

FLORIDA DEPARTMENT OF EDUCATION
RULE DEVELOPMENT WORKSHOP
PUBLIC INPUT ON DRAFT RULE TEXT

Rule 6A-5.030: District Instructional Personnel
and School Administrator Evaluation
Systems

Rule 6A-5.0411: Calculations of Student Learning
Growth for Use in School Personnel
Evaluations

Tuesday, February 26, 2013
Florida Gulf Coast University
Sugden Resort & Hospitality Management Bldg.
Room 110
10501 FGCU Boulevard South
Fort Myers, Florida

MARTINA REPORTING SERVICES
Courtney Building, Suite 201
2069 First Street
Fort Myers, Florida 33901
OFFICE (239) 334-6545
FAX (239) 332-2913

I N D E X

PAGE

Presentation by Kathryn S. Hebda,
Deputy Chancellor for Educator Quality,
Division of Public Schools

3

SPEAKERS:

Dianne Mahlman

Donna Mutzenard

Linda Corn

Tom Conner

1 MS. HEBDA: I'm Kathy Hebda. I work in the
2 Florida Department of Education, and welcome to the
3 Rule Development workshop for two different state
4 Board of Education rules we're going to cover today.
5 And I think everybody that's here has three handouts,
6 two different rules, two different rules which are
7 proposed, proposed rules, and then there's also a
8 copy of the PowerPoint that I'm going to walk
9 through.

10 Just a few instructions about how we're going to
11 work this, what our agenda will be today for the
12 workshop. We're going to do this in three parts.
13 The first part is going to be an overview PowerPoint
14 to orient you to a couple things. What is -- what
15 the law says we have to write rules on; so what we're
16 supposed to cover in this rule. And then also to
17 give you an idea of what we are intending to say by
18 the draft text that's proposed here for your
19 consideration. And we do that because one of the
20 things we need feedback on is whether or not we
21 actually got that intent by the words we used. You
22 may look at something and say, well, you said you
23 were trying to say this but that's what it says to
24 me; I think you should say it this way. So that's
25 one of the things we want feedback on.

1 The second part after that will be clarifying
2 questions and answers. And some of the Q and A just
3 happens naturally during the overview presentation,
4 and that's completely fine with me. The whole part
5 -- the whole reason for parts 1 and 2 is so you can
6 ask any questions you have or things you don't
7 understand so you have a really good understanding of
8 what's in front of you and then you can know what it
9 is you're reacting to when you give the feedback.

10 Then the third part is comments. And by the
11 time we get to the third part for your comments after
12 all the Q and A and everything else is over with --
13 and we don't debate comments. People get to make
14 whatever comments they want. They stand on the
15 record as they are and nobody challenges you or asks
16 you questions about your comments. And we have a
17 court reporter here today because whenever we do
18 face-to-face meetings we always get a transcript.
19 Then we can post the transcript online just like we
20 did --

21 Hello. welcome. We have some handouts right
22 over here and the sign-in sheet.

23 When we did the two webinar workshops week
24 before last the recording of those webinars is online
25 now, so you can see what was said and who asked

1 questions and what they were and what the responses
2 were. So there could be questions that you ask that
3 we won't know the answer to, but we'll take those
4 back and get you a response if we don't know the
5 answer today.

6 Now, we're a little worried about participation.
7 There's several kinds of things we're hoping you will
8 give us feedback on. Even it's not today, we have a
9 spot on line where all of the, as I said, those
10 recordings are and copies of these rules are where
11 you can just click a button and give us input any
12 time. And that's going to be open all the way
13 through next month. So even if there's something you
14 don't think of today, you go back and you talk about
15 it with your colleagues and anything else and you go,
16 oh, I want to tell them that, you can always do that
17 online. You don't need to wait for a workshop to do
18 that.

19 There are a couple things we're hoping we will
20 get out of this either today or in the future. And
21 if you would, so that we have a good record, when you
22 are making -- it doesn't matter as much for the Q and
23 A, but when you are actually making your comments
24 that's what the speaker card is for, and if you would
25 hand that to the court reporter, then she can make

1 sure she spells your name correctly.

2 And then it looks like we don't have that many
3 people here, so we won't have to time anybody's
4 comments to make sure everybody gets to speak.
5 You'll probably have as much time as you want at the
6 end to say whatever you like. But even if you want
7 to just say a few words today and you want to give us
8 something written that we take back and enter that in
9 the record, we can do that too; however you would
10 like to do it.

11 So this is what I was explaining a couple
12 seconds ago. We're hoping that you give us
13 suggestions for just like anything else you would do
14 for editing. Take this out; I don't agree with this;
15 take this part out; add this part in because you
16 forgot to talk about this and I think you ought to
17 take this into consideration; or I see what you're
18 trying to say here but here's how I would word it to
19 make it make more sense. Any of that. And then we
20 ask that if you're -- if you do it online or if you
21 do it today or any other time, sometimes it also
22 helps for us to know the context of why you're saying
23 that because if you're saying that for a particular
24 reason and it's in one part of the rule, it could
25 also show up someplace else and we want to make sure

1 if we change it one place everything else makes
2 sense. So you don't have to say that, of course, but
3 if you want to explain, sometimes that's very helpful
4 to us.

5 And then also, as I said, we're doing part of
6 this presentation so you can know what we intended
7 for this to mean, and so you can tell us not only do
8 you agree or disagree, but you didn't actually say
9 what you were supposed to say there.

10 So are there any questions about the agenda or
11 anything?

12 So the first thing you would want to know is
13 with our time line. For the month of February, last
14 weekend -- or two weeks ago and this week we're
15 actually holding workshops somewhere online on
16 webinars and some were face-to-face. And we'll
17 conclude those, that first set, this Thursday. And
18 throughout the months of March and April, as you
19 know, legislative session is going on and anything
20 can happen. So there could be something, a bill
21 that's ultimately passed and signed into law that
22 could affect the content of these rules. So we don't
23 have any plans to put these before the state board
24 until they meet again in June because we're going to
25 monitor session to see if anything else happens.

1 Plus, that will give us a chance in those months to
2 also run data on any suggestions we get for what the
3 standards ought to be because we're talking about
4 standards, performance standards, in these rules and
5 to see how those might impact if they had -- we
6 usually do those based on data we already have. We
7 look at 11-12 data, and if these rules had been in
8 place in 11-12, what would that have meant for folks.
9 But based on -- we've already gotten some suggestions
10 for how to do the standards differently or add to the
11 standards that we proposed. So we hope everybody
12 keeps thinking about that and gives us more
13 suggestions so we can run some impact data.

14 So we've got through March, April, and May we
15 would put out another version of these things for
16 consideration. And then our target would be to take
17 them to state board in June because the intent would
18 be that the rules would then go into effect for the
19 following school year. So nothing would change
20 for -- since these relate to personnel evaluations,
21 nothing would change based on these rules for the
22 current school year, for 12-13. It would only be
23 13-14 and forward that they would go into effect.

24 Any questions on the time line?

25 All right. We'll take these rules in order. So

1 the first rule we're going to talk about is the one
2 that's labeled 6A-5.030, 030. These rules work in
3 conjunction with each other. This is kind of the
4 overarching rule that implements how -- we have to
5 explain in this rule how a district would submit an
6 evaluation to the Department of Education for review
7 and approval based on what they're required to have
8 in their evaluation system under the law.

9 So the first thing you're going to want to know
10 is what does the law say is supposed to be in the
11 rule about evaluation systems the district submits
12 and how we approve, and then again our explanation of
13 what we put in the draft for your consideration.

14 So the first thing you'd want to know is the
15 rule authority. So what does the law say we have to
16 write the rules on? And the process for rule-making
17 for any state agency, whether it's Department of Ed
18 or Department of Agriculture, is you have to have
19 specific authority in the law to write the rules.
20 And the law very often will get down into
21 nitty-gritty on exactly what they want you to write
22 rules on.

23 So in this case what we have to do in this rule
24 is we have to establish a uniform procedure for how a
25 school district would submit and we would review and

1 approve their personal evaluation system, two kinds:
2 The instructional personnel evaluation system and the
3 school administrator evaluation system. There could
4 be other people that the district has evaluation
5 systems for, but the only two kinds the department
6 would approve based on the law would be instructional
7 personnel and school administrators.

8 Secondly we would need to talk about how
9 districts would annually report results of the
10 evaluation system and then what would be contained in
11 the Commissioner's report that has to be done by law
12 every December 1st that says what happened last year.

13 Finally, a new thing for us, because we've had
14 the requirement to review and approve instructional
15 personnel evaluations for over a decade, a new thing
16 that the law added was a monitoring process; how
17 would the department monitor district implementation
18 of the evaluation system. So we have to explain that
19 in the rule.

20 So before we go to how the rules -- what the
21 next section is, if you look at your rule copy for a
22 second I'll kind of orient you to where these things
23 are in the rule so you see how it's set up. And
24 because we use the copies of the rules, they're
25 actually the same copies that are published in the

1 Florida Administrative Weekly, there are no page
2 numbers. We're trying to be consistent with our
3 copies and not have different drafts and versions out
4 there. So it's hard for me to tell you which page to
5 turn to, but the rule starts on page 2. It's easy to
6 find page 2. And the first section you can see in
7 this rule, we have set up a section for definitions.
8 So these terms that are in the definition section can
9 be used all kinds of different ways outside the
10 context of this rule, but for purposes of this rule
11 this is how we define them to try and help people
12 understand what the terms mean for this particular
13 use. And then the definitions go through page 3 and
14 on to page 4 and then there's a second section. That
15 is the Submission Process; how would a district
16 actually send us their evaluation system for review.
17 And so that process has to be outlined in the rule.
18 And then when you get to the next page about halfway
19 down there's a third section which is the Content of
20 the Approved Evaluation Systems. Now, that section
21 is the longest because it's -- and I'll explain in
22 the next set of slides what's in that section.
23 Because that's going to be the evidence the district
24 would have to present to us that they have included
25 everything the law says is supposed to be in their

1 evaluation system. So that's where we explain to the
2 school district what we're going to be looking at for
3 review and approval. That's the longest section. It
4 goes on for a few pages. And then eventually when
5 all that is made clear, then you get to what would be
6 page 11. So several pages over there's a paren 4
7 which is called the Initial Review Process. It's
8 about a third of the way down on page 11, and Section
9 5, the approval process.

10 So in the initial review process we tell
11 districts what we're going to do and how long we're
12 going to take to do it so they know that time line
13 and we're held to that time line. And then in
14 Section 5 we actually talk about what the approval
15 status needs to be, whether you could be fully
16 approved and how do you get to have a fully-approved
17 system, how do you get a conditionally approved
18 system, and then under what conditions would we deny
19 a district submission to have an evaluation approved.

20 After those are defined, then in Section 6 on
21 the next page, on page 12 near the bottom, the law
22 says the district can modify its evaluation system at
23 any time, but only a substantial revision would have
24 to be submitted to the department for approval. So
25 we try to define what does it mean to have a

1 substantial revision. So we defined substantial
2 which means something less than substantial. So an
3 example of something less than substantial would mean
4 they just had a new election, there's new board
5 members, they changed the front cover. That's not
6 substantive. If they decide they're going to change
7 from the Marzano framework to the Danielson framework
8 that's a substantive change. If they're going to do
9 that, that would have to come back to us for review
10 and approval. Those are just examples.

11 Then the monitoring section. I told you we had
12 to do that new thing with monitoring the evaluation
13 systems the districts implements and that would be in
14 Section 7 that we tried to lay that process out so
15 everybody would know what to expect. And we based it
16 on the same process we have been using for over a
17 decade to provide feedback to districts on their
18 personnel -- excuse me, their professional
19 development system. And we put together teams of
20 district and university folks that get trained in all
21 the protocol standards for professional development
22 and we actually -- districts would report information
23 on what happened in the evaluation system every year,
24 but we would only go on-site and monitor once every
25 five years. And we would put districts on a schedule

1 like we do for the professional development system,
2 go on-site once every five years and then do annual
3 reports on what we found in each district evaluation
4 system for that year and then over the whole cycle.
5 So since that system seems to be familiar to
6 districts now and we seem to get -- they seem to get
7 good feedback that way, we modeled this monitoring
8 system on that one.

