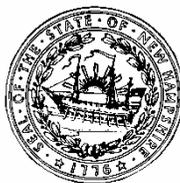


Virginia M. Barry, Ph.D.
Commissioner of Education
Tel. 603-271-3144



Paul Leather
Deputy Commissioner of Education
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**

**Commissioner's Task Force to Develop a Performance-Based
School Accountability System
August 10, 2010, noon to 3 pm, State Board Room**

NEXT MEETING: Tuesday, September 21, 9 am – 12 pm

Attending:

	Virginia Barry, Ph.D.	Commissioner of Education
X	Brian Cochrane	Director of Assessment and Accountability Nashua School District
	Paul Couture	Principal, Stevens High School, Claremont
X	Jerome Frew	Superintendent, Kearsarge Regional School District
	Molly Kelly	Chair, Education Committee, NH Senate
X	Daphne Kenyon	NH State Board of Education
X	Paul Leather	Director, Division of Adult Learning, NH Department of Education
X	Scott Marion	National Center for the Improvement of Educational Assessment, Dover
	Deborah McNeish	Principal, Conant/Rumford School, Concord, NH
	Judith Fillion	Director, Division of Program Support, NH Department of Education
X	Edward Murdough	Bureau of School Approval
	Kathleen Murphy	Director, Division of Instruction, NH Department of Education
	Emma Rous	Chair, Education Committee, NH House of Representatives
X	Vincent Spiotti	Bethlehem School Board, Bethlehem, NH
X	Deborah Wiswell	Bureau of Accountability, Curriculum and School Improvement, Division of Instruction, NH Department of Education

Guests

Representative Franklin Gould, NH House, Lebanon, NH
Tim Kurtz, Administrator for Curriculum and Assessment, NH DOE
Karen Laba, Consultant, Committee Documentation

NOTES:**Meeting Objectives:**

- ◆ Revisit, confirm student growth percentile discussions, decisions
- ◆ Examine subgroup data and graphs and reach consensus on definitions
- ◆ Provide feedback on a proposed method to construct a "system" incorporating all adequacy indicators

1. Deb opened the meeting and welcomed those who were in attendance. She provided an update on the November report due to the Legislature in November, explaining that the report is expected to describe results of the pilot of the system. This allows more time for further refinement and development on critical decisions by this committee.

Deb reported that the "Input System" is in operation, running fairly smoothly. She has received numerous inquiries for assistance, but many schools have already submitted responses. She and Ed Murdough will plan a process to respond. One school requested the text box be expanded; if additional requests come in, she will allow items to be submitted electronically outside the system rather than increasing system limits.

Anecdotally, a side benefit from the input system has been reported. It has encouraged superintendents to accelerate posting key curriculum documents on the district website so that they can be referenced in the response to the adequacy standards included in the input system.

Deb convened a group of Dept. staff to practice 'standard setting' for the student growth percentile results using actual NH schools data. The student growth percentile (SGP) data for whole schools was judged to be consistent with what would be expected, given what is known about schools and districts around the state. The SGP results do appear in line with other indicators of school performance.

Today's meeting will focus on further discussion of ways to look at student achievement and growth that offers the most useful information to allow judgments of 'adequacy' that are credible and understandable.

2. Scott began a review of recommendations from the Accountability/ AYP Task Force on the designation of subgroups of students to be included in the adequacy calculation. Recommended groups include:
Special education students (only)
Economically disadvantaged students (low SES (socioeconomic status))/ not special education or ELL
"All others" i.e., all students not special education, ELL and not low SES
Whole School

A separate recommendation for ELL students will be presented later.

Q: Are there questions that might come from governance or special interest groups about student groups who are not listed as part of the adequacy report?

A: There may be questions, but data on all student groups exist and is reported at the school, district, and state levels in the NCLB categories. So even though a particular group may not be part of the adequacy report, information about each group can be accessed. In presentations to the state board, for example, it would be wise to include all the available data and show that all students are included in at least one of these four designated adequacy groups.

Q: What about the minimum "n" for small schools?

A: Some schools may be below 5 for subgroups. We can make a case to keep the "whole school" group because of small schools who may not have enough students to be reported in subgroups.

To Do: Scott will do an analysis of Hispanic students to see if any do not meet the criteria for inclusion in the low SES group or ELL group.

3. Tim Kurtz began a discussion of minimum 'n'. Current NCLB/ AYP minimum is >10. NH believes the smaller cell size allows the smaller sub groups to be 'counted' in AYP determinations. Currently the state reports scores for groups as low as 11. Only one or two other states use such a small cell size. Option of a small n / cell size allows many more schools to all be counted on the same number of indicators, since even small schools meet the number of students in each subgroup. There is some concern about students being identifiable if using a small 'n,' and issue of privacy. Using an 'n' of 5 for special education data has been considered acceptable in the past.

