



**NH DOE Accountability Task Force
September 28, 2010, 9:00 am– 12:00 pm, Room 15**

**Next Meeting: Thursday, October 14, 8:30 am – 12 pm,
Joint Meeting with the Commissioner's Task Force, Location TBD**

Attending:

District Representatives: Patrick Connors (SAU 53Epsom); Heather Cummings (Gov. Wentworth); Donna Crook (MSD); Kathy Stavenger (SNHU)

NH DOE: Paul Leather, Sally Fellows; Mary Lane; Marcia McCaffery; Lauren Heiter; Rob Tenney, Michael Schwartz, Time Kurtz, Ken Relihan

Center for Assessment: Scott Marion, Damian Betebenner,
Measured Progress: Shannon Douglas

Meeting Objectives:

- ◆ Provide an update on the status of the INPUT component submissions, decisions
- ◆ Report on progress of the Commissioner's Task Force to develop the Performance Based component of the SB180 Accountability System
- ◆ Review sample results for representative schools
- ◆ Provide recommendations for point values / weighting across performance indicators

Notes

1. Deb Wiswell invited members to introduce themselves for the benefit of new DOE data consultant, Lauren Heiter. She reviewed the agenda for the meeting.
2. Updates on 'homework' from the July 29 meeting:
 - Special Education population: data on the distribution of disabilities by classification will be shared later during subgroup discussion
 - Hispanic students: the concern was whether the Hispanic students are being included if the system does not explicitly include that subgroup; after analysis of the data, DOE determined that, 2/3 are included in the low socioeconomic status subgroup, 1/3 of the Hispanic students are included in the whole school group.
3. Deb provided an update on the Input system, demonstrating the web access portal and how to log in. The title of the tool is, "Accountability for the Opportunity to Provide an Adequate Education." Deb reminded members of the resources available for schools, a user manual and instructional video. She reviewed the varying levels of permissions assigned that are specific to each log-in account. Marcia

McCaffery clarified that there is security and privacy associated with the various log-in levels to protect editing and viewing until the form is submitted.

Deb continued touring the system showing the tab format for each of the 12 required standards. She reported that out several problems have occurred fairly frequently. One is an oversight by principals who overlook additional tabs for some standards that have criteria cross several areas, such as the arts. Another problem being reported is neglecting to SAVE during submission. The system times-out after 30 minutes so SAVE is recommended regularly.

Deb is currently working on designing the process for Department curriculum consultants and others to review entries by schools to determine where attention is needed.

Patrick Connors reported that, as a principal, completing the submission was a lot of work. In his district, responsibility was divided across several staff members for different standards/ sections. He proofread the entries before the report was finalized, and then it was sent to the superintendent for final approval and submission to the DOE.

Questions:

Q: Where will this information be posted when completed?

A: Deb indicated this is not yet finalized. She expects there will be some public presentation of the responses, but it may be the responsibility of the schools or districts to post their submission. Ultimately, the where, when and in what form publication will be done depends on the legislature's decision. The Department will publish the 'results' of its review but not certain how at this point.

Q: Does the input system replace the School Approval Site Visits checklist?

A: No, because only 12 school approval standards are included in the Input system per the legislation. Site visit teams examine all 30+ approval standards. The submissions on the Input System will be helpful to site visitors to allow them to confirm the school's responses while on site, especially for those where schools indicate "other" or "no" as a response.

4. Deb shared a Sample summary of high school data for the state defined indicators ("Part 1") of the Performance Based System (see DRAFT data on handout).

Reminder: this chart shows ALL the proposed indicators displayed as points added together, but NO DECISIONS have been made about the relative weight of each indicator in the final decision, or whether all of these indicators will be included.

