Recommendations Regarding the Louisiana School Accountability System

Submitted to
the Louisiana Department of Education
October 26, 1998

By the Accountability Technical Advisory Committee[1]

[1]The Accountability Technical Advisory Committee members are: Jane Armstrong, Education Commission of the States; Dale Carlson, California Department of Education; Eugene Kennedy, Louisiana State University; and Brian Gong and Richard Hill, The National Center for the Improvement of Educational Assessment, Inc.

BACKGROUND

The Louisiana Department of Education has established an Accountability Technical Advisory Committee (TAC) to make recommendations about several issues regarding the design of the school accountability system. The framework for these issues was the *Recommendations for Louisiana's Public Education Accountability System (March 18, 1998),* generated by the Louisiana District and School Accountability Advisory Commission (Commission) and adopted, with modifications, by the Board for Elementary and Secondary Education (BESE) in May, 1998.

The Department of Education asked the TAC to focus its attention on those technical issues of the accountability system on which the BESE delayed adoption, along with other issues which had been little discussed. These issues included the weights for components, formulas for combining components, formulas for setting expectations for school progress, and criteria for determining consequences, such as corrective actions. This paper presents the TAC recommendations for these issues.

The TAC met in Baton Rouge, LA on September 3-4, 1998. At that meeting, the Department of Education asked the TAC also to consider several issues that were less technical and more operational. These issues covered a wide range of concerns in an operational testing and accountability program, including design of score reports, information about what other states have done regarding testing accommodations for special education students, and effective communication strategies regarding commonly misunderstood issues in assessment and accountability. The TAC's initial responses to these requests were provided directly to LDE staff at the meeting, and are not considered germane to this document. Finally, several issues were raised in the meeting which were largely policy issues. The TAC did not express opinions on those issues, considering them more appropriately addressed by policy makers within Louisiana. The TAC met in Baton Rouge, LA on October 19-20, 1998. The TAC's recommendations were refined through discussions at that time with Department staff and the Commission. Those revisions are reflected in this report.

ACKNOWLEDGEMENTS

The Louisiana School and District Accountability program has benefited from extensive work by many groups, including the Board for Elementary and Secondary Education (BESE), the Commission, the state Department of Education, and others. The TAC viewed its role as providing technical advice that builds as much as possible upon the strong foundation established by these other groups. In most instances, that entailed considering possible options and then providing recommendations about specific procedures to accomplish the accountability policies and designs already recommended by the Commission and adopted by BESE.

ORGANIZATION OF RECOMMENDATIONS BY THE TECHNICAL ADVISORY COMMITTEE

The Accountability Technical Advisory Committee (TAC) made recommendations within the framework of the design for a comprehensive accountability system already established by BESE. In general, the recommendations fill in technical details and procedures needed to fully specify an operational accountability system. Most of these deal with the establishment of scores and formulas that flesh out the school accountability design.

The recommendations of the TAC are presented in the following order:

1. *Assumptions Underlying Recommendations*
2. *10-Year and 20-Year Goals and Scores*
3. *Determining A School Score for the CRT, NRT, Attendance, and Dropout Indicators*
   - 3.1 CRT Weights and Formula
   - 3.2 NRT Weights and Formula
   - 3.3 Attendance Formula
   - 3.4 Dropout Formula
4. *Combining Components: Example of Calculating a School Performance Score*
5. *Generating Growth Targets*
   - 5.1 Formulas for School Growth Targets
   - 5.2 Examples of Expected Growth for a School
6. *Generating Accountability Judgments*
   - 6.1 Accountability Cycles
   - 6.2 Setting a Minimum Score ("Bottom bar")
   - 6.3 Setting a "Top bar"
   - 6.4 Growth Labels and Accountability Consequences
7. *Dealing with Special Cases*
   - 7.1 Start-up of Accountability System
   - 7.2 Special Cases of Student Accountability
   - 7.3 Special Cases of School Accountability
8. *Appendix – Summary*

Where several steps or intermediate formulas are involved, a recommendation is provided for each. Several examples are given that illustrate the formulas. Where the TAC discussed significant alternatives or other issues, that discussion is briefly encapsulated following the example. All the recommendations are summarized in the Appendix.