9 Do you have a question?

10 And then finally, reporting is on page 15 next
11 to the last page, what's reported by districts. And
12 then the last thing is the law refers to when
13 districts are required to do training programs, that
14 training has to be, they have to consider in that
15 training guidelines for training evaluators that the
16 department has developed. So even though a district
17 doesn't have to use these exact guidelines, they're
18 just guidelines, we thought we better put them in the
19 rule so people knew what they were. So that's what
20 the last section is.

21 So what I would like to do is on the next few
22 slides go back to that Section 3 of the rule, which
23 is the meat, what's the content that we're going to
24 be looking for that the law says, and these are the
25 things that we would expect to see in a district's

1 evaluation system, so this is really what's in that
2 section 3 of the rule.

3 The first thing is: Is the evaluation system
4 designed the way the law requires, which is designed
5 for effective instruction and student learning
6 growth, and for school principals, school leadership
7 and student learning growth? Is the district using
8 the results to develop district and school level
9 improvement plans and then is the district using the
10 results to identify professional development? How is
11 it connected to the professional development system?
12 And are they using the results for other human
13 capital decisions?

14 The law also requires four different levels of
15 performance and so we would check to see if the
16 district would have those four performance levels in
17 their system: Highly effective, effective, needs
18 improvement or developing, and then unsatisfactory.
19 And as you know today, we are right here doing this,
20 which is consulting with people on how to set those
21 performance standards, which will be in the second
22 rule that we discuss.

23 The law sets forth that there are three major
24 sets of criteria in the evaluation system:
25 Performance of students, instructional practice or

1 school leadership if it's for school administrator,
2 and professional and job responsibilities.

3 There are districts who combine these last two
4 together in one section, which is fine as long as the
5 criteria are there.

6 And the law also says that the evaluation system
7 needs to be based on sound educational principles and
8 contemporary research and effective instructional
9 practices. So one of the things that you would see
10 if you looked in the definition section is what is
11 the definition of contemporary research and
12 instructional practice. So those kind of things we
13 tried to define in the rule so people would know what
14 the expectation is.

15 This is the reminder of what the law says the
16 district needs to include in the instructional
17 practice section. Classroom teachers is the Educator
18 Accomplished Practices, and for classroom -- people
19 who are not classroom teachers but are still
20 considered instructional personnel, the accomplished
21 practices and other things related to student support
22 that are part of their job responsibilities.

23 And certainly there's an expectation in the
24 evaluation system and professional development that
25 everybody can increase their expertise from year to

1 year in something.

2 For instructional leadership for the school
3 administrator evaluation system, the law also says it
4 should be based on the standards, in this case, the
5 leadership standards instead of the Educator
6 Accomplished Practices.

7 And then the law also provides specific things
8 in addition to those that need to be included in the
9 administrator's evaluation system that mostly support
10 faculty development and supporting effective
11 instruction.

12 The other big set of criteria, performance of
13 students criteria, so we would look for evidence of
14 these things in the district's evaluation system.
15 The first part here is the actual, it's a quote from
16 the law, that performance of students, this criteria,
17 at least 50 percent is based on data and indicators
18 of student learning growth assessed annually and
19 measured by statewide assessments or, for subjects
20 and grade levels not measured by statewide
21 assessments, by district assessments that they have
22 chosen. And this little reference right here is
23 actually 1008. It's the chapter in the school code
24 that relates to student assessment programs, most of
25 which is a statewide assessment program, but paren 8

1 is specific to local assessments.

2 Then these are sort of summaries of the other
3 things we would be looking for in the district
4 system. The district can choose to reduce the
5 percentage based on performance of students down to
6 40 percent for someone who doesn't have three years
7 of their student's data available. The district also
8 splits for nonclassroom instructional personnel the
9 percentage between statewide assessment results and
10 other measures of student outcomes, not student
11 support behavior things that you would expect in the
12 instructional practice, but that are based on their
13 job responsibilities. And then -- and that's a
14 choice, districts can choose to do that. They can
15 also choose to combine state and local assessments.
16 For example, if someone is teaching seventh grade
17 social studies and the district wants to include a
18 little bit of the reading component in that person's
19 evaluation and all social studies teachers'
20 evaluations, they can choose to do that, but the
21 local assessment has to weigh more. And they can
22 choose a performance measure over learning growth.
23 Learning growth is what you're shooting for most of
24 the time, but there are some courses that have in the
25 course assessment such as industry certification

1 where it's just a pass-fail test. So it's very
2 difficult to measure growth on a pass-fail test, so
3 a district can say we're going to use achievement
4 measures for those and they can set targets for
5 different levels of performance you achieve. But if
6 they choose to do that, that's something that we
7 would expect an explanation of in their evaluation
8 system.

9 UNIDENTIFIED SPEAKER: But if there's a learning
10 growth doesn't it have to grow from one assessment to
11 another?

12 MS. HEBDA: Yes, if you're doing growth, that's
13 correct. But they can choose for a certain course
14 that doesn't have an assessment that has maybe a
15 scale or any kind of performance levels if it's just
16 a pass-fail test - hard to measure growth on the
17 pass-fail test - they could choose an achievement
18 measure for that particular course to use for
19 evaluation purposes instead of trying to measure
20 growth on a pass-fail test.

21 UNIDENTIFIED SPEAKER: But if they choose
22 achievement, then they would have to choose - this is
23 my gripe - they would have to choose one achievement
24 and then the growth of the same type of an
25 achievement. Assessments, assessments and growth to

1 the same type of --

2 MS. HEBDA: Only if they're going to actually
3 measure growth. If they are saying I can't measure
4 growth on this industry cert test, certification
5 exam, because all I get is a result for each kid
6 whether they passed or failed, then what I'm going to
7 do -- and some districts already did this with AP
8 because they're used to this with advanced placement
9 tests anyway. Advanced placement tests have five
10 levels, 1, 2, 3, 4, 5. Three and above is considered
11 passing. Some institutions want a 4 or above or
12 whatever, but 3 above is considered passing. So what
13 some districts did last year was they took those five
14 levels and said we can't measure growth on AP yet,
15 but what we do want instructionally for these
16 students is we would like to increase the percentage
17 of students that are passing the AP test. So they
18 would set performance targets for the number or the
19 percent of students in different AP courses that
20 passed the AP test. And they may also set targets
21 for students who maybe come in who don't
22 traditionally get placed in AP based on their
23 previous FCAT score where they would adjust those
24 targets or they may give additional credit as more
25 students gets a 4 or 5 on the AP. There are lots of

1 different things they did based on what the goals
2 were for the AP course.

3 UNIDENTIFIED SPEAKER: So it does not have to
4 show growth.

5 MS. HEBDA: It does not have to show growth.
6 Growth is the premium. That's what you try to shoot
7 for because you want to be able to reward people
8 taking folks from wherever they start to wherever
9 they can finish, but there are times when you just
10 can't cram a growth measure into a pass-fail test
11 because your growth measure always has to work with
12 your assessment that you have. And, again, the --
13 what we also want, of course, because we're trying to
14 increase student achievement, is it needs to make
15 sense instructionally for what's happening in the
16 class. That's really important stuff.

17 MS. CORN: Excuse me.

18 MS. HEBDA: Yes.

19 MS. CORN: If you're an itinerant teacher, you
20 go around to different schools, do they use every
21 school that you go to?

22 MS. HEBDA: They can. What the law talks about
23 a lot is being based on the students that are
24 assigned to you. So if you're assigned one class of
25 students, for example, I don't know how it works for

1 you.

2 MS. CORN: No. Individual students.

3 MS. HEBDA: Individuals. So you have three or
4 four students in each school.

5 MS. CORN: Uh-huh.

6 MS. HEBDA: Those are the students that are
7 assigned to you, not the other students.

8 MS. CORN: What if you're half and half, you
9 teach part time and you're itinerant?

10 MS. HEBDA: They should be able to combine those
11 things all together. Because there are some teachers
12 that have growth measures and teach all kinds of
13 classes. They may have a couple of classes a day
14 that are related to a statewide assessment like sixth
15 grade reading which is related to sixth grade FCAT
16 reading, but they may also have four other classes
17 during the day that are something else. So all of
18 that should be combined into what goes into the
19 evaluation. And that's what we would, in this
20 particular rule for us to approve the evaluation
21 system, that's what we're looking for is has the
22 district considered that and have they explained that
23 to us, how that works, in their evaluation system.

24 Here's the last piece of the performance of
25 students. Between now and July 1, 2015, which is

1 when this little piece of the law, this little
2 paragraph expires, it says, then districts have some
3 choices for teachers, classroom teachers, this is not
4 nonclassroom, this is just classroom teachers of
5 courses where the district has not implemented or
6 hasn't chosen or is uncomfortable with, or whatever
7 the words are, the assessment they use for the
8 students in that course right now. Because students
9 in seventh grade social studies get grades right now.
10 They take tests. They do all kinds of things and
11 projects and whatever else, but the district may not
12 be selecting those for use in evaluation yet. So if
13 that's the case, between now and 2015 these are the
14 things the district can choose to do. First thing is
15 for those classroom teachers assigned students, the
16 student learning growth must be measured by the
17 growth in learning of the classroom teacher's
18 students on statewide assessments. So what does that
19 mean? What some districts have done last year, they
20 started with - I mentioned reading for the seventh
21 grade social studies teacher. Some districts chose
22 to use both parts of the FCAT results for all the
23 teachers that they didn't have assessments for. They
24 didn't choose between reading and math. Some used
25 reading for social studies and other things and they

1 used math for science and made all kinds of choices.
2 Some districts wanted to use the value-added result
3 to measure growth for this purpose. Others wanted to
4 use learning gain. It doesn't say which one you have
5 to use. It just says growth in learning. And this
6 is a temporary paragraph until July 1, 2015. But
7 then for the other group of teachers whose students
8 don't even take statewide assessments, let's say they
9 teach eleventh and twelfth grade students all day
10 long, there's no statewide assessment for their
11 course associated with those kids even when they're
12 outside of the teacher's class, then it's supposed to
13 be learning targets that are established based on the
14 content of the course and that support the school
15 improvement plan. So as we mentioned, kids are
16 assessed, there are learning goals for students in
17 all of those courses. And so for those teachers the
18 district can choose to let it be an individualized
19 learning target between the teacher and the principal
20 similar to your individual professional development
21 plan, that kind of thing, until such time as they
22 establish the assessments they want to use for this
23 course.

24 The third choice is this last sentence here. A
25 school district superintendent may assign to

1 instructional personnel in an instructional team the
2 student learning growth of the instructional team's
3 students on statewide assessments. So in some
4 districts an entire school is considered an
5 instructional team or a whole grade level is
6 considered an instructional team. And so all of the
7 results from that entire school or that entire grade
8 level would be assigned for purposes of evaluation.
9 So that's a choice the district can make.

10 So that's -- those are the kinds of things that
11 in this particular rule for us to evaluate a district
12 system, has the district chosen to do these things
13 when they don't have assessments and which ones have
14 they chosen and under what circumstances, and all
15 that would have to be explained to us so that we
16 would know they covered all those for us to approve
17 their system.

18 And here is the list of the rest of the stuff in
19 the law that we have to check for. So in addition to
20 instructional practice or leadership practice,
21 performance of students, there's a set of other stuff
22 that we need to make sure the district has done. We
23 need to make sure they have observation instruments
24 with indicators. Remember those indicators that
25 apply to the Educator Accomplished Practices or the

1 leadership standards, one or the other. We have to
2 make sure that they are including two evaluations per
3 year for everybody who is a newly-hired person,
4 instructional person in the district. And that's not
5 just through observation to get two evaluations.
6 They should be having a chat about student data some
7 way midpoint during the year when they have that
8 other observation.

9 Evaluator training. We mentioned that earlier
10 with the guidelines for the evaluator training, make
11 sure they have a system for training all their
12 evaluators. What is their process for letting people
13 who are in the system, whether they're evaluators or
14 they're teachers in the system or principals in the
15 system, what's their process for informing everybody
16 about what's in the district evaluation system.

17 Multiple data sources. That's something else.
18 There should be multiple data sources used in the
19 evaluation system. Again, not every district is
20 ready for multiple data sources in the very
21 beginning, but we expect that by the time the
22 deadline rolls around for all the exceptions to
23 expire that they would have multiple data sources for
24 everybody's evaluation.

25 Have they linked their professional development

1 systems, evaluation systems, and vice versa? And are
2 they using the data from the evaluation system for
3 school improvement?

