years (Sanders 2005). As teacher effectiveness levels increase, lower-achieving students are the first to benefit (Sanders and Rivers 1996). If a student has an ineffective teacher for two years, this decrease in progress cannot be made up (Rivers 1999).



EVAAS “Myth Busters”

Myth: HISD administration is intentionally keeping teachers in the dark about EVAAS and has not provided information and training to explain how scores are calculated.

Fact: HISD began offering EVAAS training through a combination of face-to-face and online courses in August 2008. Since then, more than 8,100 employees – most of them teachers – have completed an EVAAS course.

Myth: HISD has decided that teachers can be terminated based on their EVAAS scores.

Fact: In February 2010, the HISD Board of Education added EVAAS scores to the list of 34 factors that can be taken into account in renewal decisions. Employees are offered individualized support and professional development opportunities throughout the year. Prior to non-renewal of a teacher’s contract, the district will review the teacher’s performance in accordance with the 34 factors. EVAAS scores are *never* the sole factor for nonrenewal.

Myth: Under a new teacher-evaluation system, EVAAS scores are 50 to 60 percent of the teacher-evaluation process.

Fact: In accordance with state law, HISD’s new teacher-evaluation system was developed at the campus level through the Shared Decision Making committees and at the district level through the District Advisory Committee. Taking into account recommendations from these committees, working groups *that included teachers and administrators* came up with proposals for the numbers used in EVAAS.

For the 2015 appraisal, the three TADS components have the following weights within teachers’ Summative Appraisal Ratings:

- Instructional Practice: 50%
- Professional Expectations: 20%
- Student Performance: 30%

The various types of Student Performance measures have different weights within the Student Performance final rating:

Measure Combinations →		VA + CG + Student Progress	CG Student Progress	+		VA Student Progress	+	Student Progress Only
Measure Weights ↓	VA + CG				CG Only			
Value-Added	20%	15%				20%		
Comparative Growth	10%	10%	20%		30%			
Student Progress		5%	10%			10%		30%
Student Performance SubTotal	30%	30%	30%		30%			30%



EVAAS “Myth Busters”

Value-Added data is never the sole measure of student performance, and is never weighted at more than 20% of the teachers’ appraisal.

Myth: EVAAS is a quartile system with 50 percent of teachers receiving a negative score every year, which means half of HISD teachers are deemed “ineffective.”

Fact: EVAAS is not a quartile system. EVAAS measures the growth of students at the classroom level relative to expected growth. Originally, the ASPIRE *Award* program used a quartile system, but it has largely moved on from this methodology for a variety of reasons. The ASPIRE Award program awards core teachers with EVAAS scores based on reaching a set standard (1.0 or higher).

Myth: EVAAS system developer Dr. William L. Sanders is not an expert in the teacher-effectiveness measurement field.

Fact: Dr. Sanders spent over two decades in the teacher-effectiveness measurement field as an academic, a researcher and a provider of value-added analyses and reporting. He was a professor at the University of Tennessee Knoxville and then director of the university’s Value-Added Research and Assessment Center. After retiring from that position, he became senior manager of value-added assessment and research for the SAS Institute in Cary, NC and was a senior research fellow with the University of North Carolina system. Dr. Sanders recently retired.

In 2006, Dr. Sanders testified before the House Education and Workforce Committee about the reauthorization of No Child Left Behind. That year, his value-added research was ranked sixth in the Education Research Center’s list of most influential research on the last decade’s national educational policy. In February 2007, Dr. Sanders shared his research in a Senate Health, Education, Labor, and Pensions Committee Round Table discussion on teacher incentives. Over the last 20 years, Dr. Sanders and his colleagues have developed and refined their methodology.

Myth: The EVAAS methodology has never been independently validated.

Fact: In the RAND Corporation’s 2003 report, “Evaluating Value-Added Models for Teacher Accountability,” the EVAAS model was studied for its validity in measuring teacher effects. The report concluded that models like EVAAS use an approach that “is likely to be preferable but is computationally demanding.” The full Rand report is available here: http://www.rand.org/pubs/monographs/2004/RAND_MG158.pdf.