1. ASSUMPTIONS UNDERLYING RECOMMENDATIONS

The TAC agreed on the following assumptions to guide the assessment and school accountability design. These design principles are drawn from and/or elaborate those recommended by the Commission and approved by the state Board of Education (BESE).

1. Report and include in the accountability formulas the four indicators:
   a) CRT (criterion-referenced test),
   b) NRT (norm-referenced test),
   c) Student Attendance,
   d) Dropout rate (Grades 7-12).
2. Scale each indicator so that each has a 10-year goal of 100 and a 20-year goal of 150.
3. Report each indicator to the nearest 0.1 on the index.
4. Create a composite school index ("the School Performance Score") by summing each indicator times its weight.[1]
5. Weight the indicators as originally recommended to BESE: a) CRT 60% of the total, b) NRT 30%, c) Attendance 10% for Grades K-8, 5% for Grades 7-12, d) Drop-out 5% for Grades 7-12.

[1] These recommendations focus on school accountability. The TAC will consider district and student accountability issues if requested by the Commission and/or the Department of Education.

2. TEN-YEAR AND TWENTY-YEAR GOALS AND SCORES

Recommendation on Goals and Scores

Issues

What should be the 10-year and 20-year goals for the CRT, NRT, Attendance, and Dropout indicators?

What numeric scores should be established as goals?

Background

The Commission recommended that 10-year and 20-year measurable goals be set for all public schools in Louisiana on each of four indicators of student achievement:

- Criterion Referenced Tests (CRT);
- Norm Referenced Tests (NRT);
- Student attendance; and
- Student dropout rates.

For each of these indicators the Commission recommended a 10-year and a 20-year goal. Those recommendations are shown below.

The Commission recommended that School Performance Scores be calculated in such a way that a school attaining the 10-Year Goal on each indicator would have a School Performance Score of 100. The Commission did not recommend what the 20-Year Goal Score should be, nor did it specify how the School Performance Scores would be calculated.

10- and 20-Year Goals Recommended by the Louisiana District and School Accountability Advisory Commission		
Student Indicators	**10-Year Goal Score = 100 2009 – Grades K-8 2011 – Grades 9-12**	**20-Year Goal 2019 – Grades K-8 2021 – Grades 9-12**
CRT (Criterion Referenced Test)	Average student performance at BASIC proficiency level[1]	Average student performance at PROFICIENT level
NRT (Norm-Referenced Test)	Average = 50th Percentile	Average = 68th Percentile
Attendance Rate	97% (K - Grade 6) 95% (Grades 7 & 8) 93% (Grades 9 – 12)	98% (K - Grade 6) 96% (Grades 7 & 8) 94% (Grades 9 – 12)
Dropout Rate (Grades 7-12)	4% (Grades 7 & 8) 8% (Grades 9 – 12)	3% (Grades 7 & 8) 6% (Grades 9 – 12)

Recommendations

The TAC agrees that a technically sound accountability system can be built using the four indicators.

CRT Goals

The TAC endorses the 10-year and 20-year goals recommended by the Commission for the CRT. These goals were an average student achievement in each school of Basic and Proficient, as 10-year and 20-year goals, respectively.

NRT Goals

The TAC recommends that the 10-year and 20-year goals for the NRT be set at 55th and 75th percentile ranks, respectively. For technical reasons, the TAC recommends that NRT standard scores be used in accountability system calculations rather than percentile rank scores.

Attendance and Dropout Goals

The TAC agrees substantially with the Commission recommendations regarding Attendance and Dropout, but recommends some slightly different goals on the basis of empirical data and for technical reasons that are discussed in more detail below.

10-Year and 20-Year Goal Scores

The TAC recommends that the 10-Year Goal Score be set at 100 for each of the indicators, and at 100 for the composite School Performance Score. The TAC recommends that the 20-Year Goal Score be set at 150 for each of the indicators, and at 150 for the School Performance Score.