4 Parental input. For both teachers and
5 administrators the law says that parents have to have
6 input into the evaluation system as appropriate, and
7 that's all it says. It doesn't say how. It doesn't
8 say how often. It doesn't say anything else. So we
9 don't have any requirements for how districts do
10 that, just that they do. So that's in there.

11 Teaching fields needing special procedures. You
12 can also consider this as special circumstances such
13 as someone being on maternity leave or extended
14 substitute, or all kinds of other considerations that
15 might be going on. Somebody who teaches hospital
16 homebound. There are lots of kinds of different
17 teaching deals that need special provisions that
18 don't cover the vast majority of teachers and
19 principals in the system. So has the district
20 addressed these things? We don't tell them how they
21 address them, but we have to make sure they have
22 addressed them.

23 Annual review by the district. The law requires
24 they look at it every year to see how things went and
25 what improvements they can make. We want to make

1 sure they have told us how they're going to do that.
2 Then they have some options, too, which are they can
3 include peer evaluations or peer review of documents.
4 They can include an assistance process for people.
5 They can include input from additional personnel into
6 the evaluation system that may or may not be peers.
7 They can include a process to amend evaluations if
8 they want to and they can include additional
9 professional responsibilities in that third set of
10 criteria beyond what's required by the state board.
11 The state board, as far as professional
12 responsibilities, considers the code of ethics and
13 the principles of professional conduct, what we have
14 covered in professional responsibilities. But
15 districts have employment requirements and all kinds
16 of other things they may expect of its employees and
17 they're free to put those in the evaluation system if
18 they want.

19 So that's everything that's supposed to be in
20 that one section 3 of the rule. That's why section 3
21 of the rule is really long, because what rules are
22 supposed to do is they're - from the agency to
23 whoever they're affecting - they're supposed to help
24 that affected group, which in this case is school
25 districts that have to have an approved system under

1 the law, what are they supposed to do to get
2 approved. So we tried to provide all the things we
3 would need as evidence to make sure all these things
4 are done in accordance with the law, so try not to
5 hide anything from them.

6 That was that rule. Any questions on that rule?

7 Yes, sir.

8 MR. CONNER: Is it substantially the same as the
9 criteria that they used for submission?

10 MS. HEBDA: Yes. It is substantially, I would
11 say 90 percent at least is the same as what they
12 already did to get their 11-12 approved.

13 MR. CONNER: What would be approved --

14 MS. HEBDA: I'm sorry. Can you ask the question
15 again?

16 MR. CONNER: The changes that are in the area of
17 approval of procedures for determining student
18 growth, that's where most of the changes are?

19 MS. HEBDA: Right. There's a separate rule that
20 we're going to talk about.

21 MR. CONNER: I mean in the submission.

22 MS. HEBDA: But in here they're pretty much the
23 same. They're pretty much the same. The monitoring
24 process, of course, is completely new because we
25 never monitored before. But I think if you sat down

1 with the original checklist that we did for district
2 for back in 2010, it started with Race to the Top,
3 you're going to see it matches up with this pretty
4 well. But one of the things when this rule was first
5 put forward we put that Section 3 in a checklist just
6 like districts were used to already and the judge
7 didn't like that. He didn't like -- didn't want it
8 in a form. So we took it out of the form and made a
9 really long Section 3. We put it all in there, but
10 it's the same kind of thing. We tried to chunk it by
11 here's the evidence for performance of students,
12 here's the evidence for instructional practice,
13 here's the evidence for school leadership, that kind
14 of thing, and then the evidence for the rest of the
15 laundry list.

16 So the next rule, so switch papers, this is the
17 companion rule, because even though we have a rule
18 that says here's how we look at districts' evaluation
19 systems to see if they have all the criteria in the
20 law, we also have to set forth a rule that does some
21 other specific things that support that process. So
22 this is the rule authority for this rule, 0411, and
23 these are the things we have to put in the rule. We
24 have to put every formula that we use, the actual
25 math that's used for measuring students' learning

1 growth that's approved by the Commissioner. And
2 we'll talk about which ones are approved by the
3 Commissioner and which ones are just district
4 options. But if it's one that rises to the level of
5 having to be approved by the Commissioner, then
6 that's the one that has to go into the rule. At this
7 point in time we only have one of those and that's
8 the one that was used with FCAT data for the 11-12
9 school year.

10 We also have to, in addition to the measure of
11 student learning growth, we have to include any
12 associated implementation procedures. What that
13 means is we define in the rule and we describe in the
14 rule what kinds of data will the district get once we
15 run this calculation using the math in the rule.
16 What can they expect to get from us, when can they
17 expect to get it, all of those different things.
18 Again, the agency always has to provide to the school
19 district what it is we're going to do. So if we
20 don't do something like that we're as accountable as
21 they are.

22 And then we also have to include in the rule the
23 process by which teachers may review their class
24 rosters. I say may because the department has --
25 provides a tool the district can use to let teachers

1 review their class rosters or they can do their own
2 process. Either way, districts have to let folks
3 review their class rosters. So we have to provide
4 and explain in the rule how we're going to do
5 something if districts can't do that on their own.

6 The other stuff we have to have in the rule are
7 these things. This is what we're going to need a lot
8 of feedback from you on. These are the standards
9 that I was talking about earlier, the performance
10 standards. So what the law requires us to do is
11 establish specific discreet standards for each
12 performance level, highly effective, effective and
13 all that. That ensures consistency in meaning across
14 school districts. That's the purpose of the
15 standards to start with. So when you're thinking
16 about using data in performance of students, then
17 establishing standards so that if it's a statewide
18 assessment or if a district is creating its own
19 standards, how were those going to be consistent
20 across school districts. That's a tall order, so
21 this is our first shot at it.

22 Then we have to establish student learning
23 growth standards that if not met will result in an
24 employee receiving unsatisfactory evaluation overall.
25 So not just the ones that would apply to only the

1 performance of students' criteria, but ones that
2 would apply to become the evaluation at some point,
3 and the same thing for setting a floor for effective
4 and highly effective that someone would need to
5 achieve that if once combined with the other parts of
6 the evaluation system the person would be considered
7 effective or highly effective. So what you will see
8 in the rule -- remember there's a category needs
9 improvement, developing. What we have done in the
10 first attempt is to define these three things, and
11 then the remainder, of course, will be needs
12 improvement or developing, depending on how many
13 years the person has been teaching. But this is
14 everybody's consideration to start with. So if we
15 end up needing to define needs improvement as well we
16 can do that as we go along. But this is what you
17 will see in a few minutes.

18 As I mentioned to you, there are two categories
19 of growth formulas. There are the kind that go in
20 statewide assessments that have to be approved by the
21 Commissioner that everybody has to use the same way
22 because they're statewide assessment data in their
23 districts, and then there are local, growth formulas
24 for local assessments that they choose to measure
25 learning growth that districts adopt, they're their

1 choices, but what we had to try to set forth in the
2 rule is what does it mean for someone to adopt an
3 equally appropriate growth formula that would have
4 the same kinds of results as you might see from a
5 statewide learning growth formula. And so we have to
6 provide examples for school districts. The examples
7 won't be in the rule, but we have tried to define
8 equally appropriate in the rules so when we do this
9 kind of technical assistance and provide those
10 examples, then districts will have some options for
11 how they can use them in their own systems and make
12 those good choices.

13 So as I mentioned, the only growth formula we
14 have so far is the one that was adopted over a year
15 ago, June 2011, to use with FCAT data, and that's the
16 one you're used to hearing called the value-added
17 model. And it's designed -- and we'll walk through
18 that a little bit because it has to be in the rule
19 itself. It's designed for measuring learning growth
20 on courses associated with FCAT. So in this case
21 it's learning growth in grades 4 through 8 for
22 mathematics and 4 through 10 for reading. So since
23 we have to have the math and the model in the rule
24 I'm going to walk you through the content like I did
25 with the other rule of what you would expect to see

1 in this draft.

2 Just a little background information to remind
3 people. You probably already know this, but if you
4 don't, here it is again, or here it is for the first
5 time. When we developed this growth model for use
6 with FCAT data, before the Commissioner approved
7 anything we established a committee under Race to the
8 Top called the Student Growth Implementation
9 Committee. And in fact they're meeting again
10 tomorrow to do the next set of things for Algebra 1
11 and other statewide assessments. But these are the
12 general categories. The committee is a majority of,
13 a majority is based on teachers or includes teachers,
14 but these other folks are also represented in the
15 committee. You can go to their website and see who
16 their names are and what district they work in and
17 all those kinds of things. And they work through a
18 process to develop the model, make a recommendation
19 to the Commissioner, and when he made his approval he
20 didn't change anything about what they recommended.
21 He just accepted it after his review the way they
22 recommended it.

23 So what did they recommend and what's in the
24 rule? The committee looked at eight different kinds
25 of growth models or ways to measure learning growth

1 that could be used in personal evaluations and they
2 settled on what's called a covariate adjustment
3 model, and so the rule sets forth what is a covariate
4 adjustment model, what does that mean. It
5 essentially means, it starts with, as you mentioned
6 earlier, a baseline set of data and then the current
7 year assessment, how did the student do on the next
8 assessment measured by grade and by subject. And the
9 covariates are those adjustments we make about
10 students based on things we know about the student
11 that I'll enumerate for you on the next slide that
12 help you establish how much growth you should expect
13 to see from that student based on things we know
14 about him or her.

15 All those covariates or variables or
16 characteristics fall in three buckets: Student
17 classroom and school characteristics. So here is --
18 this picture is not in the rule, but it's your
19 explanation today. This is actually looking at how
20 the model works based on -- starting with data from
21 one student. So we would do this very individual
22 look at how we expect students to perform. In this
23 case, again, this model is only based on FCAT. There
24 are models that could be done for all kinds of
25 assessments. This one happens to be FCAT. We would

1 take for each student who is assigned to a teacher,
2 this is one student, we would start with their prior
3 year score. And if you notice, this is kind of
4 replicating the developmental scale. It's not levels
5 1, 2, 3, 4 and 5 and the FCAT levels 1 through 5
6 don't matter a hill of beans in this. Look at their
7 prior score. Here's their current year score. And
8 if we weren't doing a covariate adjustment model, we
9 were just looking at something simple, like simple
10 growth, we would stop right there and we would
11 subtract and we would say that's the answer. But in
12 this case the covariate adjustment model establishes
13 a predicted score for each student based on what
14 their prior score was, however they started, however
15 they came to the teacher, and then also making
16 adjustments based on how students typically perform
17 across the state who share those same
18 characteristics. And I'll show -- and then once you
19 know the predicted, once you look at the predicted
20 score, look at their current score, how did they
21 actually do, there could be a difference. Could be a
22 bigger difference, could actually be difference in
23 the other direction. Maybe they did exactly as they
24 were predicted. But each one of these differences,
25 an expected and actual performance, that's what rolls

1 up to the teacher for all of her assigned students
2 and that's what produces the value-added score. So
3 it's the difference, all the differences put together
4 on what we thought the student was going to do based
5 on the things we know about them and where they
6 started and how they actually did.

7 So how do we make adjustments to that green bar?
8 What are the covariates and the variables we use to
9 adjust that green bar and make that prediction? Here
10 are the three buckets. The student characteristics.
11 First, up to two prior years of achievement scores.
12 So you mentioned earlier: what's the prior score and
13 what's the current score? In this case if we have
14 the prior score and the previous year's prior score
15 we'll use both. By far and away the prior year
16 score, when you're looking at an assessment like FCAT
17 that's offered year to year to year, is the biggest
18 predictor of anything. You can roll all of these
19 other things up together and it's not going to come
20 close to how much predicted power you get from that
21 prior year score. But even having said that, another
22 prior year score and these other things can matter to
23 an individual teacher; maybe not statewide, but to an
24 individual teacher they can matter. Remember we did
25 this model for purposes of an individual teacher's

1 evaluation. So having a lot of characteristics and a
2 lot of variables in the models makes that equation
3 longer and more complex and harder to understand, but
4 these are the things that are done that goes in the
5 evaluation for an individual teacher.

6 So two prior years of achievement scores.
7 Number of subject-relevant courses in which the
8 student is enrolled. What does that mean? That
9 means I could be, as a kid I could be enrolled in
10 both a language arts course and a reading course.
11 Well, that's two subject-relevant courses. And when
12 we looked at the data and the student growth
13 committee examined all of the data they noticed that
14 in some cases it made a big difference if the student
15 was in two courses that helped them specifically with
16 reading like that. So more mathematics. So that's
17 included as a covariate.

