FACILITATION GUIDE
VALUE-ADDED VIDEO SERIES

VIDEO 5:
UNDERSTANDING STANDARD ERROR AND GROWTH INDEX

For district, building, and teacher leaders supporting professional learning on value-added analysis

Ohio | Department of Education

Battelle for Kids

TABLE OF CONTENTS

INTRODUCTION

The Value-Added Video Series is designed for teachers and administrators as an easy, digestible way to acquire a basic, conceptual understanding of Ohio's value-added measures to foster the appropriate and effective use of value-added information.

To get the most out of the video series, you are encouraged to use the videos during collaborative learning times, and leaders may use the accompanying Facilitation Guide for each video to elicit thoughtful discussions among educators in a collegial, professional learning setting.

The series includes:

- Video 1: Introduction to the Value-Added Series (6:52)
- Video 2: Establishing Awareness: Why Growth Measures Matter! (24:30)
- Video 3: Getting the Basics: SAS® EVAAS® Multivariate Response Model (MRM), Mean Gain Approach (33:38)
- Video 4: Getting the Basics: SAS® EVAAS® Univariate Response Model (URM), Predicted Mean Approach (14:50)
- Video 5: Understanding Standard Error and Growth Index (17:28)
- Interpreting Value-Added & Diagnostic Reports (*Facilitation Guide only, no video*)
- Video 6: Analyzing and Using Value-Added Information (11:30—includes educators sharing their experiences)
- High School Value-Added Video-A Good Beginning to value-added Information (7:35)

Facilitation Guides include:

- Learning targets and length for each accompanying video
- Guided practice which includes reflection activities, questions, and scenarios
- Talking points and example responses to support the facilitator
- Additional resource suggestions for those interested in learning more

Thank you for accessing and engaging in these professional learning resources to support your understanding and use of value-added measures.

This Facilitation Guide and the accompanying Value-Added Video Series are provided through a partnership between the Ohio Department of Education and Battelle for Kids.

PLANNING FOR THE USE OF VIDEO 5: UNDERSTANDING STANDARD ERROR AND GROWTH INDEX

STEP 1

Access and review Video 5 and this Facilitation Guide to familiarize yourself with the content prior to using the resources with an audience

STEP 2

Decide when and who will be asked to participate in Video 5 professional learning. Video 5 is a continuation of content from Videos 3 and 4. The content is applicable to both the MRM and URM models and supports understanding of Ohio's value-added analysis.

Video 5 Learning Targets

- Understand that value-added reporting includes a growth estimate and an associated standard error
- Learn about standard error, and how it helps inform whether the value-added estimate is significantly different from the growth standard, or growth expectation
- Learn about the term Growth Index (sometimes referred to as Gain Index) used in value-added reporting to determine effectiveness levels

VIDEO 5

Video 5 Length: 17:28

STEP 3

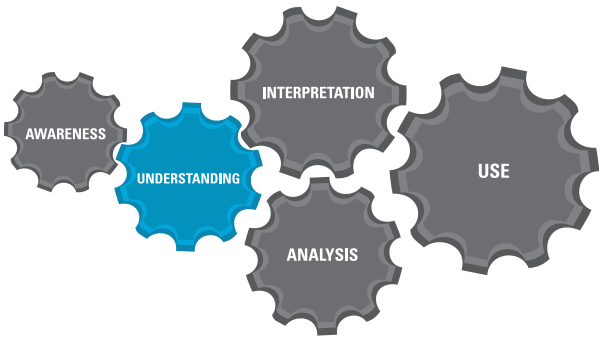
- Consider the Guided Practice activities in this guide, and plan how you will integrate these activities into the professional learning.
- Determine if you will show the video in its entirety first, or pause the video throughout to facilitate activities at suggested times. Suggested times to pause the video and engage in discussion or in an activity are included in this guide. The symbol to the right is included in the videos at any point where a pause is suggested for processing time or to check for understanding

STEP 4

Print the pages in the guide you plan to use as attendee handouts. Suggestions:

- Print pages 9–11 and 13–16 (back-to-back) for all attendees

Consider printing the pages in this guide marked "Facilitator Copy" to support your efforts. These pages include activity instructions and talking points.



