



Virginia M. Barry, Ph.D.
Commissioner of Education
Tel. 603-271-3144

Paul K. Leather
Deputy Commissioner
Tel. 603-271-3801

STATE OF NEW HAMPSHIRE
DEPARTMENT OF EDUCATION
101 Pleasant Street
Concord, N.H. 03301
FAX 603-271-1953
Citizens Services Line 1-800-339-9900



◆ **MINUTES** ◆
8th Meeting for Phase II
March 27, 2012
Londergan Hall ◆ Room 15

The meeting opened at 4:00 pm with Karen Soule welcoming everyone and asking the State Representatives that joined the meeting to introduce themselves. Representative Michael Balboni, Representative Joseph Pitre, Representative Gregory Hill, Representative Mary Stuart Gile, and Representative June Frazer all introduced themselves.

Updates

Karen Soule

Karen reminded the group that we need to create a framework, despite the fact that New Hampshire is very locally controlled when it comes to education. She told the group that Phase II is about the educator evaluation piece. Karen explained that the InTASC (Interstate Teacher Assessment and Support Consortium) is what the New Hampshire general education and professional education standards are based on. Several SIG schools are using Danielson's framework as a basis for their teacher evaluation model. Karen then updated the group with the framework that the Principal Leadership Group has created.

Scott Marion

Scott began by introducing himself. He then spoke of the draft of The New Hampshire Model Educator Evaluation System that he created. It is meant as a guide to the group so that they can provide feedback to Scott. Scott then stated that the focus of today's meeting would be on the Core Principles that Ira Glick and Carol Kierstead created. He acknowledged the work of the SIG schools and said that he expects the work that they have done and the work that we are doing will strongly resemble each other.

The Core Principles

Scott spoke of how, based on everyone's feedback, the Core Principles have been changed a little bit to infuse the Core Principles from the Phase I Task Force with those of the Phase II Task Force. Scott asked everyone to read through the handout and identify anything that they disagree with on the left side ("Draft for Phase II Task Force"). He then asked everyone to look at the right side ("From Phase I Task Force") to determine their agreements, disagreements, and any missing items.

Carol asked everyone to individually go through the Core Principles handout and identify the key ideas and principles. Each person will then share a key idea that they think is important with their group and record it. The question that Carol then posed was, "Why have Core Principles?" She stated that it reflects the values of the Task Force and that it is a foundation for everything that is going to happen around the educator evaluation pieces.

Key Ideas of Core Principles

After this portion of the group work was completed, the following Key Ideas of the Core Principles were identified:

- The role of formative evaluation is not clearly articulated.
- It is not clear that administrator (evaluation) growth and learning is as valued as student and teacher.
- Statement 1 on the Phase I side is stronger than the ideas in Phase II (#8).
- Numbers 9 and 12 from Phase II and number 8 from Phase I (the concepts) should lead the description of Core Principles.
- Numbers 1, 2, 7, and 11 in Phase II go beyond the Phase I principles.
- Missing from Phase I evaluation records: Reference to New Hampshire definition of Teacher Effectiveness, training, frequency relative to Professional Development.
- Formative and summative evaluation.
- Collaboration of all stakeholders is important and missing from "key principles" (to support and build the model).
- Number 1 from Phase II is important, but might be better in a preface.
- Multiple measures for an evaluation.
- Student learning is an essential goal. It should inspire teachers to improve.
- Performance level descriptors are a necessary component.
- Differentiated rating system – experienced versus novice.
- Classifying teachers in performance categories.
- Collaboration is important, but it is missing on the Phase II side.
- Purpose of effective teaching and evaluation is growth.
- All New Hampshire schools, as determined by their districts, will classify all personnel.
- Number 8 from Phase II: Student learning is at the center of the teacher development continuum, pre-service through experienced educator.