18 UNIDENTIFIED SPEAKER: So if they're in two
19 mathematic courses, then more than likely you want --
20 they're predicting a higher score.

21 MS. HEBDA: That's right. And actually in the
22 technical report that we put out every year, there's
23 one there now for 11-12 on the website, if you go to
24 the back of it, then actually all of those
25 coefficients and how much we would expect by grade

1 and subject, something to go up or go down, the green
2 bar to go up or go down, are actually listed there so
3 you can see how that relates. It also tells you in
4 that report whether that difference is statistically
5 significant. Because it may look like it's a big
6 number or a really small number, but if you don't
7 look to the right of that and see if it's
8 statistically significant, then it may look like it's
9 having a bigger impact on somebody's value-added
10 score than it really is. So all those things are
11 kind of important to take into context.

12 Student disabilities are included and this is
13 important the way the committee decided to include
14 this or recommended to include this. It's not
15 whether or not the student is disabled, yes or no,
16 but it's actually by disability. Because when you
17 look at the coefficients we were just talking about
18 you can see that different disabilities have
19 different effect sizes on how we would expect a
20 student to do. And all those are based, they're not
21 based on theory or any kind of random weighting
22 procedure, they're based on how kids actually do on
23 the assessment around the state. That's where they
24 come from.

25 English language learner status. That is a yes

1 or no. Either they're in ELL services or they're not
2 the first two years.

3 Attendance is what we call a rolling variable
4 because it's the number of days present throughout
5 the whole school year. As you would expect, the more
6 days a student is present, the better they're
7 expected to do. If they're not in school they're
8 probably not learning.

9 Gifted status. That's yes or no.

10 Mobility and number of transitions. That means
11 in or out of school. One of the things the districts
12 report to us in a standardized way throughout the
13 state is how many times the student leaves a school
14 and comes back. It could be to the same school and
15 it could be to different schools. Doesn't matter.
16 We capture the number of transitions period, and that
17 has an effect.

18 Difference from modal age in grade, indicator of
19 retention, but also an indicator of promotion.
20 Sometimes you have really young kids that are way
21 advanced, and rather than just looking at the data to
22 say retained or not retained, the difference in modal
23 age is a much more exact way to capture all of those
24 different circumstances. Because you could be behind
25 for all kinds of reasons, not just because you're

1 retained.

2 So those are those things. Also classroom
3 characteristics, class size. The homogeneity of
4 students' entering test scores in class. That's an
5 icky way of saying, one of the things the committee
6 discussed, the teachers brought up, was what if I
7 have students who are basically all level 1
8 performers in my class, does that matter, or does it
9 matter if I have all level 4's and 5's? Does it
10 matter if I have a mix? So we ran data and sometimes
11 it can matter. And we run this data by grade and
12 subject, so you won't necessarily expect identically,
13 because it's based on how students do, you won't
14 expect to necessarily see the exact same coefficient
15 for seventh grade math as fifth grade reading.
16 They're going to be different. But that's how we do
17 it, by grade and subject.

18 And then school characteristics. There are
19 schools that really have a lot of growth, schools
20 that really don't have much of any growth at all.
21 And when they're that different - most schools are
22 somewhere in the middle - but when the schools are
23 that different the committee felt it was important to
24 capture that as a covariate, essentially, because
25 that can even influence more of student learning that

1 the teacher can if you're in a really high-growth
2 school or really low-growth school.

3 So as I mentioned, it's an ongoing process. The
4 only thing in this rule to date that's been adopted
5 is the FCAT model. But, as I said, the student
6 growth committee meets tomorrow and they're going to
7 consider models for Algebra 1, biology and some of
8 the other end-of-course tests. So this will be
9 ongoing. If those get approved by the Commissioner,
10 then we go back to the rule and put those in the rule
11 too.

12 MS. MAHLMAN: I'm an Algebra 1 teacher. So the
13 evaluation process that we just now got from last
14 year's students, that may be out the door with
15 something new that's coming through for Algebra 1.

16 MS. HEBDA: But not this year. That would be
17 next year. This year it's going to be the same as it
18 was last year, unless your district has chosen to
19 make a change. The same thing for two years in a
20 row, and then if the committee recommends an Algebra
21 1 model and the Commissioner approves it, then that
22 would go into the rule and would be effective next
23 year. If it's done this spring and we provide
24 districts with data they can choose to use that this
25 year if they wanted to but they wouldn't be required

1 to.

2 MS. MAHLMAN: So it will just be Algebra 1 and
3 biology will be the new ones?

4 MS. HEBDA: Those are the ones we're looking at.
5 Plus what's the third one?

6 MS. MAHLMAN: Is it geometry?

7 MS. HEBDA: We haven't done geometry yet.
8 Algebra 1 and biology was on the agenda for tomorrow.
9 And that meeting is webcast live, but then the
10 recording is also put up on that same website I
11 showed you for SGIC, student growth committee, right
12 here. So it's in your PowerPoint. So if you want to
13 follow from the department's home page or from their
14 home page you can or you can go back later and look
15 at the recording or look at the presentation
16 materials or anything you want. All of our committee
17 proceedings are always webcast live.

18 So remember I mentioned to you that we had to
19 put in the rule those other implementation procedures
20 and what are the other kinds of data and what are the
21 things that we need to understand about how this
22 particular model works so we can even set performance
23 level standards, what would the numbers even mean.

24 The first thing we need to understand is what
25 zero means. In this kind of model and in many models

1 that are like this, covariate adjustment models and
2 others, what you're looking for is some difference,
3 as you saw the difference in the student's individual
4 score, you're looking for a difference in typical
5 expected growth. And zero means that the student
6 performed typically like other students across the
7 state that shared those same characteristics,
8 including prior score and everything else as well.
9 So a score of zero actually means that's good.
10 That's exactly what everybody expected. It also
11 means if it's above zero, then students on average
12 for that teacher, once all those residuals are rolled
13 up to the teacher, on average outperformed what we
14 expected, or if it's a negative score that means on
15 average they did less than what we expected. An
16 individual student may have done one or the other
17 things, but that score for the teacher puts all of
18 those data points together to arrive at the score.

19 Another piece of data we get, and some districts
20 use this instead of the score in their evaluation
21 system, is the -- because we do that for each
22 individual student, we know which student met or
23 exceeded, or did not, their expected score. So for
24 any teacher's class roster what the district has is
25 an indication of whether or not each individual

1 student met or exceeded their expected score. So
2 what some districts did instead of using the
3 value-added score itself the first year, they used
4 the number or the percent of students in each
5 teacher's class that met or exceeded that green bar.
6 Okay. Now, one of the things you lose in that -- you
7 don't lose -- you lose the amount of gain. It's
8 almost like going back to learning gains where it's
9 just a yes or no question. Did you make a gain or
10 not, yes or no. You don't get to take into account
11 like the score does how far you may have moved a
12 student, but it is a piece of information that can be
13 helpful, because one of the questions some people
14 want to know is how many students is that teacher
15 reaching.

16 Another thing we get is a standard error of
17 measurement. And the standard error is important
18 because any time you use any kind of statistical
19 measure, any kind for any reason, you want to use it
20 wisely. So in this case one of the things you want
21 to know is how much variability is in that score,
22 meaning not should it have been something this year,
23 but what if this same teacher was assigned another
24 group of students, similar, but different kids, would
25 the results have been the same or how much -- what

1 kind of range would we expect that score to be in.
2 And we relate it to something, it's the same process,
3 although I think this is a less important process,
4 because I'm not the president, is a presidential
5 popularity poll. The question is, people ask them
6 all the time, is the president doing a good job or
7 not, yes or no. And they'll sample so many voters,
8 so many registered voters in three states and they'll
9 say of the registered voters polled or sampled
10 50 percent said he's doing a good job. And they'll
11 say 50 percent plus or minus three points. You have
12 seen that before. That plus or minus three points is
13 something called a confidence interval and it's
14 constructed by that standard error, the standard
15 error of measurement. Standard errors tend to be
16 smaller when you poll or you sample or you have lots
17 of data points. So one of the reasons why the bill
18 puts a premium on having three years' worth of the
19 teacher's data in her evaluation or his evaluation is
20 because the more data points you get the more
21 confident you can be that this is typical performance
22 for that teacher. It doesn't change what the score
23 is for that year. The score represents what
24 happened. But the confidence intervals can be used
25 to say if I had, in the presidential poll, for

1 example, if I had gone to the same three states and I
2 had sampled another set of registered voters,
3 different people, but still met the same
4 qualifications, just like our students who would
5 share the same characteristics, I would be confident
6 that that would necessarily be somewhere between 47
7 and 53 percent. That's what that means, plus or
8 minus three from what the actual number was. Those
9 registered voters that were polled said, 50 percent
10 of them said he's doing a good job. The confidence
11 interval said if I had sampled another group I'm sure
12 that it would have been within this range. So
13 knowing that's a good way to use statistics and
14 because we're trying to use these numbers not just to
15 rank people, but to say in combination with
16 instructional practice and all these other measures
17 we have in the system, how confident am I that this
18 is how this teacher would normally do in any other
19 circumstances, then the confidence interval becomes
20 important.

21 I knew I would put the clicker down somewhere
22 and I wouldn't know where it was.

23 And that gets us to how are we then setting up
24 our standards. So if we're going to use our
25 statistics wisely and use them the way we should,

1 then one of the things that we want to do is to try
2 to construct a system where we aren't just ranking
3 people every year, so we actually set criteria for
4 the performance level that people can work on
5 exceeding every single year. And so if everybody
6 ends up effective and highly effective because they
7 have jumped over these standards every year, great,
8 because that means kids are learning.

9 So one of the other pieces of information that
10 would be important for you to know is what does the
11 VAM number mean, what does the score mean. The
12 aggregate VAM score actually is a portion of the
13 year's growth. Since we know what typical growth is
14 every year by grade and subject, how much students on
15 average moved that year, then when we look at a VAM
16 score and a VAM score turns out to be .2, for
17 example, .2 would be 20 percent over typical
18 performance, which means that teacher's students on
19 average when you put them all together grew
20 20 percent more than the state average.

21 So what we would like to do, our proposal is -
22 people are talking about whether this is a good idea
23 or not - we would start with 11-12 data and whatever
24 the typical growth was for 11-12, we would use that
25 to set our standards and then people could try to --

1 could then exceed those standards every year. We
2 would provide districts the standard error so they
3 could use the confidence intervals, and I will show
4 you how we do that in just a minute. And then these
5 would be used beginning next school year, not this
6 school year, but next school year, just to classify
7 teacher and principal performance and performance of
8 students. And that's part of the evaluation system
9 and related only to those courses that are related to
10 FCAT data. Okay. That's what these performance
11 standards are for.

12 So what are the standards we are proposing?
13 Here's a visual example of the things we just talked
14 about. In our visual the diamond represents the
15 score, the 50 percent of voters that said what they
16 said. This is the teacher's value-added score for
17 that year, what actually happened. We can construct
18 two levels of confidence using the standard error.
19 One level of confidence uses one standard error, the
20 red section, and that says I'm 68 percent confident
21 that if this teacher had been assigned another group
22 of students with similar characteristics, the score
23 would have either been here or somewhere in this
24 range. If I use two standard errors to construct my
25 confidence interval, then you can think of that as

1 plus or minus one, plus or minus two, then I'm
2 95 percent confident that it will be somewhere in
3 here. Doesn't change what the score was. The score
4 is what it is and that's what happened that year, but
5 what you're trying to do is say how confident am I is
6 that typical performance for that teacher so I can
7 combine it with the other things and make a really
8 good, solid judgment for evaluation purposes.

9 Yes, ma'am.

10 MS. CORN: How do you get a negative confidence
11 score, confidence interval?

12 MS. HEBDA: The confidence interval is, whatever
13 the standard error is, let's say the standard error
14 is .2 for a particular score, so the top of the
15 confidence interval is plus .2 and the bottom of the
16 confidence interval is minus .2. So you take
17 whatever the standard error number is and you add it
18 to the score here and you subtract it from the score
19 there and that's your whole VAM. And then you do
20 that for the two standard errors. So if the standard
21 error is .2, that means I take the score and I add .4
22 to it and I subtract .4 and that's my VAM. That's my
23 VAM where I think 95 percent confidence, as positive
24 as anybody could possibly be using numbers, that if
25 this teacher had been assigned another group of

1 students her value-added score would have been
2 somewhere in whatever this range is. Now, you don't
3 see any numbers on this page because I'm going to
4 show you numbers on the next page. I want you to get
5 the visual. Score, one standard error, 68 percent
6 confidence; two standard errors, 95 percent
7 confidence. We're trying to judge whether or not
8 based on state average how kids typically do in
9 Florida, if this teacher's students outperformed or
10 underperformed that average.