GUIDED PRACTICE 1: STANDARD ERROR, GROWTH INDEX, AND PERFORMANCE LEVELS

PURPOSE

The purpose of this activity is for attendees to understand the concept of standard error that accompanies the growth estimate and how it helps inform the growth index and value-added performance levels.

SUGGESTED USE

1. **Prior to showing the video:**
   - Share Video 5 learning targets with the audience.
   - Print Guided Practice Handout 1, pages 9–11 (back-to-back) ahead of time for attendees.
   - Although it is suggested that this activity be completed in its entirety after the video, consider distributing the handout prior to the video so attendees can think about the questions while watching.

2. **After the video:**
   - Allow for at least 10 minutes for attendees to work with a partner or their table group on the questions in Guided Practice Handout 1.
   - If time permits, call on a few tables to share highlights from their conversation.
   - Share the Example Responses from your Facilitator Copy with attendees as needed.
   - If attendees would like more explanation than what was provided, suggest the Additional Resources provided at the end of this Facilitation Guide. The Technical Documentation of EVAAS® Analysis, located on the Ohio EVAAS® login site http://ohiova.sas.com, contains more in-depth information.

GUIDED PRACTICE 1: STANDARD ERROR, GROWTH INDEX AND PERFORMANCE LEVELS

EXAMPLE RESPONSES

Review the information provided on two Ohio schools from the 2014 value-added analysis, and respond to questions 1–3.

SCHOOL A

- 6–8 building
- 240 students
- Growth Measure: 3.0
- Standard Error: 2.5

SCHOOL B

- 6–8 building
- 480 students
- Growth Measure: 3.0
- Standard Error: 1.5

1. Although the growth measure is the same for both School A and School B, what may have contributed to the difference in the associated standard error for each school's growth measure.

- The difference in the number of students is the most likely explanation for this example. School A has 240 students and School B has 480 students.
- Another potential factor that would impact standard error is the amount of missing prior test scores. School A may have had more students with missing prior test scores. By 6th grade, many students in Ohio would have had prior state test data from grades 3–5.

2. What would the Growth Index be for both schools?

- School A: 1.2 (3.0/2.5 = 1.2)
- School B: 2.0 (3.0/1.5 = 2.0)

The growth index is the growth measure divided by its associated standard error

FACILITATOR COPY

GUIDED PRACTICE 1: STANDARD ERROR, GROWTH INDEX AND PERFORMANCE LEVELS

EXAMPLE RESPONSES

3. Reference Ohio's value-added performance levels (below). How would you describe the 2014 value-added result (the growth index) for both schools?

	Index: 2 or greater Growth is 2 standard errors or more above the Growth Standard (0) Significant evidence of exceeding the growth standard, expectation.
	Index: equal to or greater than 1, but less than 2 Moderate evidence of exceeding the growth standard, expectation.
	Index: equal to or greater than -1, but less than 1 Evidence of meeting the growth standard, expectation.
	Index: equal to or greater than -2, but less than -1 Moderate evidence of not meeting the growth standard, expectation.
	Index: Less than -2 Growth is 2 standard errors or more below the Growth Standard (0) Significant evidence of less progress than expected.

- School A: The growth index is 1.2 resulting in level 4/light green. This means that there is moderate evidence that gains in this school exceed the growth standard (or growth expectation). Remember that the growth standard is a comparison to the average growth of students across the state for the same grade/subject level.
- School B: The growth index is 2.0 resulting in level 5/dark green. This means that there is significant evidence that gains in this school exceed the growth standard (or growth expectation) or the average growth of students across the state for the same grade/subject level.

GUIDED PRACTICE 1: STANDARD ERROR, GROWTH INDEX AND PERFORMANCE LEVELS

EXAMPLE RESPONSES

Review the information provided on Teacher A's value-added results from 2013 and 2014, and Teacher A's Multi-Yr Trend that incorporates the results from 2013 and 2014. Respond to questions 4 and 5.