- Combine numbers 6 and 7 from Phase II: Multiple measures are used to assess student learning and teacher effectiveness and effectiveness determinations shall be based on multiple years of data.
- The primary purpose of a framework for effective teaching is to enhance growth and learning of students and teachers.
- Classify teaching performance, not the teacher.
- Differentiation.
- Student learning is at the center.
- Common categories but different descriptors.
- Multiple measures.
- Not applied mechanistically.
- To be determined: Specificity of rubrics and are all personnel licensed?
- Define performance level descriptors.
- A “model” from which communities can pick and choose.
- Four Levels: Highly Effective, Effective, Approaching Effective, and Ineffective.
- Student learning at the center of effectiveness.
- Student learning is more than performance on NECAP. It includes creativity, thinking, problem solving, etc.
- System: Goals driven.
- Collaboration among stakeholders.
- The model on the left (Phase II) does not tie in to professional development and professional growth.
- Student growth appropriately measured – multiple, fair.

After everyone expressed their group’s key ideas, there was some conversation and one of the questions that was raised was how to delineate between classifying the person versus the performance? Scott answered by saying that we talk about this as teacher effectiveness, but it should really be teaching effectiveness.

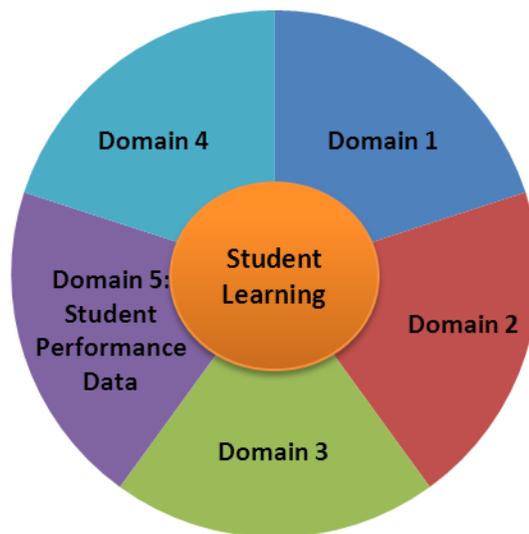
Dimensions/Domains of a Model System

After the key ideas group work, Carol asked everyone to get the handout that Scott had done (The New Hampshire Model Educator Evaluation System) and go over pages two through four and the top of page five. She told the groups that they were going to do a process called “A-B Teach”. A will read Dimensions and Domains and teach their B partner through summarization, identifying key points, and sharing thoughts and questions. B will read Standards of Practice and teach their A partner through the same methods. Carol then asked everyone to summarize their key thoughts and questions regarding the Dimensions and Domains and Standards of Practice and put them on the flip charts.

The following is a list of what was written on all of the flip charts.

- Student data is important and 20% is good.
- The correspondence between Danielson and Phase I is convenient.
- Danielson terms are good.

- Work has been done with her rubrics and that would be nice/valuable to share and use.
- Why define?
- What are nominal and effective weights?
 - Is 20% for student performance enough? What if it's less?
 - Are the data points equal for all teachers? Use of whole school data?
- Standards – There is a challenge for alignment.
 - Admin Rules (610.02) = INTASC derived
 - Danielson framework is convenient; it is well developed to be used for evaluative description but it is not organized like the admin rules.
 - Will Danielson be a good way to measure these?
 - Will Danielson framework be updated?
- Which student performance data points will be used? How? E.g. Value-added? Growth? Status? “May” use versus “Must” use.
- Evaluators – Training
- Competency



*All five domains are equal

- Absolute – The definition of effective teaching must come from Phase I work!
- Weighting of domains should be consistent across the state?
- Living document.....
- Difference between nominal and actual weights (1/7 or 20% - 5 Point Domains)
- Subdomains

- Phase I – Phase II – Do they truly align? (For example: Would learner and learning be lost in classroom environment?)
- Descriptors – Above and beyond
 - Danielson’s domains?
 - Look at work of Kim Marshall – Simplification, Walk-throughs
 - How will we determine what specific descriptors equal effectiveness?
- Key Question: Can we use portfolio for domain 2 and 3, too?
 - If A is the Phase I Report and B is the model that we are looking at right now, they don’t really agree. And if B is the model and C is where we want to go, then we still don’t see that we agree.
 - $A \neq B$ } Phase 1 = 5 Domains
 - $B \neq C$ } Where is core value student learning driver?
 - Student voice – “It’s not like this every day.”