Highlights:

Sample Schools: selected to represent geography, size, known performance challenges or successes

Reading Index scores: rubric points allocated by those scoring "above the state average" earning 3 or 4 performance points; subdivisions roughly equal

Reading Percent Proficient: the state average is lower than for index scores, so the point ranges differ; still "state average or above" earns 3 or 4 points;

Math Index: math scores statewide much lower than reading; state average of 66.4 separates 3 or 4 points from lower points earned;

Math Percent Proficient: particular problem because state average very low (31); allocated points along a subjective score distribution;

Improvement points: the Department ranks schools according to changes in their index scores from the prior year to the most recent year; in this Sample, high schools in the top 20 of the improvement ranking for reading and math each earn 1 "bonus point"

Science Index: used last year's data; state average separates 3 or 4 points from 1 or 2;

Graduation rate: uses 2009 'new' cohort calculation; state average = 81.5, 3 or 4 points awarded if above state average, 1 or 2 points if below, with at rate of 75 as the bottom

Dropout rate: this sample uses the 2009 data without GED or college; state average = 9.4, 3 or 4 points awarded if better than state average (lower dropout rate); 1 or 2 points for higher/ worse rates;

Discussion:

-- what about other indicators of high school performance, such as course taking/ course credits, college retention, AP participation, AP scores, or SAT scores?

Response: indicators on the sample summary are those selected by the Commissioner's Task Force to date; if other data is available, the system can be amended

-- prefer to be cautious about using data that only meet the strictest standards of validity and reliability; high statistical standards restricts the use of data that may be "fair" measures of performance, or consistently "trustworthy";

-- reminder that the Commissioner's Task Force has proposed an optional "part 2" of the performance based system that allows schools to propose locally defined indicators, goals and targets to demonstrate they have provided the opportunity for students to receive an adequate education

-- one key question in the design of the system is how the data will be used; making decisions with the end use of the data in mind can inform planning judgments

-- in the end, the system will be judged by whether it reports something meaningful and whether the conclusions of the system are defensible

Deb asked participants to offer their thoughts about relative "weight" of the different indicators on the sample set. In the "additive" system on the sample, state test scores earn points for whole school and at least two subgroups, and, if using a 4 point range, schools would accumulate from 4 to 12 points for each of the two core subject areas and for science. She also asked if there should be any "absolutes," for example if a schools earns a "1" on any category, they cannot be judged as meeting the adequacy standard.

Comments:

-- it seems justified to allocate high value to student performance on state assessments since that is the core function of schools

-- is valuing the status score more important than judging the trend in scores, especially for high schools where there is currently insufficient student achievement data to calculate a student growth percentile.

-- including both index AND percent proficient seems to overvalue NECAP in comparison to other performance measures already collected; (discussion seemed to prefer index as a value more sensitive to differences in distribution of scores)

-- will have to include an appeals process, particularly for student achievement at high schools, to allow a school to explain why an illness epidemic or other catastrophic event renders a particular test result an invalid measure of typical performance

TO DO:

1. Deb will prepare distribution graphs of existing scores for each of the sample measures for review by this committee and the Commissioner's task force.

5. Scott transitioned the conversation to a review of work on the elementary and middle school performance indicators, beginning with decisions of the Commissioner's Task Force.

Overview of decisions to date (slide 2) –

Open question = how to accommodate ELLs; subcommittee charged with developing a proposal

Participation measure options: (slide 4)

Slides 5 through 10 shows statewide data for participation; note that all schools report 95% or above participation, so there is little discrimination between schools using this measure

Some considerations: Does it make sense to develop a rubric to disaggregate minor differences between rates of different subgroups? Is participation rate important enough to be included?

Should participation rate be 'conjunctive,' i.e., all schools must meet 95 for all subgroups or "no" points are earned participation? (E.g. 1 or 0 points toward adequacy) (slide 11)

Remember that the federal rule stays in effect (if participation is lower than 95%, the school does not make AYP); however, the state does not have to follow the federal rules for the adequacy system

Discussion:

- give point to those doing what "we" want to represent 'adequate'
- mathematically, allotting points for participation doesn't change a lot, but philosophically makes a statement about what a school is supposed to be doing
- in a way, participation is like a primary grades report card that includes "S" or "N," satisfactory or not; while it is not a "grade" it still provides information to the parent/ reader
- what about schools where there are not enough students for a subgroup? Currently the state examines data over 3 years if needed to accumulate enough data to meet the current minimum "n" of 40 students; could use a similar procedure for the adequacy system
- another option is to average actual participation rate across all subgroups and allocate points for the final average participation rate;

Recommendation of the Group: If the school achieves 95% participation for either 1 year or on a 3 year average if needed to meet minimum "n", the school earns a point towards adequacy.