TEACHER A'S VALUE-ADDED RESULTS				
Year	Growth Measure	Standard Error	Index	Level
2013	-1.0	2.0	-0.5	Level 3/Average
2014	-1.5	1.75	-0.86	Level 3/Average
Multi-Yr Trend	-1.4	1.25	-1.12	Level 2/ Approaching Average

4. Notice that the Multi-Yr Trend for Teacher A resulted in an index of -1.12, or a Level 2 result. How is this possible given that neither the 2013 result nor the 2014 annual result from this teacher was at this level?

- The standard error associated with the Multi-Yr Trend is smaller. The Multi-Yr Trend represents a larger number of students—both Teacher A's 2013 students and 2014 students. The number of students increased, therefore the amount of data increased and standard error got smaller. The reliability of the multi-year growth measure estimate is greater.

Tip: Value-added reporting in the SAS® EVAAS® website, https://ohiova.sas.com, provides additional information such as diagnostic pattern reports each year by grade/subject area. Analysis of this information each year is important to identify areas of strength and opportunities for improvement. By analyzing and responding to the data each year, teachers should not be caught off guard when they see the results of Multi-Yr Trend information.

5. Discuss with your partner/table group:

Why is it important to look at both the growth measure and standard error information when interpreting and analyzing value-added results?

- Standard error represents the degree of statistical certainty in the growth measure estimate. It's possible for the standard error to be large because of a low number of students or lack of prior data. In this case, looking only at the growth measure would be misleading.
- Both the growth measure and standard error are used to determine the result or performance level.

FACILITATOR COPY

GUIDED PRACTICE 1: STANDARD ERROR, GROWTH INDEX AND PERFORMANCE LEVELS

Review the information provided on two Ohio schools from the 2014 value-added analysis, and respond to questions 1–3.

SCHOOL A

- 6–8 building
- 240 students
- Growth Measure: 3.0
- Standard Error: 2.5

SCHOOL B

- 6–8 building
- 480 students
- Growth Measure: 3.0
- Standard Error: 1.5

1. **Although the growth measure is the same for both School A and School B, what may have contributed to the difference in the associated standard error for each school's estimate?**

2. **What would the growth index be for both schools?**

HANDOUT 1

GUIDED PRACTICE 1: STANDARD ERROR, GROWTH INDEX AND PERFORMANCE LEVELS

3. **Reference Ohio's value-added performance levels (below). How would you describe the 2014 value-added result (the growth index) for both schools?**

	Index: 2 or greater Growth is 2 standard errors or more above the Growth Standard (0) Significant evidence of exceeding the growth standard, expectation.
	Index: equal to or greater than 1, but less than 2 Moderate evidence of exceeding the growth standard, expectation.
	Index: equal to or greater than -1, but less than 1 Evidence of meeting the growth standard, expectation.
	Index: equal to or greater than -2, but less than -1 Moderate evidence of not meeting the growth standard, expectation.
	Index: Less than -2 Growth is 2 standard errors or more below the Growth Standard (0) Significant evidence of less progress than expected.

HANDOUT 1

GUIDED PRACTICE 1: STANDARD ERROR, GROWTH INDEX AND PERFORMANCE LEVELS

Review the information provided on Teacher A's value-added results from 2013 and 2014, and Teacher A's Multi-Yr Trend that incorporates the results from 2013 and 2014. Respond to questions 4 and 5.

TEACHER A'S VALUE-ADDED RESULTS				
Year	Growth Measure	Standard Error	Index	Level
2013	-1.0	2.0	-0.5	Level 3/Average
2014	-1.5	1.75	-0.86	Level 3/Average
Multi-Yr Trend	-1.4	1.25	-1.12	Level 2/ Approaching Average

4. **Note that the Multi-Yr Trend for Teacher A resulted in an index of -1.12, or a Level 2 result. How is this possible given that neither the 2013 result nor the 2014 annual result from this teacher was at this level?**

5. **Discuss with your partner/table group:**

Why is it important to look at both the growth measure and standard error information when interpreting and analyzing value-added results?

