

District TA Follow-up Call on Value Added

September 30, 2011

10:00 a.m. via Conference Call

Agenda

Intro – Chancellor Grego

Thanks you to all districts for participating in August 1-2 technical assistance meeting in Orlando. All meeting evaluations were positive. Our contractor, AIR, and DOE staff did a great job. Depending on today's call and your needs we may have more calls. We were asked to produce the state average growth by grade and subject and that has been provided to districts. Juan has some updated info on aggregated value-added (VA) scores that were also requested. We will continue to improve upon the model.

Kathy Hebda reviewed the agenda for today's call.

I. "Frequently Asked Questions"

A. Regarding September 30th deadline:

Race to the Top districts need to submit the following by September 30th

- A letter signed by the superintendent verifying that implementation for 2011-2012 will begin in accordance with the district's approved system, including completion of any school board approval and/or collective bargaining needed; This varies by district and is determined locally. September 30 is Year 2, Quarter 1 deadline for grant purposes and we have to follow-up with USDOE.
- Final version of the 2011-12 evaluation system documentation that was submitted June 1, reflecting any changes made since then (related to any of the criteria on the checklist), and
 1. How you are using the state's value added calculation and the student growth rating criteria for all teachers (methods of "classification" and "aggregation" and the performance scale or cut points) for various groups of teachers
 2. How the student growth component combines with the instructional practice portion and any other metrics to determine the summative rating.

Questions

Paul Felsch, Leon: Must we set the cut points set by Sept 30 or just provide assurance that they will be set?

Kathy: Set the cut points.

Paul: We could have different cut points for diff groups of teachers.

Kathy: Yes, because teachers are using diff assessments

Paul: For example, we are using Algebra 1 EOC and we don't know passing scores yet.

Juan Copa: The information submitted 9/30 can be revisited by districts even for 2011-12 as new information is known. This is not the end of the road but a key milestone to ensure districts are implementing in 2011-12. For Algebra 1 EOC, local assessments, etc. plans can be amended for 2011-12.

Chancellor: The same thing will happen with other EOCs as they are implemented; this is a process.

Marc Baron, Palm Beach: does the timeline apply to non-RTTT districts?

KH: By December 1, 2011, you must submit to us fully revised instructional practices piece, and this is the same deadline as the student growth piece for you.

- B. Juan Copa provided information on the school component and its relationship to what's in the data file

Using the elements in your data file, the formula is:

Teacher Value Added Score = Teacher Effect + 0.5*School Component.

The key part to understand is the Teacher Effect is the teacher's performance relative to the school and the School Component is the school's performance (in the particular grade and subject area) relative to the state.

Thus, if a teacher has a teacher effect of 6 and the school component is 2, the teacher's value added score is $6 + (0.5*2) = 6 + 1 = 7$. (Example 1)

On the other hand, if a teacher has a teacher effect of 6 and the school component is -2, the teacher's value added score is $6 + (0.5*-2) = 6 + -1 = 5$. (Example 2)

That leaves the impression that a positive school effect "helps" the teacher's score and negative school effect "hurts" the teacher's score. However, it is quite the opposite.

Remember, the key part to understand is the teacher effect is the teacher's performance relative to the school – her deviation from the school component. Thus, in Example 1, the teacher's students performed 8 points above state expectations (school component is 2; teacher's deviation from that is 6; thus in reference to state expectations, the teacher's students performed 8 points above expectations). In Example 2, the teacher's students performed 4 points above state expectations (school component is -2; teacher's deviation from that is 6; thus in reference to state expectations, the teacher's students performed 4 points above expectations).

Those scores (8 and 4) would be the value added scores for the teachers if the model fully attributed the school component to the teacher – in other words, if the theoretical framework was that after controlling for the factors in the model, all else is attributable to the teacher. However, the committee felt strongly that the school itself exerts an influence on student learning as well – and that school impact is partially driven by the teachers at the school and partially driven by other factors at the school (e.g., leadership, school environment). It is basically a question of attribution. The students performance relative to the state is the observed outcome, but by parsing out a teacher

effect and a school component one is determining how much is of that observed outcome is the teacher held responsible for.

Therefore for the teacher in Example 1 – her score is 7, not 8. The performance of the students is driven by the teacher, but also impacted in part by the school – and the teacher is not fully held accountable for that.

In Example 2, her score is 5, not 4. The performance of the students is driven by the teacher, but also impacted in part by the school – and the teacher is not fully held accountable for that.