[1] The Commission report states, "100% of students at an average proficiency level of BASIC," which might be interpreted as meaning that all students should be at the Basic level. The TAC was informed that the intent of the Commission was that the CRT 10-year goal is that the average performance across students within a school should be Basic. The same note applies to the CRT 20-Year Goal.

The TAC recommendations on school accountability goals and scores are shown in the table below. The TAC felt these goals were ambitious but attainable, given the commitment and resources that will be devoted to improving educational achievement in Louisiana over the next several years.

School Goals and Goal Scores Recommended by TAC				
Student Indicators	**10-Year Goal**	**10-Year Goal Score (100)**	**20-Year Goal**	**20-Year Goal Score (150)**
CRT	Average student performance at BASIC proficiency level	100	Average student performance at PROFICIENT level	150
NRT	Average Composite standard score corresponding to the 55th percentile rank in the grade level	100	Average Composite standard score corresponding to the 75th percentile rank in the grade level	150
Attendance	95% (Grades K-8) 93% (Grades 9-12)	100	98% (Grades K-8) 96% (Grades 9-12)	150
Drop-out	4% (Grades 7 & 8) 8% (Grades 9 – 12)	100	2% (Grades 7 & 8) 4% (Grades 9 – 12)	150

Discussion

Goals are very important. School improvement goals should be clear and credible to educators, policymakers, and the public. The goals should be high, yet realistic, so schools can show progress whatever their current state of achievement, and be ultimately obtainable. The goals and formulas for each indicator should meet the scale requirements of equaling 100 for the 10-year goal and 150 for the 20-year goal. The goals and formulas should proportionally acknowledge progress on each indicator; i.e., the effective weights should be similar to the nominal weights. The goals and formulas should lead to reports that are feasible to produce, understandable, and useful.

The TAC believes that the goals and associated formulas it recommends meet the criteria described above. The particular goals are discussed more in each specific section following.

As an example of considering whether goals were intuitively understandable, the TAC did briefly discuss emphasizing the 20-Year Goal by making it 100. The TAC did not pursue this discussion since there was general agreement on the desirability of setting 100 as the 10-year goal for policy and scaling properties.

3. DETERMINING A SCHOOL SCORE FOR THE CRT, NRT, ATTENDANCE, AND DROPOUT INDICATORS

Issue

What numeric scores should be set for the 10-Year and 20-Year Goals?

Background

The Commission recommended that a composite score be calculated for individual schools, called a *School Performance Score*. The *School Performance Score* would range from 0 to beyond 100, with 100 indicating a school has reached the *10-Year Goal*.

The Commission recommended that factors entered into the formula have varying weights to indicate their relative importance in determining student academic success. The Commission recommended using the following weights to calculate School Performance Scores:

Indicators	Grades Administered	Weights
Criterion-Referenced Tests (CRT)	Grades 4, 8, 10, 11	60%
Norm-Referenced Tests (NRT)	Grades 3, 5, 6, 7, 9	30%
Student Attendance	Grades K-6 Grades 7-12	10% 5%
Dropout Rates	Grades 7-12	5%

The Commission recommended the components and their relative weights, but did not address how to calculate a School Performance Score, nor how scores for each indicators should be calculated.

Recommendations

20-Year Goal Score

The TAC supports all the recommendations made by the Commission regarding the School Performance Score. In addition, the TAC recommends that a score of 150 be used to indicate that a school has reached the 20-Year Goal.

This report provides detailed recommendations on how to calculate an overall School Performance Score and a school accountability score for each indicator: CRT, NRT, Attendance, and Dropout (for middle and high schools). Those recommendations are given below.

3.1 CRT Weights and Formula

Issue *How should a school score be calculated for student performance on the CRT (criterion-referenced test)?*

Recommendation The design goal was to generate a single composite index score of school performance on the CRT. To achieve this, the TAC makes recommendations for each of the following:

1. Determining values for each performance level of student achievement
2. Combining student achievement into a CRT index score for each content area
3. Generating a school CRT index score that combines all content areas included on the CRT.