11 So here's what we propose. We've already gotten
12 suggestions on how we can modify and add to this, so
13 we would love to take more, but here's what we're
14 starting with. We're going to call this zero, this
15 black line zero. We know what zero means in typical
16 performances. And I'm going to start with the
17 outside edges first because they're the clearest. So
18 to be considered for performance of students section,
19 highly effective, you can see that I have a
20 value-added score that's above zero, so my score
21 itself is above typical. And I am 95 percent
22 confident, 95 percent confident that in any other
23 circumstances my score would have been completely
24 above zero. So my whole confidence band of my range
25 of scores that could possibly have happened based on

1 all the things I know is still above state average
2 performance. That would be highly effective. And,
3 conversely, if I have a score that's below zero and I
4 constructed a 95-percent confidence interval that
5 under any other circumstances all of my on-average
6 student performance would have been less than state
7 average, then that would be an unsatisfactory result
8 for performance of students. There's still
9 instructional practice, there's still professional
10 and job responsibility, but that would be the piece
11 of the evaluation that pertains to performance of
12 student.

13 So there's highly effective, there's
14 unsatisfactory. Here's our proposal for effective.
15 Two different definitions for effective. In the
16 first one you can see the actual score itself is
17 above zero, so it is better than the state average
18 and that's what actually happened that year. When I
19 tried to construct a 95-percent confidence interval
20 that given any other set of circumstances it would
21 still be above zero, some of that confidence interval
22 falls below. So there's a chance, given a different
23 group of students, that I might not have done as
24 well, but this year I did. So that's effective
25 performance. I may not be totally positive it would

1 always be that way, but for this year it was, so
2 that's effective.

3 Now, at the same time my score might be just
4 below state average when you roll everything together
5 and that's what my value-added score is. But when I
6 construct my confidence interval, even if I only use
7 68 percent confident, it still could have been above
8 zero given a different group of kids. So in this
9 case I can meet the definition for effective for
10 performance of students either way, either starting
11 out with a score that's above zero, period, with or
12 without a confidence interval, or having a score
13 that's just below zero but with a confidence interval
14 could have been above by a change of kids. So then
15 what you see is the remainder would be either needs
16 improvement or developing, depending on which year of
17 teaching the teacher would be in. So that's what we
18 propose to start.

19 Now, one of the suggestions we have already
20 received, which is an interesting one, is starting
21 with this set of definitions, but then taking into
22 account that -- remember one of the things I said is
23 one of the ways to reduce your potential for error is
24 to have a lot of data points. So there's some
25 teachers that even over the course of three years

1 don't have that many students assigned to them. So
2 by virtue of that they may have a larger standard of
3 error or a larger confidence interval. We may be
4 less sure than we would about a teacher who maybe
5 teaches middle school and has 120 different kids
6 every year as opposed to teaching fifth grade or
7 fourth grade where I have only 20 different kids or
8 22 different kids every year. So what could happen
9 is you could have two teachers with an identical
10 score for the same year, but because of their
11 confidence interval, one confidence interval may
12 stretch out because we don't have as many data points
13 and the error potential is pretty big, so that it
14 touches below zero, and for the other teacher I could
15 have a whole big bunch of data points, a small
16 standard of error, and I could be highly confident no
17 matter what I did she was going to be above zero, but
18 the real difference is that level of confidence. So
19 the suggestion was to rectify that was not to be less
20 confident, but to say that if we had a second cut
21 point for highly effective, that a second definition
22 for highly effective, and this is the one where we're
23 95 percent confident this is the way it's going to
24 be, that if I for that year, because the score is
25 what happened, if I for that year had a rockin' and

1 rollin' value-added score even if my confidence
2 interval might be too big, if I have a score that's
3 this high, whatever that second cut point is, I could
4 still be considered highly effective for that year.
5 That would be another way to define highly effective.
6 That's not in our rule right now but I wanted to
7 present you that suggestion to let you know what some
8 people have thought about and been considering so you
9 can consider the same kinds of things in your
10 feedback.

11 MS. MAHLMAN: Then on the opposite side below
12 zero, then are you saying that the needs improvement
13 is that, so you have two above, one effective, and
14 then two below?

15 MS. HEBDA: In that case we could still keep the
16 two effective definitions. We would just add a
17 second definition for highly effective as well. In
18 other words, we would keep these same four, the
19 definition for unsatisfactory, the two definitions
20 for effective and then add a second way to be highly
21 effective.

22 MS. MAHLMAN: But if you add a second way to be
23 highly effective shouldn't you have a second way
24 below the line?

25 MS. HEBDA: Well, you could. You wouldn't have

1 to, but you could. That could be a suggestion. It
2 may be that the only way you want to determine
3 somebody is unsatisfactory is if you're 95 percent
4 sure this is the case. It may be, though, that like
5 you're saying that even if I'm not 95 percent sure
6 somebody's value-added score is so low --

7 MS. MAHLMAN: No. I was thinking between
8 unsatisfactory and the --

9 MS. HEBDA: Oh, for needs improvement.

10 MS. MAHLMAN: Yes.

11 MS. HEBDA: For needs improvement, right. So
12 what we could do is figure out or define what the
13 remainder is. Not just say the remainder is going to
14 be needs improvement, but actually put the definition
15 around needs improvement. We could do that too.
16 That's another suggestion.

17 Tom.

18 MR. CONNER: So each teacher's data points would
19 determine their individual standard error of measure?

20 MS. HEBDA: They do. That contributes to it.
21 It contributes to it. A larger number of data points
22 allows you to be more confident than what you saw
23 could happen --

24 MR. CONNER: An elementary teacher, like you
25 said, who has 24 kids or whatever, you're saying that

1 their standard error measure would be determined on
2 these 24 data points, not fifth-graders who have the
3 same --

4 MS. HEBDA: No, no, no. The individual student
5 prediction is still measured on how those kids do
6 around the state and the value-added score is still
7 the difference between how each one of her kids was
8 supposed to do or we thought they were going to do
9 and how they actually did. That's still the
10 teacher's score. In addition to the score we get the
11 measurement of what the variability in the score
12 could be if I had been assigned this group of kids
13 and that's what the standard error is. It's a
14 separate piece of information from the score.

15 MR. CONNER: But every teacher will have their
16 own individual standard error of measurement.

17 MS. HEBDA: They do right now. Each individual
18 score has its own standard error of measurement.
19 That's right.

20 MR. CONNER: And that standard of error of
21 measure is only determined by the data points of that
22 individual.

23 MS. HEBDA: Not only, but it's largely
24 influenced by that. There's also some measurement
25 error in any test you give. There's not as much

1 measurement error in the FCAT as there would be for
2 textbook, but there's always a little bit of
3 measurement error taken into account anytime you do
4 assessment statistics, but primarily the standard
5 error is influenced by the number of data points you
6 have.

7 MR. CONNER: Just for clarification then, there
8 could be a substantial variance between the standard
9 error in measure of a fifth grader in Duval and one
10 in Dade?

11 MS. HEBDA: For any teacher, for any teacher.
12 It's not the fifth grader himself. It's the
13 teacher's score, the actual value-added score.
14 Because it's a result of a statistical calculation
15 there's variability in that score just like in the
16 presidential poll. If I sampled a different group of
17 people I might have gotten a different result, but
18 based on my statistics and what I know about all
19 those people I'm confident that within three points
20 either side it's going to be in that range. That's
21 the same thing we do with the value-added score.

22 MR. CONNER: For some reason when I read it I
23 thought it was all fifth-graders who had similar
24 characteristics and VAM scores.

25 MS. HEBDA: It is how those fifth-graders do on

1 the assessment, what their developmental scale score
2 is. That's what the student's prediction is based
3 on, the developmental scale of the FCAT and how --
4 let's say they're predicted to do a 220 and the
5 student actually did 230, so there's a ten percent
6 difference. And I take those ten points and I -- ten
7 points, and I take all the other points from all the
8 other students' differences in how we felt they were
9 going to do and how they actually did, and all the
10 other stuff in that green bar, the prior score, the
11 student characteristics, all those are the other
12 contributions, including the school, are the other
13 contributions to student learning, that difference
14 between what we predict based on all those things we
15 know and what actually happened, that's the part
16 added by the teacher and that's why those residuals
17 are rolled into the teacher's value-added score
18 because the other stuff is already accounted for in
19 the model. And that's why the committee went with
20 the covariate-adjusted model, because this was before
21 an individual teacher's evaluation or an individual
22 principal's evaluation, and then when you measure
23 learning gain you just do the subtraction, just
24 subtracting learning gain, anything could have
25 influenced that learning that day.

1 UNIDENTIFIED SPEAKER: Is there a percent that's
2 needed of the number of students that you used or
3 should be used for sampling?

4 MS. HEBDA: We don't have any strict guidelines
5 on that.

6 UNIDENTIFIED SPEAKER: I suggest you do.

7 MS. HEBDA: Okay. That's a good thing. That's
8 a good suggestion. When we get to the comment part
9 make that suggestion, we'll work on that. We
10 certainly would say two is not enough. But what is
11 the right number? I'm not positive, but we can
12 research that.

13 UNIDENTIFIED SPEAKER: There's still some
14 confusion about the standard error around the
15 teacher. Basically comes down to if you have -- if
16 Cathy has a class of 24 fifth-graders and I have a
17 class of 24 fifth-graders, our standard error
18 measures, our confidence intervals are going to be
19 different because we have different students, but
20 it's all informed by the statewide information that
21 we've gathered from all fifth-graders.

22 MR. CONNER: It would be compared to the --

23 THE COURT REPORTER: I'm sorry, sir, I couldn't
24 hear you.

25 MS. HEBDA: He said you'd be compared to the

1 state mean.

2 MR. CONNER: In relationship to the state mean.
3 That was just -- I had thought it would be like
4 teachers who would determine the standard error
5 measure for teachers as opposed to an individual
6 teacher who might have -- she had mentioned a small
7 number.

8 MS. HEBDA: So those are what we propose to
9 start with and we want your feedback on whether you
10 like those, don't like those, which ones you do and
11 don't like and the other suggestions you have already
12 started to make about what we should include and
13 anything else.

14 Most districts around the state, because right
15 now in 11-12 and in 12-13, districts set their own
16 cut point for how these numbers are used, how the
17 value-added results are used, how instructional
18 practice is used, all those things, and they have
19 their own scoring system. The majority of the
20 districts around the state use the score and use the
21 confidence intervals now. As I mentioned to you,
22 there's probably eight or so districts that don't use
23 the score, they use the percent or number of students
24 in the class that met or exceeded their predicted
25 score. And within the districts who used the score

1 with the confidence intervals, some have this -- for
2 example, Lee County was telling me they use almost
3 the exact same thing in their evaluation system right
4 now. Other districts did what I mentioned a minute
5 ago, which is why they had the suggestion, that they
6 start with this and then set a separate cut point for
7 highly effective and effective and a cut point for
8 unsatisfactory. So one of the benefits in not doing
9 this as a state yet, although we know we want to do
10 the consistency part across the state, is letting
11 people work with what makes sense to them to start
12 with actually helps us get feedback on how we have to
13 set them in the state. Because if we weren't
14 explaining this and the districts weren't using this
15 already for a year and a half as it worked for them
16 in their own district, then we wouldn't probably get
17 any suggestions at all and nobody would understand
18 it.

19 So now here's the other part. Remember there
20 are two things we have to do with these performance
21 standards. I'm going to go back to the -- based on
22 what the law requires. We had to establish them for
23 performance of students so we have consistency and
24 meaning, but then we also have to establish them such
25 that they are used to determine the final evaluation

1 for unsatisfactory and to set the floor for effective
2 and highly effective. So two uses for these
3 standards in the rule. So one of the things we
4 thought would be very important is not just to
5 establish what everybody thought were the right
6 performance standards, but how would they be used for
7 that second purpose and under what conditions would
8 they be used for that second purpose. So here are
9 the conditions that we thought would be important to
10 apply for using the performance standards, not just
11 performance of students, but to determine the final
12 evaluation.