The PowerPoint presentation from the June 7 Student Growth Implementation Committee is a good resource for this issue.

<http://www.fldoe.org/committees/pdf/June7Materials.pdf>

Chancellor: If districts have questions after looking at website, call Juan. By talking through them I became a better student and could move forward. We have had successful conference calls with districts in 15-20 minutes. This is not easy stuff. We want to keep it simple but that may not always get us to our goal. We are following direct recommendations of the SGIC.

Juan: This is absolutely a challenging concept as documented in committee deliberations. We're here to provide as much understanding and assistance as you need.

C. Juan discussed aggregation options

Two options presented August 1-2. First, beginning on slide 47 for the meeting PowerPoint (<http://www.fldoe.org/committees/pdf/august12tampres.pdf>), this method requires districts to standardize VA scores on common metric. Growth at different grade level varies. VA score of 10 in grade 4 may mean something different than a 10 in grade 10. Standardizing scores is required. Data provided August 1-2 is one option and is based on average amount of growth per grade level and subject. AIR will perform that calculation, and provided tools August 1-2, but we can ease the burden and AIR will provide aggregated score for each teacher even if he/she teaches across multiple grade levels over multiple years. One score of combine reading performance, one score combined math performance. We had hoped to provide last week but earthquake and hurricane delayed AIR delivery. AIR assured delivery COB today. Hope to have to districts tomorrow. Advantage of this approach is you set one scale, one set of cut scores for reading and math on FCAT.

Second option deals with differences by grade level and subject by classifying teachers independently within grade level and subject. If one teaches multiple grade levels and subjects, they will get a classification for each grade level and subject grouping. Summative rating – weighted average of all classifications. Some districts are moving this way. Added complexity of multiple scales but intuitive and easy to understand.

Third option that some districts are pursuing is to use a piece of data in the file – the percentage of students meeting or exceeding expectations. Easy to understand, metric similar to learning gains we are accustomed to, but you are losing some of the information that VA score provides like magnitude by which students are moving. Could have teacher with 100% of students meeting or exceeding expectations by only just barely and their VA score would be close to 0. Could have teacher where 50% of students exceed expectations but they are exceeding by wide margin. VAM score won't necessarily correspond to % but % provides some info. May be a transition step.

Application of standard error. Presented August 1-2. There have been some questions as it relates to 4 different performance levels. Set cut point based on nominal score (VA score) to determine whether in top 2 categories or bottom 2 categories. Then use standard error to determine extreme categories. You have some degree of statistical certainty. If a teacher's score still clears cut point with standard error applied, then you have better sense that teacher is actually highly effective.

II. District Questions

Marc Baron – Palm Beach: I have emailed questions on frequency distribution and standard deviations. Will we get that? Want all frequency statistics we can get.

Juan: Will get that to you for state level data.

Richard Itzen – Lee: Are districts considering giving separate rating per year and then aggregating? If we did, do we go through standardization method but not across years?

Juan: Yes, this is hybrid between option 1 and 2. Option 2 would require setting different cut points for different grade levels and subjects and possibly across years. You could standardize each year independently so set one scale for a subject. Classify independently and aggregate across.

Richard: would we only be weighting, for example, teacher who teaches 2 grade levels in the same year?

Juan: Don't do simple average but weight by # of students in each category. If taught twice as many 5th graders than 4th graders, weight 5th more heavily.

Richard: The state has included any student associated with a teacher in survey 2 or 3, correct?

Juan: Yes, in your historical data.

Richard: Is that a recommended way of doing it or can district set different criteria for including students in calculation?

Juan: we are working how best to facilitate if districts did that. Keep in mind with Survey 2-3 match issue, the model controls for student attendance and mobility on an individual basis. Different than school grades. If student only there for 1 survey student will have attended fewer days than someone there both surveys, and likely higher mobility.

Driven by course level and, especially at HS, more semester courses. That's why 2 or 3 – don't want to throw out students just because course is a semester.

Richard: attendance isn't course or teacher specific right?

Juan: Daily attendance in school as a whole. Began discussion of starting new collection of attendance at course level but not for 2011-12. Models run on historical data driven by daily attendance.

Richard: If a student was with me for survey 2, then left before survey 3, would the model show he had fewer days with me?

Juan: We will verify and get back to you.

Chancellor: We will distribute questions and responses to all districts.

Paul Felsch – Leon: Talk about corrections process for data in the past or this year, like we do for FCAT. Is it true that we assign proportion of value to each school if student at one school 2/3 of year then another 1/3?