3.1.1 Recommendation on CRT Performance Level Values

Issue *What points (values) should be assigned for the CRT performance levels of Advanced, Proficient, Basic, Approaching Basic, and Unsatisfactory?*

Background The Commission noted that the new CRTs have five performance levels: Advanced, Proficient, Basic, Approaching Basic, and Unsatisfactory. Focusing primarily on student achievement levels promotes understanding of the standards and eases communication. However, a way must be designated to provide numerical values for the performance levels. By assigning a point value to each performance level, a school average for the criterion-referenced tests can be calculated.

Recommendation

The TAC recommends that the values for Performance Levels on the CRT be:

Performance Level	Values
Advanced	200
Proficient	150
Basic	100
Approaching Basic	50
Unsatisfactory	0

Discussion

Values for Basic and Proficient

The TAC started from the goals established by BESE and the Commission. If the 10-year goal is to have the average student performance be at the Basic level, and the 10-year goal is an average index of 100, then the value of a Basic score should be set to 100. Similarly, if the 20-year goal is to have the average student be Proficient and the 20-year goal is an average index of 150, then Proficient should be set to a value of 150.

Values for other performance levels

In reaching this recommendation, the TAC also considered the important issue of whether the values should be evenly spaced between performance levels. Considerations included whether movements between the categories were equally desirable, and whether the movement between categories was roughly equally difficult. The TAC recognized that the values also should produce an understandable and credible score.

The question of what values to give the performance levels other than Basic and Proficient is largely a policy issue, but the experience of the TAC in observing other states has been that even spacing of points between performance levels is generally considered equitable and the most easily understood system. Thus, assigning values of 0 to Unsatisfactory and 50 to Approaching Basic seems most reasonable, as does the assignment of 200 to Advanced. However, BESE could change the weights for these levels without negatively impacting the basic accountability design as proposed.

Initial results available to the TAC indicated that these values, in addition to being understandable and credible, would have technically desirable properties. The school scores for Mathematics, using the proposed values, was distributed:

Mathematics	Obtained School Index Score[1]
Maximum school score	105
75th percentile in state	70
50th percentile	48
25th percentile	33
Minimum school score	3

[1] These results are for the sample of spring 1998, LEAP field tests in Mathematics using tentative standards that have not yet been approved by BESE. The sample included approximately 100 schools at both grades 4 and 8. The example reflects the proposed weights for Advanced, Proficient, etc., of 200, 150, etc., respectively.

3.1.2 Recommendation on Formula for CRT Index for each Content Area

Issue

How should student scores on the CRT (criterion-referenced test) be aggregated into a form suitable for school accountability?

Background

The Department is implementing the new Louisiana criterion referenced testing program (LEAP for the 21st Century, or LEAP 21), which will include English, mathematics, science, and social studies. The Commission referred to tests being designed by Louisiana to measure whether a student meets state expectations of what a student should know and be able to do. The CRT tests are designed to be harder and require students to use more complex thinking skills. These tests will be taken by students in Grades 4 and 8, and the Graduate Exit Examination will be taken in Grades 10 and 11.

Recommendation

The TAC recommends that a performance index for each of the four content areas be calculated for each school. The CRT index for a content area would be the average performance level of the students tested in the school on that content area. The average performance level would be calculated as the number of students at each performance level times the values for those performance levels, divided by the total number of students.

Example

The following example is for a hypothetical school with 50 students in Grade 4. The example shows the students' performance on the CRT in mathematics, with the distribution of students from Advanced to Unsatisfactory as shown:

Example School: Mathematics, Grade 4			
Performance Level	Number of Students	Points/ Student	Total Points
Advanced	1	200	200
Proficient	9	150	1350
Basic	10	100	1000
Approaching Basic	15	50	750
Unsatisfactory	15	0	0
Total	50		3300
Mathematics Index = 66.0 (=3300/50)			

To calculate the index for a content area, such as mathematics,

- multiply the number of students at each performance level by the number of points for each student at that performance level. For example, multiply 1 Advanced student by 200 points for each student = 200 total points; multiply 9 Proficient students by 150 points per student = 1350 total points.
- add the total points together. For example, 200 total points for Advanced plus 1350 total points for Proficient, plus the total points for Basic, Approaching Basic, and Unsatisfactory = 3300 total points.
- divide the total points by the total number of students to give an index for the content area. The 3300 total points divided by 50 students equals 66.0, the CRT Mathematics Index for the sample school.

