VAM and Evaluation Results
Transcript

Slide 1: Good afternoon and welcome: primary accountability contacts, VAM contacts, evaluation system contacts, district personnel directors, and district staff development directors.

Slide 2: "We will not evaluate ourselves to greatness, but we need evaluation to help coach and support teachers to greatness." Commissioner Pam Stewart

Slide 3: Goal of evaluations – to provide honest and actionable information to support continuous improvement. One way we can do that is to provide data like the red and green sheets to school and district leaders and to have discussions about effective ways to use this data.

Slide 4: Our objectives today are to share new educator evaluation data, offer ideas for how to use this data and describe how participants can ask questions.

Slide 5: We began providing 'red/green' sheets to districts in 2013-14 as a way to help districts and principals begin having data-informed conversations to improve teaching and learning within their schools." We will post these to sharefile after this webinar. Once uploaded there, district primary accountability contacts, or their designees with access, can download them and share them with you.

As we go through the presentation if you have questions you may submit them using the chat function. At the end, we will go through these and answer as many as we can in the time remaining.

Slide 6: It is important to remember that VAM, and indeed any measure of student performance, is only one component of a teacher's evaluation. As a result…

Slide 7: Here you can see a snapshot of what a red/green sheet looks like at a glance. We will go over the individual fields in the next several slides, but here you can see that there are 4 different colors represented. Green represents VAM scores that are classified as HE, yellow represents VAM scores that are classified as NI, and red represents VAM scores that are classified as U. Keep in mind that these classifications are based on the new SB rule that goes into effect for 2015-16 and MAY NOT represent how the score was classified in 14-15 by the district using locally negotiated cut points. In addition, the classification of the VAM score may be different that the final summative rating the teacher receives.

These sheets will only contain teachers who have BOTH a VAM score and a final evaluation rating for 2014-15. Teachers missing either one will not be on the list. Teachers who teach at multiple schools within the district will appear on the list for each school they taught at.

Slide 8: These fields are all pretty self-explanatory.

Slide 9: Here we see the fields that contain the VAM scores and other important information used to interpret them for 4 different teachers. Based on the color, you can see that each is classified into a different performance category using the VAM data.

Looking at this first column, this represents up to 3 years of READING (although this became ELA in 2015 and will be reported that way next year and going forward) data across all grades the teacher may have taught reading in. The score itself is standardized so that it can be interpreted as the percent above (or below) average or expected learning growth achieved by students on the teacher's roster. When compared to how other, similar students performed on the same assessments. So by moving the decimal over two places, you can convert this score to the percentage that it represents. Positive scores represent above average learning growth, while negative scores represent below average learning growth. You can see right away that the green line is positive, representing 62% above expected learning growth, while both the white and red lines are negative, representing below expected learning growth of 33% and 50% respectively. The yellow line is blank in this field because that teacher did not teach reading/ELA and so did not receive a reading/ELA score.

Note there are also Math and Combined scores. These are interpreted the same way the reading score is, and represent Math and the combination of both subjects together respectively. And in the case where only one score (reading or math) is included, that score will be the same as the 3 year combined aggregate. As a result, the 3 year combined aggregate is the official VAM score that will be used going forward because it contains the most complete set of information about the teacher available across all grades, subjects and years.

In addition to these 3 scores, each one also has an associated standard error. The standard error is a statistical term that works the same way a margin of error does. It is an measure of the degree of precision surrounding the impact the teacher had on student learning growth as measured by the VAM score. The standard error is impacted both by how many students are included in the calculations, and how consistently students perform relative to their expected scores. The more students used to compute the VAM score, the lower the standard error tends to be. And the more consistently students perform with respect to their expected scores, the lower the standard error tends to be.

The standard error is important to consider because without it we may draw unwarranted conclusions about performance. In the examples above, when you look at the 3 year aggregate combined VAM score, 3 of the 4 scores in this table are negative. But the standard error provides us with important information about how different the teachers really performed, statistically, with respect to their student's average learning growth.

Slide 10: These columns illustrate the upper and lower limits of the confidence intervals which are used to classify the VAM score into one of the 4 performance level categories. They are computed using the 3 year combined aggregate and the standard error. Once computed, the sign is all you need to know to make the classification distinction. When all four of these are positive, the score will be classified as HE. When all of them are negative, the score will be classified as U. When both lower limits are negative and both upper limits are positive, the score is classified as E, and when only the 95% upper limit is positive, the score is classified as NI.

As you can see, all 4 of these scores are classified differently meaning that the teachers' impact on student learning growth was different statistically, and in the case of the green and red, in very important, statistically significant ways.

However, when you look at the final ratings these teachers' district awarded them, they all received the exact same final summative rating. This means that, absent a walk-through of this data, these teachers may be unaware of how different their impact on their student's learning growth actually was.

Though to be clear, there is no expectation or requirement that the final rating and VAM classification necessarily align; the fact that there is no differentiation among these teachers means that observational data is not providing useful information that differentiates among effective teaching practices that are reflected in better student outcomes. Though they are different measures, observation and student performance data should tell similar stories. If not, then they are not measuring the same things. And since we know the VAM data is measuring the teacher's impact on student learning growth, it becomes unclear what the observation might be measuring instead.

Slide 11: We want to be thought partners with you to continue to think about ways to use student and teacher performance data to inform and improve teaching and learning. We've listed some of the levers here.
- An obvious thing principals can do is to identify teachers who may be able to share best practices with their peers. These teachers might be willing to be mentor teachers or to open up their classrooms for other teachers to see what they're doing that is having such a positive impact on student learning.
  - Many districts have at least several principals who have learned how to facilitate learning walks in non-threatening ways through the Commissioner's Leadership Academy. I'd be happy to share those principals with you.
- Another thing principals can do is to identify themes present in classrooms of HE and U teachers. This can be as simple as a principal asking his/her leadership team…(next slide)

Slide 12: What do we know about teachers, our teachers whose impact on student learning is statistically, significantly positive AND What do we know about teachers, our teachers whose impact on student learning is statistically, significantly negative.

Slide 13: We can also us the data from the Red/Green sheets to:
- Make data-informed decisions about teaching assignments at the school and at the district level.
  - You might look at your elementary school teachers who teach math and language arts, for instance, and discover that a an "Effective" teacher had a significantly positive impact on student learning in math and a negative impact on student learning in language arts.
  - One of the teachers we recently recognized as a "High Impact" teacher for having a significantly positive impact on his student's learning is teaching PE this year. This teacher teaches at a low performing school. I bet the principal wishes he/she knew about this teacher's impact on student learning.
- One last thing we'll share here is to use the red/green sheets to review differences between VAM score classification and final evaluation among teachers in your district.

Slide 14: We know the most about those teachers whose impact on student learning is significant – at either end.

Not every teacher in red is a bad teacher. In fact many are not. We know of countless stories that shed some light on factors that contribute to teachers not having the impact they'd hope to have on student learning:
- South Florida principal whose "red" teacher was teacher of the year, also a coach, SGA sponsor, NHS sponsor – was out the day I was there.

- Central Florida district leader whose teacher when he was a principal was red, she was an expert – out of class all the time supporting others.
- Central Florida principal who discovered that two best friends, both excellent teachers, one green one red, were a month apart on their pacing guides.

The key takeaway for me with all of these examples is that these principals wouldn't have even known about these teachers' negative impact on their students' learning had they not see their VAM scores.

Slide 15: Remember that these red/green sheets can be obtained by your district's primary accountability contact, or anyone they have designated access to your district's sharefile account to. If you do not know who the primary accountability contact is in you district, you can email us and we will provide it to you.

Slide 16: www.FLDOE.org