VALUE-ADDED DATA QUALITY BLOG RESOURCE

PURPOSE

The purpose of this resource is for attendees to read the blog, reflect upon what they have learned about Ohio's value-added models, and discuss the key takeaways. Content of the blog includes information that was discussed throughout this value-added video series.

SUGGESTED USE

1. **Prior to showing the video:**
   - Print Resource Handout, *Value-Added Data Quality: Too Often, It's in the Eyes of the Beholder*, pages 13–16, (back-to-back) ahead of time for attendees.
   - Wait until after the video to pass out the blog handout.

2. **After the video:**
   - Pass out the blog handout and share the purpose of this resource as stated above.
   - Give attendees 10–15 minutes to read the blog.
   - Assign tables/groups to discuss one of the three discussion prompts found at the end of the blog resource:
     - How have value-added measures prompted you to a new way of thinking about academic measurement?
     - Which element on the VAM Top Ten resonated with you the most?
     - How can value-added information extend beyond teacher evaluations to inform professional practice?
   - Allow 10 minutes for attendees to work with a partner or their table group on their assigned discussion prompt. Depending upon the time left in the session, you might assign groups to work on more than one discussion prompt.
   - If time permits, call on a few tables to share highlights from their conversation.

VALUE-ADDED DATA QUALITY: TOO OFTEN, IT'S IN THE EYES OF THE BEHOLDER

Blog Post by Jamie Meade, 4/30/14

If you're keeping up with educational blogs, tweets, and other op-eds these days, you're likely encountering a barrage of opinion pieces on the use, and misuse, of value-added measures (VAMs). Since I count myself among those deeply committed to the impact of high quality data analysis and use, I'm especially interested in the recent heightened debate that has emerged around VAMs, now that they are being used in many educator evaluation systems across the country.

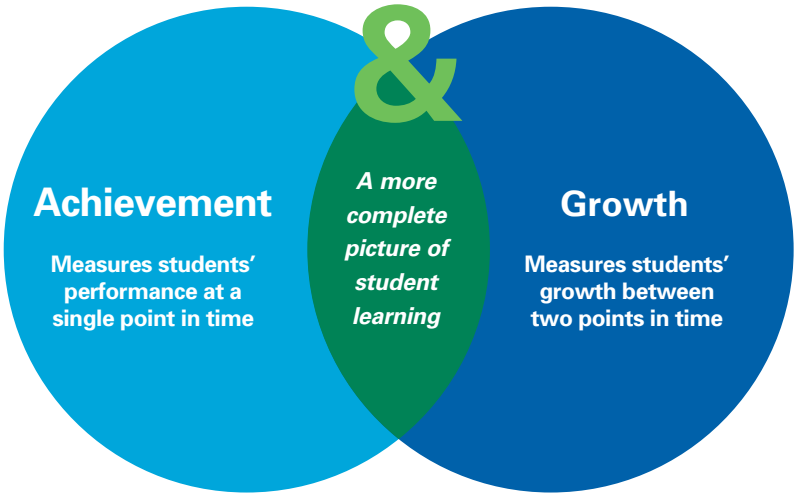
I was introduced to value-added measures more than a decade ago. Compared to the policy-driven manner of value added implementation experienced by many educators today, I consider my introduction to be more, let's say, on the softer side. I count myself among the educators who were early adopters, the pioneers of "using" value-added measures. You might consider this group a coalition of the willing. We were introduced to the differences between achievement measures and growth measures—and the Power of Two— and invested deeply in professional learning. **These early adopters of value-added measures welcomed an educational paradigm shift: a new way of thinking about academic measurement.**

I learned that value-added is not about measuring which students are "passing a test," and that robust value-added models are not about "one test, one day." These ideas resonated with me since I was, and continue to be, mindful that achievement measures (or passage rates on tests) are often influenced by variables unrelated to the effect of a school or teacher. As a former high school teacher, I recall the earliest of accountability measures introduced more than 20 years ago: achievement measures, which were based upon the percent of students who "passed the state test." For sure, this concept of measuring effectiveness of our schools and classroom teachers endures across our nation today—if not formalized in public policy, certainly in public perception and real-estate sales.