Scott pointed out a distinction between conjunctive vs disjunctive systems to explain the importance of considering 'n' or cell size. The current NCLB/ AYP system is a conjunctive system, where any one failure is a failure across the whole school, which is why 'n' (or cell) size is important. A state with a low 'n' holds schools accountable for more of their students. A disjunctive system considers each indicator independent of the others and an acceptable rating on one is sufficient to satisfy the criteria. The adequacy system being developed here is a compensatory system, where "scores" on each indicator add up to a collective total which will then be evaluated to determine adequacy. Each indicator has value that contributes to a school's total.

Tim proposed an 'n' of 5 for this system, and Scott indicated he used that value to prepare the data reports for today's discussion.

4. ELL special consideration: NCLB requires English learners to take the state's content area assessments after only one year of instruction in schools. NH currently uses the ACCESS English Language Proficiency (ELP) test, which results in a score rating a student's English language proficiency and a second score describing that student's rate of progress toward English proficiency.

Discussion followed on the particular NH challenge of English proficiency. For example, in Manchester, 100+ dialects are spoken as the primary and/or home language, sometimes with little literacy in the first language. Yet, students are expected to be prepared to take the state content assessments after only one year in school.

Paul Leather pointed out that there is a state commitment to develop policies around English learners. A committee has been convened and will meet over the next few months. Scott Marion noted that this committee (CTF) could report to the Legislature in November that another state committee is preparing recommendations for ELLs.

Tim Kurtz restated the recommendation to use the ELP progress measure for adequacy, until students demonstrate proficiency in English. In mathematics, the data is less clear that English proficiency is necessary. Tim cautioned that this would not replace federal accountability rules currently in place.

Scott pointed out that English learners tend to show dramatic growth when they reach English proficiency. Deb and Tim reported that data has been run to show relationships between ACCESS scores and NECAP results and will be included in documentation of the adequacy system.

5. Tim began a discussion of the possibility of using "zeros" in place of indicators such as "participation" rate. In recent discussions, schools either met the federal participation standard (95%) or not, earning a 1 or a 0 for that indicator in the adequacy system. To overcome the possibility of some

schools not testing low performing students, Tim proposed allocating a "zero" score to all students not tested, rather than setting a minimum participation rate as in the federal system.

Options proposed: use the federal minimum of 95% and allot score of 0 or 1 (not met/ met); allot a score of 0 for students who are not tested; allocate points based on a range of participation rates as proposed for graduation and dropout rates (see handout).

Discussion:

- concern about penalizing a school for students not tested through no fault of their own
- schools already held accountable for participation through the federal system
- eliminate participation rate as an indicator for the adequacy system
- current participation requires a minimum of 40 students (and can be aggregated over three years)

Further conversation needed – Scott will develop some numbers to see how various participation options impact the adequacy judgment.

6. Scott reviewed the recent proposal and discussion to use a "4 quadrant" approach to allocate points to schools using the "bubble graphs." Last week's AYP meeting included a discussion of how to define the "quadrants" and therefore the point values for the sections of the bubble graph. One option is to decide where to draw horizontal and vertical lines, creating 4 quadrants. Another option considered was to create bands (slide 8) that give the lowest point values to those in the lower left corner of the grid. Higher points would be allocated the further the school scored on both achievement and growth. In one suggestion, the bands would be created using horizontal then slanting vertical lines to add greater (or lesser) value to growth or achievement.

When Scott met with others at the Center for Assessment, they discussed the defensible nature of the "bands" approach. The decision was to try to return to a criterion-based measure, namely the actual and projected target SGP (student growth percentile). At the May meeting, members of this committee examined sample graphs showing actual and projected student performance against proficiency (horizontal lines, slide 11). These discussions addressed whether students were on target to attain proficiency.

Discussion followed about the differences between the current growth model and this proposal.

Questions were raised about how to explain the target calculation/ model/ graph.

- reference to "panty hose" charts as examples many can understand
- greater familiarity with bubble charts among the general population
- caution that sometimes simple presentations/ solutions aren't the most valid

Scott presented his suggestion to aggregate the observed SGP (student growth percentiles) for each group of students in a school and calculate the median of those values. Then, conduct a comparison with the median aggregated projected targets for SGP. If a school's average observed SGP is greater than the median target SGP, then the school – on average -- is serving its students well so that they are meeting or exceeding the average SGP targets. Graph of this data shows number of students who met their target (green line); median growth target (circle T) and median observed/ actual SGP (white circle M). [Slide 14]

Discussion

- when comparing the bubble charts with the O-M-T graphic, the bubble graph is easier to understand and explain
- school names on the O-M-T line graph leads to direct comparisons but not always to the correct interpretation of those comparisons
- concerns expressed about how the information will be presented in the media
- how the information is displayed can inform or lead to misconceptions