13 In the law you may remember that the district
14 has the option to reduce the percentage of the
15 evaluation based on performance of students if
16 there's not three years' worth of data for a teacher,
17 so we thought it would be important to not apply
18 these standards as the ultimate evaluation rating if
19 we don't have three years of data for that teacher.
20 So until there's three years of data for a teacher it
21 would only be used for performance of student
22 calculation.

23 The second thing is we saw that paragraph that
24 had all the temporary measures you can use until July
25 2015. If a teacher is being evaluated on one of

1 those things like team data or learning targets or
2 something like that, then you would not apply the
3 standard to affect the ultimate evaluation. It will
4 just be applied to performance of students and that's
5 all. And because nonclassroom instructional
6 personnel could use those other student outcomes,
7 then we're suggesting for your consideration that
8 nonclassroom instructional personnel would not be
9 included for this use of the standards. The district
10 would need to set something comparable for
11 performance of students but they would not be used to
12 determine the final evaluation for nonclassroom
13 instructional personnel.

14 So those are the conditions we included in the
15 rule. They're set forth. One other thing I'll call
16 your attention to in the rule if you want to look at
17 it is the comparable standard by school districts and
18 their option for doing that. Because, remember,
19 these standards are related to FCAT data and the FCAT
20 model and we would have to set standards for an
21 Algebra 1 model at some point and all those things
22 once they're approved, but then districts will be
23 setting those comparable standards for their own
24 assessment, their local assessments and how they
25 might do AP or any of those things.

1 So what would be page seven if they were
2 numbered, there's a little paren C that says
3 district-established performance level standards for
4 local assessment. And we provided two methods by
5 which districts could establish their own standards,
6 how they would be comparable. One of the methods
7 uses that percent meeting expectation. So if a
8 district went back and looked at how many teachers in
9 their district, all their highly-effective teachers
10 in their district based on the score and then looked
11 at how many percent on average of the kids in those
12 rooms met or exceeded expectations, they could use
13 that to possibly set performance standards for local
14 assessments and say I expect 75 percent of the
15 students, for example, just off the top my head, on
16 average, in all their highly-effective teachers'
17 classrooms met or exceeded their expected score, then
18 perhaps they'd set 75 percent as the bar for highly
19 effective for number of students that passed the AP
20 exam or number of students that passed the industry
21 cert exam or something like that. So they could use
22 the data they already have on their other teachers'
23 performance to look at as one criteria and to look at
24 setting comparable performance standards.

25 The other would be an equal amount of learning

1 growth. If they are measuring a learning growth on a
2 district assessment with a pre- and post-test, then
3 we know the value-added score, the aggregate score is
4 turned in proportion of a year's growth. And so, for
5 example, in 11-12, mathematics teachers around the
6 state that were rated highly effective by a school
7 district on average had a value-added score of
8 basically, I think it was .19. So on average
9 mathematics teachers that were rated highly effective
10 in the state, their students on average grew 20
11 percent higher than typical. So a district could
12 take that kind of information, a portion of the
13 year's growth on these measures and say, well, on my
14 own learning assessment, my own pre- and post-test,
15 then I'm going to expect 20 percent above the mean
16 growth for the class for someone to be rated highly
17 effective on performance of students for those tests.

18 The third option is if neither one of those
19 work, then propose what you want but give us the
20 instructional rationale for why.

21 Those are our three options for district setting
22 comparable standards on their tests.

23 I don't think there's anything else I need to
24 tell you about this rule. There's only one document
25 that's incorporated in the school which is, I

1 mentioned the technical report that has the
2 coefficients in the back, because we have to put the
3 formula itself in the rule, we put the formula in the
4 rule and we explain the same thing you saw on the
5 slide, but then because full transparency, because we
6 own the value-added model, it's not licensed from a
7 company or anything else like that, the technical
8 report, the methodology section of the technical
9 report actually tells you exactly what you would do
10 if you were some other statistical company or anybody
11 who wanted to replicate this model, you can do it.
12 So that's why that methodology section is
13 incorporated in the rule. So even if it doesn't
14 matter to you because you're not going to go home and
15 try to calculate the value-added model for the entire
16 state of Florida, because you don't have access to
17 that data anyway, it's full transparency. So the
18 whole model is here in the rule.

19 All right. I think that is everything. There
20 are some slides at the end that are only there for
21 people's information on what has happened in the past
22 and how districts worked on their evaluation systems
23 and things like that, but that's just for your
24 information. It's not really pertinent to what we
25 would be getting feedback on today.

1 So are there other questions before we go into
2 comments? Anything else you think you want to ask at
3 this point?

4 Yes, ma'am.

5 MS. CORN: So again when you were doing the
6 performance standards, that last slide, should it be,
7 when you're a half-and-half person, who decides which
8 half --

9 MS. HEBDA: The district does. That's why it's
10 important that we have those in there.

11 MS. CORN: Do we have that in writing in our
12 district?

13 MS. HEBDA: What do you mean?

14 MS. CORN: That there's -- who -- which one
15 is -- are chosen to do that, to do the -- which one I
16 would be?

17 MS. HEBDA: It's actually going to be aggregated
18 together. So one of the things that we did say in
19 the rules is, here's a teacher who teaches two
20 classes a day that relate to state assessments and
21 four classes a day that don't. And if the districts
22 got the assessment they want for those courses and
23 they're using those, maybe two of them are
24 performance measures and two of them are growth
25 measures, but the district has set those comparable

1 standards we just talked about for each one of those
2 courses. And so for your evaluation, if you have two
3 and two and two, then the third-year performance of
4 students' criteria is measured on the state
5 assessment and the third is measured on the
6 performance standard, or achievement standard they
7 have for those other two courses, and a third is
8 measured on the growth standard they have for the
9 third --

10 MS. CORN: Are they using the roster
11 verifications to figure out which students were
12 itinerant?

13 MS. HEBDA: They should. I don't know how
14 roster verification works for itinerant teachers.

15 MS. CORN: Well, everybody is supposed to --

16 MS. HEBDA: Should be. Did your district use
17 the tool last year to verify rosters?

18 MS. CORN: Yeah.

19 MS. HEBDA: Okay. Did you get to go in and do
20 your --

21 MS. CORN: A couple of the four, of the five
22 schools I went to but not all of them.

23 MS. HEBDA: But not all of them. Well, that's
24 something that needs to be worked on.

25 MS. CORN: I didn't hear from them all and I'm

1 not labeled the same way at all.

2 MS. HEBDA: That's something that you need to
3 talk to somebody about for sure. Absolutely.

4 Yes, sir.

5 MR. CONNER: The list of courses that are
6 associated with the base year.

7 MS. HEBDA: Yes.

8 MR. CONNER: So if the student was enrolled in
9 one of those courses and took the FCAT that year,
10 that data is in.

11 MS. HEBDA: And if they had a prior score from
12 the previous year.

13 MR. CONNER: Calculus AB here and everything
14 that's on the list?

15 MS. HEBDA: All of these courses --

16 MR. CONNER: So if they were in the course and
17 took the FCAT that year, then they're in this base
18 data.

19 MS. HEBDA: They are in the data as long as they
20 also had a prior score from the previous year.

21 MR. CONNER: Yeah. I mean prior score.

22 MS. HEBDA: That's right. What these courses
23 here are are those that are associated with FCAT.
24 That's all the courses that are here, are the ones
25 associated with FCAT because it's the only model we

1 have. Once we have an Algebra 1 model, we'll use
2 Algebra 1, the Algebra 1 courses that are affected.

3 MR. CONNER: The student database, when they
4 looked at it, the kids in this class that took the
5 FCAT, so they're in a base year. If they have two
6 prior years --

7 MS. HEBDA: That's right, that's right. One
8 prior year. We use the two prior years if we have a
9 second prior year, but we have to have the one prior
10 year to measure growth.

11 Yes, ma'am.

12 MS. CORN: Another question about the
13 FCAT-related scores or FCAT-related courses, and
14 there are course codes in there that are actually
15 only inclusive of ESC like special diploma kids or
16 alternate assessment kids. How can those courses be
17 considered FCAT-related?

18 MS. HEBDA: Some of those courses, the kids in
19 those courses can be FCAT.

20 MS. CORN: But what if all of them in that class
21 did not?

22 MS. HEBDA: They wouldn't be included because
23 they don't have a score. They have to have the score
24 and they have to have the prior year score, but
25 students enrolled in those courses, those are all the

1 courses where a student could be enrolled who would
2 be eligible to take FCAT, but if they don't take it,
3 then they're not included. We can't include them if
4 they don't have new data.

5 MR. CONNER: Could you tell us a little bit, you
6 and I were talking about this briefly before, if the
7 base year always stayed 11-12, then you would have a
8 point from which you improve instruction. And our
9 history in Florida has been if we improve it, then
10 the board might just put another base year, another
11 base year.

12 MS. HEBDA: That's a great question. What we
13 tried to do in the rule was say -- and there's still
14 debate yet on whether we should use 11-12. There are
15 districts that are saying maybe we should wait until
16 we get 12-13 and do 11-12 and 12-13 together and
17 figure out that's going to be the baseline. There
18 are lots of -- this is our suggestion, because what
19 we were trying to prevent was every year you don't
20 know where you're going to end up because you've got
21 to wait and see how the other teachers do first.

22 MR. CONNER: Statistically you're going against
23 yourself if you change the base year.

24 MS. HEBDA: That's right, it could, it could.
25 So what we wanted to do was keep that solid, and I

1 think what we said in there was that the state board
2 would have to go back in five years and review the
3 standards, so try to put a time clock on it because,
4 you know, we could go on forever and nobody ever does
5 anything. We don't want that to happen. I think we
6 worded it so you wouldn't have to actually change the
7 standards every five years, but you would have to go
8 back and review them every five years to see if they
9 were set right. And we're going to need to review
10 them anyway because once we change from FCAT 2.0 to
11 the PARCC Assessment in a few years, then we'll have
12 to readdress again and start from whatever that
13 baseline is. So this is -- this would work for what
14 we have right now this very minute.

15 MR. CONNER: The base year data the districts
16 have received they could compute it themselves?

17 MS. HEBDA: Several of them already have, yes,
18 using these standards.

19 MR. CONNER: The state is in the process of
20 computing it for them so districts that do not have a
21 robust R and R would then receive data saying this
22 would be the impact that if it had been implemented
23 in the base year.

24 MS. HEBDA: Correct, correct. That's what we're
25 working on now. We're working on what we had

1 proposed right now and we're going to provide that to
2 the state board at the next meeting in March and just
3 so they can see and have an update. But then also,
4 as I said, we're getting these other suggestions. So
5 what we would like to do during the month of March is
6 also see if we employed those other suggestions too
7 or instead of or how that would work, what the impact
8 would have been at least looking backward, what that
9 would be.

10 Now, one of the considerations in using 11-12 as
11 the baseline you're also -- that folks have brought
12 up to us is that one of the benefits of running it,
13 of using the distribution you get from that year in
14 that year is it takes into account what happened that
15 year. That's how kids did that year. You know, the
16 reason for wanting to set a performance standard or
17 criteria and make it a criteria and reference
18 standard is we do want to have something that becomes
19 the anchor that people can then know what they're
20 shooting over. You know, is it 20 percent more
21 growth or 10 percent more growth or whatever that is.
22 But, you know, especially with being 11-12, some
23 people are concerned about that too because sometimes
24 there's an advantage of just using what happened that
25 exact year.

1 MR. CONNER: Are they going to equate FCAT 2 and
2 PARCC?

3 MS. HEBDA: There's going to have to be
4 something done just not for the sake of this, but
5 just to go from one test to another. I'm sure there
6 will be a lot of things that states have to do,
7 because we're in the 26- or 28-state consortium, so
8 we're not at this alone. There will be a lot of
9 states that are going from their own state assessment
10 to the national assessment.

11 MR. CONNER: There could be a good chance there
12 would be another baseline based on Park.

13 MS. HEBDA: Oh, yes. Right, right. Just like
14 we were saying. If the assessment changes we're
15 going to have to do this again, too. One of the
16 things we don't want to lose is how well this model
17 works. What I don't have in here is the impact data
18 from 11-12 that shows for teachers of these courses
19 how well it actually worked to level the playing
20 field with all those characteristics and everything
21 else in it. So we don't want to lose that going
22 forward. We started discussions with the folks that
23 do PARCC, and other states are in the same situation
24 we're in because lots of folks have their own growth
25 models and evaluation systems and all those things

1 that are running around the country now. So it's not
2 just Florida that's worrying about that, other people
3 are too.