In the early days of my teaching career, I instinctively believed that other important measures were missing from the accountability landscape. **I knew then, and I know today, there are other measures that matter!** Learning and teaching is dynamic; and therefore, hard to measure. As such, all educational measures are imperfect, including classroom assessments used to generate letter grades and GPAs that often inform high-stakes decisions about where or if a student will go to college or the amount of scholarship support they might be awarded.

No single instrument can capture all that matters with regard to school and teacher effect on the present, and future, of a child. Only through a combination of multiple, valid, reliable measures can we begin to capture a more comprehensive picture of the difference our schools and teachers are making with students. So, perhaps the **dialogue surrounding quality VAMs is not as much about the pursuit of perfect measures, but rather, better measures used within a comprehensive system of measures.**

VALUE-ADDED DATA QUALITY: TOO OFTEN, IT'S IN THE EYES OF THE BEHOLDER



Today, the concept of value-added growth measures has taken root, and as a nation, we're realizing the **Power of Two: the powerful insights to be gained when we analyze and use academic data through the combined lenses of both achievement measures and growth measures.**

As student growth measures (SGMs) are now included in many educator evaluation systems across the country, a wide variety of growth measures, including VAMs, have emerged from multiple statistical providers. As a result, **multiple, varied statistical analyses have become clustered together under a common term, "value-added measures."** Local policy also influences the value added metrics—yielding very different results in districts and states across the country. **These variations in VAMs make it increasingly important that our nation's educators and policymakers become more data literate—and in a hurry!** National educational literature is brimming with references, accolades, and stern panic-inducing warnings, with regard to VAMs.

If unequipped with essential, fundamental data literacy skills, our nation's educators are left to respond to value-added information with two limited options:

1. Reject all value-added measures based upon the literature review out of mistrust or fear of becoming a victim of unsound measures used to evaluate, rank, file, and fire teachers.
2. Accept all value-added measures by entrusting policymakers and other influencers to adopt reliable value-added growth measures.

Clearly, **educators must become more assessment and data literate to inform their professional practice, accelerate student learning, and impact educational policy.** Rejection of all educational measures and data will only exacerbate the issues associated with poorly designed, unsound measures used to evaluate educators. Thus, our intensive national dialogue and debate around the quality of value-added measures.

So, how do we determine data quality with value-added measures? The answer is often in the eyes of the beholder and likely depends upon the purpose, or intent, of the measure. Often, the purpose of the VAM drives the design of the measure. Through the inclusion of policy rules or local preferences applied to the analytics, the outcome of a teacher value-added score or label associated with the value-added effect can differ

VALUE-ADDED DATA QUALITY: TOO OFTEN, IT'S IN THE EYES OF THE BEHOLDER

Educator evaluation systems and value-added measures designed only to be summative measures, that is, to serve the sole purpose of scoring, ranking, or labeling teacher performance, are often designed and implemented without regard for how the data can or may be used to inform and support teacher practice. For example, if the design of the VAM is limited to the purpose of generating a value-added percentile score to rank teachers; then, one may have a different idea about quality value-added measures. On the other hand, value-added measures, designed with regard for potential formative use of the results, produce data and robust reporting that can be interpreted and used to generate valid inferences about professional practice and student learning. In this scenario, the quality of the VAM can be defined by the degree to which the data and associated reporting can be used to diagnose strengths in professional practice and opportunities for professional growth.

At Battelle for Kids, we've had over a decade of experience supporting educators with the appropriate design, implementation, and use of value-added measures to support school improvement efforts, teacher practice, and ultimately, accelerate student learning. We've learned that VAMs can, and should be, designed to support teachers and students in their learning! We've learned that educators need not be statisticians to be sufficiently data literate to interpret, analyze, and use value-added measures designed for formative use. **We believe the question is not "whether or not to measure student academic growth." Rather, the question is "how best to measure student academic growth."** While not intended to be all-inclusive, here's our top 10 list of the most important elements of value-added measures you can trust—and really use!