6. Damian discussed the bubble charts for subgroup performance using Student Growth Percentile (SGP).

Whole School (Bubble Graph 1): : each bubble is the median student growth percentile for a given school; pop up also tells the target median value and the count of students; includes IEP, low SES and ELL; broader distribution of bubbles/ SGPs

Students with IEP (Graph 2 (pink bubbles)): Deb reported the results of her research on this group. She met with Mary Lane to research the distribution of classifications of student disabilities. Those students with speech, language, specific learning disabilities, ADD, etc. for whom specific interventions can improve academic achievement account for 90% of the students identified as students with disabilities. Only 10% of students in the IEP subgroup have significant cognitive challenges and are not appropriately measured with NECAP. Some are assessed with the ALT assessment, which is included in the statewide reports. These data suggest that it is valuable to examine results for the IEP subgroup, especially since the student growth percentile is calculated by comparing each student to those with similar assessment history.

Low socioeconomic status (Graph 3 (green bubbles)): low SES students only, excluding those included in the IEP subgroup; the distribution of bubbles reflects higher academic achievement than those for the IEP students; there is a wider range of growth (X axis); pop-up shows actual median SGP, count of tested students in the subgroup, and median target SGP; second example shows a school with high achievement, low target, in which the school doesn't meet its growth target

"All Other" group – (non IEP, not low SES, not ELL); generally higher achieving, distribution along the X axis (growth) centers around 50th percentile

The distribution of “bubbles” on the charts are consistent with expectations for the different subgroups, but also provide visual evidence that the range and distribution vary within groups across schools.

7. Scott initiated a discussion of another way to use student growth percentile (SGP) to assess adequacy. Previously, the group had considered creating “bands” on the bubble graphs to allocate points across the distribution from high growth high achievement (upper right) to low growth/ low achievement (bottom left). The challenge of that approach is how to justify where the bands are placed. It could appear arbitrary to schools.

Scott proposed an alternative using a procedure similar to that developed for Colorado. The Colorado process looks at two factors: (1) whether the school meets the median growth percentile target and (2) the observed/ actual median growth percentile for the current year. This approach takes into account the fact that schools with high achieving students will generally have a lower “target” which is a projection of the growth needed for students to attain or maintain proficiency within three years. Schools with lower-achieving students will have a more aggressive target. Points are awarded based on actual median SGP, but score ranges needed to earn points vary by whether the target was met or not met. (see slide 18)

Scott presented the M-T- graph (slide 16) on which the M is the actual median SGP, T marks the median Target SGP, and the length of the green bar shows the percent of students who met their individual SGP targets.

To use the Colorado approach, check first if the observed median SGP (M) meets the target SGP (T). If $M > T$, find the actual median SGP in the YES column to determine how many points are earned (slide 18).

If the observed median SGP (M) DOES NOT meet the Target SGP (T), find the actual median SGP in the NO column to determine how many points are awarded.

	Did median SGP exceed target SGP?		
POINTS	YES	NO	POINTS
4	56- 99	70-99	4
3		55-69	3
3	45-55	40-54	2
2	30-44		1
1	1-29	1-39	1

Discussion:

- could we allocate full points if a school meets its target, with no differential for the “toughness” of the target?
- this two step process may be difficult to explain; is there a simpler way to make fair judgments about adequacy using SGP?

8. Scott asked for input from the group on relative weighting of the achievement, growth, attendance and participation indicators(slide 35).

Discussion:

- weight the indicators based on what’s more important; academic subject areas are certainly important

- should growth count more than status (e.g., yearly scores on state assessments) ? What about science which doesn't yield a growth measure?
- could we look at growth through other indicators, such as AP participation? For example, previously schools could make AYP if their graduation rate was improving; USED eventually I disallowed this exception; not bound by federal regulations in this state system

9. Meeting concluded at 12:20 pm.

TO DO:

1. Calculate the results from schools across the state using the Colorado approach to SGP. Check to see if the results are reasonable and realistic.
2. Graph the distribution of "yes" and "no" for selected sample schools.

Next Meeting: Thursday, October 14, 8:30 am – 12 pm, Joint with the Commissioner's Task Force

(chart papers)

High School Questions:

- Participation
- Attendance
- Truancy Rate
- Improvement data (?) one year; over time
- Index vs. % proficient

Other High School Data

- College Retention
- Course completion (variety, rigor)
- AP courses
- State Scholars