4 Other questions before we go to comments? Okay.
5 We have 25 minutes for comments, so you should have
6 plenty of time to say whatever you would like today.
7 And then also keep, keep in touch with us and let us
8 know whatever else you would like. When we go to the
9 comment section, again, we won't respond and nobody
10 else will respond, so you will give your comments,
11 the court reporter will capture them, they will be on
12 the record as they are and we'll just say thank you
13 when you're done and then we'll adjourn the meeting
14 if there are no other questions after that.

15 So since there are just a few of you we'll start
16 from the front and go towards the back.

17 MS. MAHLMAN: First of all can I note that there
18 should be a percentage of samples?

19 MS. HEBDA: You can say anything you want in
20 your comment.

21 MS. MAHLMAN: Okay.

22 MS. HEBDA: Any suggestions, anything else you
23 want. That's what we want to hear. That's great.
24 And I'm going to retreat.

25 MS. MAHLMAN: Massive crowd.

1 My name is Dianne Mahlman and I'm a teacher at
2 Mariner High School in Lee County. And we just
3 recently got our VAM scores like everyone else and I
4 was shocked. So this is personal but it's also some
5 comments from other teachers. It's not against you.

6 Our teachers teach seven of eight periods, block
7 scheduling. There's approximately 175 students per
8 teacher. This is my 21st year of teaching. I have
9 taught elementary, middle and most recently high
10 school. Every year that Lee County provided the
11 opportunity for teachers to receive a monetary bonus
12 for student performance I received it except for the
13 first year of the portfolio system which I did not
14 apply. During the silo years I even received extra
15 money based on the fact I was in the top ten percent
16 of my silo in the county. All of these performance
17 bonuses were based on student growth comparing FCAT
18 to FCAT. Although my current overall VAM score was
19 effective, the student performance section showed I
20 needed improvement. The growth was comparing the
21 eighth grade FCAT to the ninth grade Algebra 1 EOC.
22 Where is the statistical growth comparing apples to
23 oranges?

24 Another concern of our teachers is the number of
25 students representing in the sampling. A sample is

1 the representation of the whole from which it is
2 taken. How can the scores of just five students
3 represent 175, yet one of our teachers received a
4 needs improvement on student performance based on
5 just that. while basing evaluations on reading, one
6 teacher had a sampling of five students. Four
7 exceeded the expected scale by a considerable growth
8 while only one did not. That means 80 percent
9 exceeded, yet the teacher's score was only 1.59,
10 barely in the effective range. How was that score
11 determined? Others were based on 14 and 17 students.
12 Please explain to me the logistics of evaluating a
13 teacher based on only approximately three percent of
14 his or her students.

15 Our teachers are certified in their specific
16 area of expertise. I have a math certification,
17 while others are certified in science, history, PE,
18 et cetera. The only teachers who are evaluated in
19 their specific field are Algebra 1 and reading. This
20 current year geometry will be added. The rest of the
21 teachers are evaluated on the reading scores of the
22 entire school. Yes, we all integrate reading into
23 our curriculum, but to evaluate nonreading teachers
24 whose students excel in their particular class is not
25 justified. Mariner has one teacher who taught upper

1 level math classes last year, had a sampling of 17
2 students. It was based on reading scores and
3 received a needs improvement student performance.
4 All but one of those 17 students passed the reading
5 FCAT but may or may not have dropped -- and made some
6 drop from a five to a four. Even the student who did
7 not pass the FCAT met the expectation. And we wanted
8 to know why decimals for -- why there are decimals
9 for expectation versus rounding for the actual
10 results. That can make a difference.

11 I understand that the Florida DOE has no policy
12 for credit denial due to absences, yet we are held
13 accountable for students who rarely come to school.
14 Our VAM scores were based on students from 211 to
15 212. We received those scores just last week. It is
16 now over halfway through the third quarter.
17 Evaluations are a tool for teachers to change his or
18 her teaching in order to increase the students'
19 success rate. I have approximately eight weeks
20 before the Algebra 1 EOC. What do you expect from me
21 to better my chances of not receiving another needs
22 improvement?

23 I have seen the VAM equation. Even as a math
24 teacher I find it extremely complicated. How do
25 teachers know that their scores are accurate and

1 fair?

2 I'm proud to say that Mariner High School
3 maintained its A grade. It's apparent that our
4 administrators are placing the right teachers in the
5 classes. We have no control as to which students are
6 placed in those classes. I currently have three
7 students -- I currently have students, one with a .3
8 AGPA and another with a .13 GPA. I can only do so
9 much.

10 Although the VAM evaluation sounds good, it is
11 not all realistic. There are too many variables that
12 don't make the evaluation fair and consistent for all
13 teachers. We need to be recruiting new teachers who
14 are eager to enter the profession and keep the
15 veteran teachers who have the experience and
16 expertise to make students successful in their
17 educational career. The current VAM evaluation
18 process is not going to do that.

19 MS. HEBDA: This is just a clarifying statement.
20 Didn't you want to say something else about the
21 percentage of students --

22 MS. MAHLMAN: Oh, yes.

23 MS. HEBDA: That was just a reminder.

24 MS. MAHLMAN: Because of the -- yes. Because of
25 what I said, when one teacher has 175 students and

1 five was the sampling, that was not just one. There
2 was a couple. That's preposterous. That's
3 ridiculous. There's got to be a percent. I realize
4 -- I have two daughters that teach elementary school.
5 I realize - and I taught too - that you only have so
6 many students, but maybe you could come up with a
7 percent of further sampling, because that was
8 ridiculous.

9 MS. HEBDA: Thank you. On the next row.

10 MS. MUTZENARD: Dianne covered about all of it.
11 There's one other factor that can't be --

12 MS. HEBDA: I'm sorry. Could you give your
13 name.

14 MS. MUTZENARD: Oh, sure. Donna Mutzenard. I
15 do have -- I did write it out. M-U-T-Z-E-N-A-R-D.
16 And I'm the Executive Director for Island Coast FEA
17 Service Union, which is the teachers and support
18 association here. As I said, Dianne covered about
19 all of it. The only other issue is -- and it was in
20 today's paper about tragedy strikes Immokalee again.
21 Our students, of course, today was Florida Writes,
22 which isn't part of the VAM score necessarily, but
23 this could happen the day before FCAT Reading or
24 anything else, a student was killed this week. And
25 in 21 months they've had eight Immokalee students

1 killed in some -- either through car accidents or
2 shootings or whatever. And those students today had
3 to do FCAT Writes. How do we figure that into a VAM
4 score? I know we have all those covariates in there
5 but you can't figure something in that and those
6 students are taking that test today. That's
7 something that can't be figured in anywhere, and yet
8 that could affect every teacher in that building,
9 what their VAM scores would be and what their ratings
10 would be.

11 MS. HEBDA: Thank you. Tom?

12 MR. CONNER: One comment. Tom Conner, Heartland
13 Educational Consortium. The little c, parens,
14 district-established performance level of standards,
15 is just a lot for my six districts that I represent
16 to swallow. We just really are struggling. And the
17 other alternatives are assigning what some of the
18 other folks have asked questions about, general FCAT
19 scores to teachers to -- that they would assume the
20 reading scores of students by being a part of the
21 literacy program for the school. But I don't really
22 know how we're going to do that well. And in meeting
23 with some high school principals who had to meet with
24 their teachers and discuss with them why they
25 received effective instead of a highly or needs

1 improvement based on data that is loosely connected
2 to their students that they teach and what they teach
3 is pretty demoralizing to those teachers. The AP is
4 workable. The industry certifications are workable.
5 But there's just so many teachers who aren't directly
6 connected to the data for which we would attach
7 significance to and all of our folks who are just
8 really struggling with how, how is that going to
9 happen. And when we looked to the state and the
10 state clearly says we can't do that for you like we
11 have done with FCAT, we cannot, and I'm not sure we
12 would want the state to develop any of the courses
13 for band or everything else that there is. But that
14 one the state board really needs to understand is
15 pretty frustrating to teachers. It's one thing to
16 teach reading in fifth grade and have reading FCAT
17 scores that you're accountable for. It's quite
18 another thing to teach fine arts at the high school
19 and be held accountable for reading scores or a
20 reading score from prior years that might predict an
21 end-of-the-course score for the course they teach.
22 They're worried about that, how to handle it.

23 That's all.

24 MS. HEBDA: Thank you. Anybody on this side of
25 the room wants to speak, make comments or

1 suggestions?

2 MS. MAHLMAN: Can I ask a question?

3 MS. HEBDA: You can ask anything you want.

4 MS. MAHLMAN: This is probably back on the other
5 part.

6 MS. HEBDA: That's okay.

7 MS. MAHLMAN: Geometry is coming forth. I mean
8 it's -- the ones, the scores we just got, geometry
9 teachers, it was based on the reading. I don't know
10 if that's statewide or if that's just something the
11 district comes up with or what.

12 MS. HEBDA: The district.

13 MS. MAHLMAN: It's the district. Okay.

14 MS. HEBDA: Which is what Tom was just talking
15 about.

16 MS. MAHLMAN: Okay.

17 MS. HEBDA: It's that little last sentence in
18 that paragraph that says a superintendent can --

19 MS. MAHLMAN: Right. Okay. Then maybe you
20 would know. With geometry coming into play for this
21 coming year or the year we're in now, what is it
22 going to be compared to? Is it going to be compared
23 to the Algebra 1 EOC? So with sophomores the growth
24 is going to show from the Algebra 1 EOC to the
25 Geometry EOC, which again is apples to oranges?

1 MS. MUTZENARD: I don't know.

2 MS. HEBDA: Here's one thing I can offer you.
3 As far as measuring growth on geometry, like we're
4 working on a model to measure growth on Algebra 1,
5 that's something the student growth committee is
6 going to work on doing and maybe we can find a growth
7 model that works and maybe we can't.

8 MS. MAHLMAN: As opposed to FCAT compared to
9 EOC?

10 MS. HEBDA: Right, right. And it also may be
11 that, and what is perfectly acceptable now is we
12 talked about how a district can use the data that
13 they have just to do an achievement measure instead
14 of a growth measure. For example, I've heard of
15 teachers that teach in eleventh and twelfth grade and
16 they have courses of just FCAT retakers. And their
17 focus, their goal for those kids is not necessarily
18 to make so many points of growth, but to get into a
19 path so they can graduate high school, for crying out
20 loud. So it is perfectly acceptable for the
21 district, because we don't include FCAT retakers in
22 the growth model, for them to say what we would like
23 for you to do, your standard is here's how many kids
24 typically across the state will pass FCAT on a retake
25 and that's what we're shooting for for you for

1 effective, and if you can do better than that, that's
2 highly effective or whatever those standards are.
3 They can use the data that they have that's related
4 to what the person actually teaches rather than
5 assigning a school rating.

6 what I think Tom is getting at is that -- and
7 I've heard third grade teachers saying it. There's
8 no growth for third grade FCAT, but people have third
9 grade results and if they're shooting the moon on
10 those third grade results and doing great, then they
11 want to get credit for that rather than being --
12 working to the school score. Of course they do.

13 I think where districts are having a struggle
14 with that, because that sounds perfectly logical, it
15 makes perfect sense, that's exactly what you want
16 because it's based on what the person is responsible
17 for teaching, there are, there are times when little
18 districts feel like they can't keep track of all
19 that. And so one of the things that's difficult for
20 them to do is to make a decision about what can I
21 logistically handle and what really is the best thing
22 for that individual teacher in the evaluation system.
23 And it's not just little districts. We talked with
24 Broward, Palm Beach and Dade yesterday, we were on
25 that side of the state, and their problem is the same

1 one, not because they don't think that's a great way
2 to look at that person's own data, but they have so
3 many people they can't keep track of them. They have
4 R and D departments, like Tom was talking about, they
5 have staff that do this kind of thing, multiple staff
6 that do it, but they still have to have a data system
7 where they feel like they can keep track of all that.
8 So I think it's going to take a lot of people working
9 together, pitching in and saying here are the
10 suggestions we have as teachers for the third grade,
11 here are suggestions we have for teachers of
12 eleventh- and twelfth-graders on what can be used and
13 here is a -- here are our suggestions for how we can
14 keep track of this at the school level for right now,
15 maybe later at the district level; here are the
16 procedures we use; everybody knows it's above board
17 and we're being fair about this. I think we're going
18 to have to have a lot of dialogue about that because
19 it's, it is -- I don't know how to explain that
20 someone's being assigned school data when they don't
21 touch those kids at some point during that school
22 year. But in absence of that, then what is done?
23 what is to be done and how do we keep track of it and
24 how do we work together to make sure that those
25 people get the data they need and that it's kept in a

1 way that everybody feels comfortable, that it's being
2 used properly and everybody is getting a fair shake
3 and all of those kind of things. It really does take
4 everybody's minds on it and everybody dedicated to
5 it.