This procedure would be followed to generate the CRT Index for each content area on the CRT.

3.1.3 Recommendation on Formula for Total CRT Index (Composite of All Content Areas)

Issue

How should the student performances in the four CRT areas of English, mathematics, science, and social studies be combined into a form suitable for school accountability?

Background

The Commission recommended that one accountability indicator be reported for a school's performance on the criterion-referenced tests, but did not specify how scores from multiple content areas on the CRT be combined.

Recommendation

The TAC recommends that the CRT Index be calculated using student-level data, by assigning points to each student for each content area, based on the Performance Level achieved by the student; summing all points earned; and dividing by the number of students times the number of content areas.

Multiple grades

If a school has multiple grades of the CRT (e.g., a K-8 school would have the CRT in grades 4 and 8), the same procedure would be followed in calculating the overall school CRT Index.

Example

An example is given below of how a school's total CRT index would be calculated.

Sample School		English/Lang. Arts		Mathematics		
Performance Level	Points/ Student	Number of Students	Total Points	Number of Students	Total Points	Total
Advanced	200	3	600	1	200	
Proficient	150	6	900	6	900	
Basic	100	9	900	9	900	
Approaching Basic	50	17	850	12	600	
Unsatisfactory	0	8	0	15	0	
Total		43	3250	43	2600	5850
Index			**75.6**		**60.5**	**68.0**

In the example, only English/Language Arts and Mathematics are included, as will be the case in 1999 for Grades K-8. Assume that the school has 43 students who take the CRT. These students are

distributed as shown. For English/Language Arts and Mathematics, the index scores are calculated as described previously, yielding an index score of 75.6 for English and 60.5 for Mathematics.

The school's total CRT index is calculated by taking the total points for both content areas and dividing by the total number of tests taken. This yields an overall CRT Index of 68.0.

Discussion

The TAC assumes that each content area is of equal value, and so does not recommend differential weights (e.g., weighting English/Language Arts more than science).

This CRT index has the properties of having a value of 100 when the average student performance on the CRT tests is Basic (the recommended 10-year goal) and a value of 150 when the average student performance is Proficient (the recommended 20-year goal).

The TAC recommends using student-level data to calculate the school index. This is more accurate and operationally robust than averaging across content index scores at the school level. In the example, note that if the average were calculated of the English Index of 75.6 and the Mathematics Index of 60.5, then the composite CRT Index would be 68.1, due to rounding error, instead of 68.0.

3.2. Recommendations on NRT (Norm-referenced Test) Formula

Issue *How should student scores on the NRT (norm-referenced test) be included in accountability?*

Background The Commission recommended including norm-referenced tests in the accountability system. Louisiana has recently adopted Riverside Publishing's *Iowa Tests* for the statewide norm-referenced testing program. Based on a recent change to the testing law, norm-referenced tests will be given in grades 3, 5, 6, 7, and 9. In 1997-98, the *Iowa Tests* were used for the first time and several metrics were reported, including standard score and percentile rank of the average standard score.

Recommendation The TAC recommends that a school's NRT index score be put on a scale of 100 for the 10-year goal and 150 for the 20-year goal, as are the other indicators.

To achieve this, the TAC makes recommendations for each of the following issues:

1. Selecting a unit for the NRT
2. Generating a school index for the NRT that has the desired scale properties
3. Deriving a school score for multiple grades if necessary

3.2.1 Recommendation on Unit for NRT Scale

Issue *What metric should be used to calculate and report the school results on the NRT (norm-referenced test)?*

Background The Commission discussed various metrics for the school NRT accountability index. It desired a metric that was simple to use, familiar to the public, and technically sound for accountability purposes. The Commission did not recommend a metric, although it

discussed standard scores, percentile ranks, stanines, quartiles, and quintiles.