- if schools see that high SES schools have lower median SGP 'targets,' that may diminish support for this approach
- on the 'bands' graphic, a suggestion was made to use concentric circles, with 0,0 as origin, and points allocated for schools that fall within each circle/ ring

Scott pointed out Colorado's solution to the high SES, low targets problem (slide 16). CO allotted points to schools that met or didn't meet their SGP target, with variations for level of difficulty; namely the rate of growth the school needed to achieve earned different rubric points which rewarded high growth schools. Scott described how to read the results he ran using NH data (slides 20-25). For the "all others" group and the low SES groups, the median observed SGP is greater than the target SGP. For the SWD (students with disabilities) group, the opposite case appears, observed SGP is lower than the Target SGP. This is consistent with historical data, since students with disabilities typically fall below proficiency and need to grow at a faster rate to reach their targets.

Scott also described the rubric scores for the actual data (mathematics) for all three subgroups, averaged and graphed (slides 26-34). Using Colorado's approach of a rubric to allot points for meeting SGP targets would still show schools with large percentages of students with disabilities earning fewer points than others.

Due to the difficulty in explaining the O-M-T graphs and likelihood of misinterpretation of the information, consensus among committee members was to make best use of the bubble graphs to represent achievement/ growth results.

7. A discussion followed whether dramatically poor performance on any one indicator for any one subgroup would lead to a finding of 'not adequate' in this system. Scott reminded the members that a compensatory system allows for success on one indicator to balance out lack of success on another. But, should poor rating (a "1" on the rubric just described) of any group of students be understood to mean the schools is not providing an opportunity for an adequate education?

Discussion:

- a compensatory system makes more sense to people, seems to be a more common sense approach; the state system is not only about reliability and validity, as demanded by the federal system
- we don't want to continue the major flaw of the federal NCLB AYP system, where failure of one group designates the whole school as failing
- maybe we do not want a fully conjunctive or fully compensatory system, but maybe include a trigger where dramatically poor standing on a key indicator results in a finding of "not adequate"
- what's the distribution of SWD? Which ones could reach the achievement targets? Some Dept. data is being collected and examined to answer this question
- also remember that the system is being designed so that schools can demonstrate adequacy using locally defined targets if the state-defined ('level1') measures do not reflect the school's work
- in addition, success on the input system can earn a school the adequacy rating
- the state system should offer the flexibility to demonstrate a school is providing the opportunity for an adequate education -- or better -- in some areas, though may not be in all areas
- Idea -- could there be different outcomes for different findings? For example, if SWD in lower left corner could still be declared 'adequate,' but scoring in that range would identify the school for special services/ support/ attention from the Dept.
- the legislature seems to have intended a compensatory system, assessing adequacy across the whole school; it would be difficult to have one measure or subgroup trump the accumulation/ balance of all the other indicators or measures
- it appears this committee needs further discussion of whether "not adequate" can be found for a low SGP bubble chart rating;

- concern that the system can't make "everyone OK"; if all are found "adequate," the system won't pass the test of common sense
- strong inclination to continue to bring attention to school's efforts to serve ALL students, not just "most" or the easy to educate students, which would be more like the current federal system where all subgroups must make targets

Scott reiterated the central question: If a particular subgroup scored particularly poorly, would the committee accept an 'adequate' designation because other subgroups or other indicators added up to a satisfactory rating? Or withhold the designation?

- from an administrator's perspective, sometimes a "lousy" evaluation offers an opportunity to improve;
- withholding the designation of 'adequate' would be acceptable, if it is in response to the rare "abysmal" performance;
- don't forget to consider districts that have clustered or focused special education services in particular schools in order to provide more efficient services;
- Scott proposed that 20th percentile growth be considered "abysmal" performance in this discussion
- additionally consider what the real data looks like, which and how many schools would end up in the lower left corner of the bubble graph.

TO DO: Scott will create bubble graphs for the four groups discussed to date (whole school; low socioeconomic status; students with disabilities; 'all others' – not SWD, not ELL, and not low SES).

8. Closing discussion: changing a lot, much of it may be hard to explain. If keeping the system the same where one group has a large impact on the overall adequacy judgment, it will be a hard sell. The bubble charts leave no place to 'hide.'

On the other hand, the system may not be as hard to communicate as anticipated. Decisions in this group are complex because we need to consider all the options and each in depth. Once the components of the system are finalized, presenting it will be more straightforward than our discussion here.

TO DO:

1. Run the scores for actual schools for all groups to see who falls into the lower left most section of the bubble charts.
2. Run multiple sets of data in multiple ways – 9 boxes, quarter concentric circles, grid, etc.
3. Run the various participation options (as in section 5 above).
4. Do an analysis of Hispanic students to see how many do not fall into other subgroups (see item 2, above).

Next Meeting: Tuesday, Sept. 21, 9 am – 12 pm, Board Room