6 MS. MAHLMAN: But in the meantime with all this
7 tweaking going on I have an evaluation in front of me
8 that says -- my overall evaluation said effective,
9 but I have an evaluation in front of me that says the
10 student performance needs improvement. I have never
11 had that in my life, you know. So two years down the
12 road if I'm still teaching, if retirement doesn't
13 come soon enough, and they finally tweak it to make
14 it reasonable and reliable and statistically okay, I
15 still have this, this current VAM score that I just
16 got. I mean I just --

17 MS. HEBDA: Here's the benefit, here's bonus
18 part of that. This is the good news piece. Your
19 evaluation for that year is what it is. I do not
20 know why you just got your results this past week.
21 That I could not possibly tell you.

22 MS. MAHLMAN: We did.

23 MS. MUTZENARD: Because our research department
24 was double-checking and triple-checking everything.

25 MS. MAHLMAN: They were supposed to come out

1 and then the deadline changed, and the deadline
2 changed and, you know, we just got them.

3 MS. HEBDA: So putting that aside --

4 MS. MUTZENARD: That was a local decision.

5 MS. HEBDA: Okay. Let's go back to the question
6 of what about this year's evaluation two years from
7 now, what does that mean.

8 MS. MAHLMAN: Right.

9 MS. HEBDA: You get evaluated every year the way
10 the law lays out the process and that one year
11 starting in 11-12, all we have is one year of data.
12 But at the end of this school year we'll have 11-12,
13 plus we'll have 12-13 data. So the way the student
14 growth portion is written, all data that you have
15 should be used in your evaluation, but you don't take
16 a needs improvement and average it with something
17 else. You just take all the data points and put them
18 together and you get a brand-new score. So it
19 doesn't matter how the district sets the cut point.
20 It mattered to you this year how they set the cut
21 point for 11-12, but if they reset their cut points
22 for 12-13 to deal with the issue that maybe it
23 shouldn't have been a needs improvement or whatever
24 they decide, I don't know the answer to that
25 question, but you're still taking all those data

1 points from those two years and putting them together
2 in this new evaluation. It's one year of
3 instructional practice, but it's up to three years
4 for the student learning data. Then the next year
5 when we hit 13-14, then you will have 11-12, 12-13
6 and 13-14 data, but those results for that year,
7 whatever they were, and they stand and they're over
8 with, every year you get a brand-new evaluation and
9 it's based on as much as three years' worth of data
10 if you have it and one year of instructional practice
11 for that year. So you don't have to -- it doesn't --
12 districts are changing, they're improving their cut
13 points, and eventually we'll have state cut points at
14 some point, but then those will count for that year,
15 You won't go back and change what happened in
16 previous years, and those points all get rolled up
17 together into data points determined under these new
18 standards. What does that mean? So what would have
19 been a needs improvement in one year actually could
20 end up being better the next year once it's all
21 rolled into a second data point.

22 MS. MAHLMAN: And then the other question that
23 teachers were asking, that third year down the road,
24 that's when money is tied in, right?

25 MS. HEBDA: As of this point for 14-15 that's

1 when the performance salary schedule starts. If
2 you're a PSC, if you have a professional services
3 contract, that's an opt-in choice for you. You don't
4 have to opt in.

5 MS. MAHLMAN: Right.

6 MS. HEBDA: You can stay on the grandfather
7 schedule. For brand-new teachers the performance
8 salary schedule will be there for them. They'll be
9 on it with the exception of anybody who is on those
10 temporary measures that we talked about, the 7(e)
11 paragraph, they stay on the grandfather schedule.

12 MS. MAHLMAN: Okay. Let's say - I won't - but
13 let's say three years down the road I would go to the
14 performance, which means then I would go on annual
15 contract, is that state?

16 MS. HEBDA: Yeah.

17 MS. MAHLMAN: Okay. People want to know what's
18 the difference in the money from highly effective to
19 effective. Nobody knows.

20 MS. HEBDA: Well, they don't know yet because
21 the negotiations aren't done yet. All the salary
22 stuff is negotiated. Just like salaries right now
23 vary from district to district because they're
24 locally negotiated. That's no different in
25 performance. There's a different framework for the

1 performance salary schedule but it's still locally --

2 MS. MAHLMAN: Even though it's a state -- even
3 though the VAM is state, it's not going to be a set
4 amount for the state, it's by the district?

5 MS. HEBDA: Salary or a mandatory selective
6 bargaining.

7 MS. MUTZENARD: There is a formula, though, that
8 the grandfather schedule has to be different than the
9 performance schedule by a certain percentage
10 depending on, so there is --

11 MS. HEBDA: So here's the difference between the
12 grandfather and the performance schedule. Within the
13 performance schedule alone, put the grandfather
14 schedule over here, in the performance schedule alone
15 there's a relationship that the framework requires
16 between the effective increase, and it's a salary
17 increase, it's not a bonus that comes and goes. You
18 get it and then it's your new base salary going
19 forward. Yeah, nice. So if you get an effective
20 increase one year, that's your new base salary for
21 the next year. Even if you're needs improvement the
22 next year you don't go back down. That's your new
23 base salary. So it has a relationship between the
24 base salary increase for effective and the base
25 salary increase for highly effective, which it has to

1 be either 50 percent or 75 percent of a relationship
2 between those two increases. So there has to be some
3 differential. You can't make it a dollar more.

4 Now, the relationship between the grandfather
5 schedule and the performance schedule is that
6 whatever the highest salary increase is on the
7 grandfather schedule no matter what step it is or
8 anything else, let's say it's \$5,000 and it's at step
9 23 or something like that, then the increase, the
10 salary increase for highly effective has to be
11 greater than that. So whatever the highest increase
12 is in the grandfather schedule, highly effective has
13 to be bigger than that. Doesn't say anything about
14 how big effective has to be in relation to that, only
15 highly effective.

16 MS. MAHLMAN: And when it comes to the
17 principal's evaluation of you, which is 50 percent of
18 the VAM --

19 MS. HEBDA: That's right.

20 MS. MAHLMAN: I think that I --

21 MS. HEBDA: Fifty percent of the overall
22 evaluation.

23 MS. MAHLMAN: Right. I think I counted like 21
24 criteria, 20 or 21 criteria, and maybe that's
25 district. I don't know.

1 MS. HEBDA: That's district.

2 MS. MAHLMAN: And then 19 of the, you have 19 of
3 those 21 to be highly effective. Okay. Chuckle,
4 chuckle.

5 MS. HEBDA: Yes. And then we'll go to Tom.

6 MS. CORN: I did have a comment.

7 MS. HEBDA: Okay. Official comment. State your
8 name.

9 MS. CORN: Linda Corn, C-O-R-N. I'm a teacher
10 at Fort Myers High School and five other schools at
11 the time on my VAM scores, which I am also very
12 disillusioned with because, like she, I have an
13 effective, you know, with my school and then I just
14 got a needs improvement based on six kids. So my
15 problem is that we were not told, you know, anything
16 about the roster verifications were all messed up and
17 they still are. So I know that's the district's
18 responsibility, but I just want it on the record.
19 I'm supposed to retire in two years and I'm in DROP,
20 so I'm actually thinking of, for my sanity, giving up
21 that \$50,000, you know.

22 MS. HEBDA: Any other comments or questions?
23 Tom.

24 MR. CONNER: I had one that one of our folks
25 raised. The use of the word --

1 THE COURT REPORTER: The use of the word what?

2 MS. HEBDA: Summative.

3 MR. CONNER: I'm sorry. They don't hear me at
4 drive-throughs either.

5 The use of summative still is a marriage of the
6 classroom walk-throughs, or in this case for this
7 individual that was rated needs improvement but got
8 effective overall, is that in 13-14, will that still
9 be the case?

10 MS. HEBDA: The way, the way this proposed rule
11 works based on what the law's required us to do, that
12 beginning next year it's still the combination of
13 50 percent, 40 percent, whatever the different things
14 are in the evaluation system in the district to
15 arrive at the summative rating, which is highly
16 effective, effective, needs improvement, developing
17 or unsatisfactory. That's what we're saying, that's
18 the final evaluation rating.

19 MR. CONNER: Okay. So it's --

20 MS. HEBDA: Now, let me finish though, okay.
21 Let me finish. That's beginning next year if this
22 rule -- if everything stays in the plan and
23 da-da-da-da-da-da. Sorry. You probably can't type
24 da-da-da-da-da-da. I was a music teacher. I'll
25 start singing in a minute and then you'll be really

1 in trouble. However, the law says we have to
2 establish those standards also so that they would
3 become the final rating of unsatisfactory or the
4 floor for effective and highly effective. So what
5 we said in the rule is that would only occur, we
6 would only use the performance standards for that
7 instance where they become the final rating in the
8 conditions where the teacher had three years worth of
9 data on the subject she was actually teaching. So
10 none of the temporary measures in the 7(e) paragraph,
11 only actual performance measures, the students that
12 you taught on your subject matter for three years
13 could it be used to determine the summative rating
14 alone.

15 MR. CONNER: But it would determine the
16 summative rating.

17 MS. HEBDA: In that condition that's what the
18 law says it needs to do, so that's why we put the
19 conditions in there. And when the conditions aren't
20 there, then it's just used for performance of
21 students and it's combined with the district's rating
22 and everything else.

23 MR. CONNER: So stating it another way, the
24 other 50 percent of classroom walk-throughs and
25 everything could not mitigate an unsatisfactory, an

1 effective or highly effective.

2 MS. HEBDA: It could not mitigate the
3 unsatisfactory rating if the conditions are met.

4 MR. CONNER: That's what I thought.

5 MS. HEBDA: It would contribute to or be the
6 floor for effective and highly effective. So if you
7 met whatever the standard was for effective, if all
8 the conditions were going in and this is what they're
9 used for, then the district's other part of the
10 evaluation would still be determinant on whether you
11 would be effective or highly effective. But you
12 couldn't be anything greater than effective --

13 MR. CONNER: Couldn't be greater but you could
14 be lower.

15 MS. HEBDA: -- unless you met the standard for
16 highly effective.

17 MR. CONNER: So it could only --

18 MS. HEBDA: I mean you could even be needs
19 improvement if you met the standard for effective in
20 that case. Depends on what happened in the rest of
21 your evaluation.

22 MR. CONNER: I would say as soon as you can get
23 the data out to the districts that would be great
24 because there really wasn't any about what the effect
25 would be. You said Broward had just a whole chunk in

1 development --

2 MS. HEBDA: That was Dade County. That was Dade
3 County.

4 MR. CONNER: Dade County?

5 MS. HEBDA: Uh-huh. Broward's was more evenly
6 distributed, they said.

7 Other questions? I know we're at time.

8 Okay. I appreciate everybody coming out and
9 asking questions and everything else. And if you
10 think of other things there's a website with an
11 automatic e-mail that blips us a comment and a
12 question or whatever else. And when the data are
13 ready and we present them to the state board those
14 will be available. We'll let you see what those are,
15 everything else.

16 MS. MAHLMAN: Just shocked this is all that
17 came.

18 (Workshop concluded at 6:04 p.m.)
19
20
21
22
23
24
25

1 STATE OF FLORIDA

2 COUNTY OF LEE

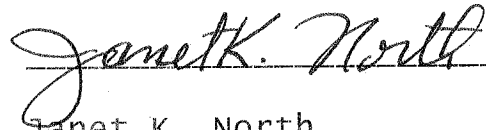
3 I, Janet K. North, RPR and Notary Public, do
4 certify that I was authorized to and did stenographically
5 report the foregoing proceedings and that the typewritten
6 transcript, consisting of pages numbered 1 through 99, is
7 a true record.

8 Dated this 14th day of March, 2013.

9

10

11


Janet K. North

12

13

14

15

16

17

18

19

20

21

22

23

24

25