Juan: We are working in conjunction with survey 2 and 3 a process for roster verification. Current thinking is that processes for school grades to verify demographic data and assessment match will continue, Envisioning roster verification system will serve as membership file for this purpose. Once final closure of assessment corrections for school grades we would match against verified rosters that will serve as basis for VA for 2011-12. Do not have plans in place to go back to historical data to verify. What many districts are doing, since SB 736 is for 2011-12, is they are treating 2011-12 as year 1 of the evaluation. Moving teachers to 40% of evaluation and then building out. Historical data is a reference point to help see how model works and inform classification decisions.

Paul: Linked directly to survey or another data element?

Juan: Survey 2 and 3 will feed system but roster verification will be a stand alone process.

Paul: How will teachers be involved, will they interface with stand alone?

Juan: We are part of a project with Gates Foundation, working with pilot districts, to create web application where teachers view rosters and indicate whether students are in their class. Final correction and sign off would take place at school or district level. Working on functionality to print and hand deliver for sign off. Another option is to provide districts with file format and verify and upload certified roster. With any option, we don't envision teacher making the correction. Current model does not parse out students proportionally. Each student has different characteristics. Assigned to both teachers. Will verify.

Chancellor: If you have a student absent 45 days, the model levels the playing field – their expected learning outcome is lower. Same with mobility. Expect student with high mobility to have fewer learning gains. Bring to both teachers.

Paul: Both teachers get full gains even if students only there for a month?

Chancellor: Yes, but with controlled variables.

Juan: The model shows students with fewer days or high mobility have lower expectations.

Owen Roberts – St. Lucie: FEA notified us by phone that we will not be getting school VA data this year. Is that correct?

Juan: It will be provided this year.

Raymond – Escambia: We are probably going with option 3, take to student level and determine predicted and actual score. Does predicted score have standard error that we could use to adjust score downward?

Juan: I don't think you want to go that route go with prediction and met/exceeded, those % calculated for you. **Will verify**

Raymond: If you're within a % of expectation you are considered effective, etc., but seems more accurate to adjust at student level.

Amy – Pasco: Some of us participated in a focus group related to database on DOE website that will provide VA scores for teachers in October. Is that still in the plan?

Juan: As part of AIR contract they will develop reporting system of VA results. We are moving in that direction cautiously and there will not be a full blown release based on historical data in October. Working on developing that reporting system, go slow, pilot, resolve issues, but yes eventually will have reporting system that provides public information – not necessarily at teacher level, but how districts doing. SB 736 includes reporting requirements like performance data related to evaluation system.

Sylvia Jackson – Gadsden: Are teacher effect and school effect provided by state? Tell us more about approach 3, first 2 seem quite complicated and we might need full-time statistician.

Juan: Teacher VA score includes teacher effect and school component. That info is provided in files received August 1-2. State will always calculate that. Option 3 looks at whether students meet or exceed expectations. If so, there is a gain. Percentage is sum of all students who met gain divided by sum of all students teacher taught by grade and subject. Uses VA formula to inform that %. Some districts like it because intuitive and it's a %. Statisticians ask me to point out – you're taking VA score that provides range of how much teachers are moving students and reducing it to a Yes/No determination. You lose information of how far teachers are moving students. But easy to understand. Some districts we know are going this route. It's a possibility at least for a transition.

Sylvia: If option 3 could we take into consideration attendance and mobility of students assigned to teachers?

Juan: It's already taken into account. Determination of gains made are individualized based on characteristics included in model.

Sylvia: We will ask for a follow-up call with you.

Chancellor: Please do.

III. Wrap Up (Juan and any final comments by Chancellor Grego as needed)

Chancellor: We have the opportunity to make this work or stay confused, let's make this work. This will be transition year for most districts. We are learning also, trying different ways, benefit of pain through transition. These conversations are helping us all. We are looking into your great questions. This dialogue will make the model stronger in the end.

Kathy: We will send notes out. Submit September 30th information via e-mail to Racetop@fldoe.org.

Chancellor: Deadlines established by us and USDOE. The deadline was originally Sept. 1 and was extended to Sept. 30. We appreciate hard work of districts.

Jason Cochran - Polk: when and how receive information on applying the value-added results to non-FCAT subject teachers?

Juan: We have had discussion with AIR on best method to apply VA results to come up with score for those teachers. Likely focusing on 2011-12 results.

Kathy: Will provide an update as soon as possible.

John Boyd – Osceola: Please restate requirements due September 30.

Kathy restated requirements.