Nominal and Effective Weights

Scott was asked to explain the difference between nominal and effective weights. Scott said that the five domains appear to be equal weighting at 20% each, but if something has a lot of variability, one of the components has a lot of variability and true variability where people are spread out with their knowledge skills, and another one of those domains where they are not spread out, the thing with the greater variability will carry more effective weight than the stated weight of 20%. For example, if we are all being rank ordered on two things, and we all score the same on one of those things, which one is going to carry all of the weight in that rank order? The thing with variables. So we need to think about how we are going to weight these? Things with higher liability would often carry more weight. Scott said that we are not going to worry about these right now, but we are going to have to think about them in the future. If you have something that doesn’t vary as an indicator, you either have a problem in the thing you’re measuring, a problem in your measurement, or both.

Danielson’s Framework

Carol spoke of Danielson’s framework and how it is the only framework that it is tied to student outcomes. She was saying that in this line of work, it is asked that research based and evidence based practices be used. She then asked why you would go with something else when this is a proven framework and has a high level of rigor.

Scott cautioned that there will need to be some fidelity checks because 80% of the districts say that they are using Danielson. However, Scott said they are not truly using Danielson if they only observe once every three years.

Commonality and Comparability

Scott said that simply having a common definition of what effectiveness is does not necessarily mean that we are going to get commonality. How important are commonality and comparability? For example, would an effective teacher in Fall Mountain be an effective teacher in Manchester? Or vice versa?

Student Voice

It is difficult to see how we can ask 3rd and 4th grade students to evaluate their teacher, (or high school students, for that matter). People say that they will tell children that it is not consequential, but how

long will it be before they realize it is consequential? Student voice is important. The question is how do you do it in a way where it's not more dangerous than it is helpful?

Tracy Bricchi spoke of how a number of districts have used the "My Voice" survey to identify what the tone is across our district. That could be placed under the "shared attribute" piece of student data. Part of the student data is shared amongst the building and it could be looked at as, "What is the overall pulse of the building, (rather than having it on a specific teacher)?"

Scott asked if school climate was in the Principal's Evaluation and it is. Randy Bell then asked the following questions:

- Who knows the teacher best?
- Who interacts with the teacher best?
- Who spends the most time with the teacher?
- Who is most impacted by the teacher?

And the answer to that was that it is not the principal. It is the children. So how we can get our hands on legitimate student data is difficult and dangerous, but we should not avoid it. Someone can observe a teacher three times per year, but the students live with that teacher 180 days per year. That is data and insight that we somehow need to get our hands on.

We could collect information that could be useful to help inform the practice, even if we don't put it in the accountability system, per se.

"Bring Me Some Rocks"

Scott is asking that everyone read the Student Learning Objectives handout and think about it. This is one approach for non-tested subjects. This will be talked about at the next Teacher Effectiveness Task Force meeting on April 10, 2012.

Scott asked what people needed to see to try to advance in this process. Student learning is definitely at the center of this.

- Someone recommended going with Danielson. Scott said that if we are going to use Danielson, then we are going to use the Danielson rubrics. He also believes that once you prove you are an effective teacher, you could be evaluated every other year, rather than every year.
- Flexibility for local control is important. The SIG schools will provide an excellent example of how Danielson is working for them and how they have changed it a little to meet their needs.
- There needs to be some autonomy about the data and not just some hard number. There needs to be some conversation or dialogue about the data.
- Could we outline the decision points and prioritize them?

Questions to Consider

- In terms of the Domains, do we want to recommend that it is Danielson, or not?
- Do we want to make a recommendation that we should have a common set of performance level descriptors across the state?
- In the general measurement framework, do we want to talk about weighting where we have about 20%, at least a nominal weight for each of the five areas?

Feedback

Scott should be sending out more information by April 6, 2012. Please send your feedback to Scott after reading the Student Learning Objective handout and the next document that he sends out to everyone.

The meeting convened at 6:31 pm.

Submitted by Trisha Allen

March 30, 2012

DRAFT