1. Since VAMs are not comprehensive measures of all that matters, **they must be used in combination with other measures** to inform decisions.
2. At best, **VAMs are estimates** and must be accompanied by measures of precision (standard error) and regarded within ranges of statistical certainty.
3. Since VAMs are estimates, they become **more reliable with multiple years of data**.
4. Strong, robust VAMs are based upon the growth of a group of students, using historical testing data.
5. A **strong, reliable roster verification process is essential** to improve accuracy of a teacher's class roster of students used in the value-added analysis.
6. **VAMs are only as reliable as the academic assessments used to generate the analysis;** sound assessments that generate data from which valid inferences can be made are essential for reliable VAMs.
7. Value-added measures built upon strong academic assessments are further **enhanced when accompanied by complete, transparent item-analysis data from these sound assessments.** When assessment items connect back to articulated expectations for student learning, teachers can then use the data to reflect upon instructional delivery of particular lessons or units of instruction.

VALUE-ADDED DATA QUALITY: TOO OFTEN, IT'S IN THE EYES OF THE BEHOLDER

8. Labels, or descriptors, associated with school and teacher effect in VAM reporting **must be accurate—and carefully chosen**—so as to elicit an appropriate response by educators and other stakeholders.

9. Robust, reliable VAMs **must be thoughtfully implemented with due regard for educator preparation** (data literacy support) for how to access, interpret, analyze, and use the data to accelerate student learning and improve professional practice.

10. VAM's **must be implemented by data literate school leaders** to ensure conclusions are not drawn from simple correlations. Robust, reliable VAMs can detect causation, but are not strictly causal.

I entered the teaching profession because of my strong desire to make a difference, and I continue to be deeply committed to that cause. **While not all that matters can be measured, high-quality educational measures, including value-added measures, can provide essential information to advance our efforts in understanding what works in education.** Thoughtfully designed growth measures, used in combination with other measures are important to ensure educators have access to information about what works, so that we may sustain and replicate these practices across our schools in the interest of making a difference for students.

Jamie Meade is Managing Director of Learning and Leading at Battelle for Kids.
Follow her on Twitter @meade_jamie.

BLOG DISCUSSION PROMPTS

How have value-added measures prompted you to a new way of thinking about academic measurement?

Which element on the VAM Top Ten resonated with you the most?

How can value-added information extend beyond teacher evaluation to inform professional practice?

NEXT STEPS & ADDITIONAL RESOURCES

STEP 1

Follow up with attendees:

- Show them where to access the resources from Video 5 of the Value-Added Video Series located on the Ohio Student Progress Portal: www.BattelleforKids.org/Ohio.
- Communicate the plan for additional professional learning that will lead to effective use of value-added information.

Tip: Now that attendees have a basic understanding of Ohio's value-added models, they may want to move into report analysis and use. It is recommended that before moving into this stage with real data, they be exposed to sample reports and practice report interpretation. Sample reports and interpretation activities can be found in the next facilitation guide: Interpreting Value-Added and Diagnostic Reports.

STEP 2

Does the staff need more time with the content from Video 5?

- Share the Additional Resources information, below, for more professional learning with the entire staff, groups/teacher teams, and/or individuals before proceeding with the Value-Added Video Series.

ADDITIONAL RESOURCES

The following are free to Ohio educators and available for use to go deeper into the content discussed in this video and guide.

- The Technical Documentation of EVAAS® Analyses can be found on the Ohio EVAAS® login site at: https://ohiova.sas.com.

- Online Value-Added modules:
  o Coming spring 2016: New courses on Understanding the MRM and URM Analyses

  Log into the Ohio Student Progress Portal to access online courses through "MyLearning" and Enroll in Learning: www.BattelleforKids.org/Ohio

Many online courses already exist for Ohio educators including courses on both value-added and formative instructional practices.