Recommendation The TAC recommends that the NRT 10-year and 20-year goals be the 55th and 75th percentile ranks, respectively.

The TAC recommends that the NRT Index for each grade be based on the average standard score on the NRT Composite score.[1] Since standard scores (and by extension, average standard scores) have no inherent meaning, the TAC recommends that the percentile rank of the standard score be reported along with the standard scores. The standard scores recommended as goals were taken from the publisher's percentile rank-to-standard score conversion tables. The TAC recommends that the standard scores selected for the 10-year and 20-year goals at each grade be the publisher's standard scores for those percentile ranks. More specifically, the recommended standard scores for 10-year and 20-year goals are given in the following table.

ITBS Composite Standard Scores Equivalent to the recommended 10- and 20-Year Goals, by Grade Level[1]

		Grade				
Goals	Percentile Rank	3	5	6	7	9
10-year Goal	55th	189	220	232	245	266
20-year Goal	75th	201	237	253	268	290

[1]Source of percentile rank-to-standard score conversions: *Iowa Tests of Basic Skills, Norms and Score Conversions, Form M* (1996) and *Iowa Tests of Educational Development, Norms and Score Conversions, with Technical Information, Form M* (1996), Chicago, IL: Riverside Publishing Co.

[1] The Composite score is the score based on the student's overall performance on the reading/language arts, mathematics, science, and social studies sections of the test.

Discussion

Goals

The TAC recommended the 55th and 75th percentile ranks as the 10-year and 20-year NRT goals. The TAC recommends the higher goals based on several reasons, including: a) achievement in the state on the NRT is starting fairly close to the 50th percentile on average, and a significant number of schools are already above the 50th percentile; b) experience in other states has shown that NRT scores usually increase rapidly over the initial several years of instituting a higher-stakes testing program, so it is reasonable to expect even higher scores on the NRT in the next 10-20 years; and c) the higher NRT goals correspond more closely to the achievement standards embodied in the CRT portions of the accountability system and to the implicit standards of what other high-performing states' NRT scores might be.

There was considerable discussion of what goals would appropriately reflect Commission's intent of significant, achievable growth. The ITBS data available to the TAC indicated that the average state performance in 1997 was approximately at the 40th percentile, including special education students. Achievement was approximately at the 43rd percentile for students who took the test under standard conditions. However, in Grade 4 mathematics, over 23% of the schools were above the 50th percentile in 1997, and 1% were above the 68th percentile. In light of the current achievement of the schools, the higher goals seemed reasonable and appropriate to the TAC.

Index

The TAC considered generating an NRT index from the students' individual scores in Reading/Language Arts, mathematics, science, and social studies, but decided that this approach was unnecessary for school accountability purposes and needlessly complex. Using the NRT Composite score for accountability purposes will not prohibit the reporting of student scores in each content area, and schools, teachers, parents, and students may be expected to use this more detailed information from the NRT in a variety of ways.

3.2.2 Recommendation on NRT Index Formula

Issue

How shall NRT scores be related to the 10-Year and 20-Year Accountability Goals?

Background

NRT results are reported in a variety of ways by the test publisher. However, the usual reporting statistics are not related to Louisiana's goals and are not suitable for direct use in the school accountability system. There needs to be a method for converting the performance of students in a school on the NRT to an accountability NRT index score.

Recommendation

NRT Index

The TAC recommends that an NRT index be created, based on a straight line for which the standard score for the 10-year goal is associated with a score of 100 and for the 20-year goal, 150.

Each grade will have its own relationship between standard scores and the index, since there are different standard scores associated with the 10-year and 20-year NRT goals for each grade. The formulas for Grades 3, 5, 6, 7, and 9 are included in the summary. The examples in this paper are based on the 55^{th} and 75^{th} percentile rank goals recommended by the TAC, but the principles could be easily extended to other goals.

Calculating the NRT Index

The TAC recommends that the school NRT Index be calculated by using the formula to convert each student's standard score into an index score, then averaging the student index scores.

Lowest score

The TAC recommends that zero be the lowest NRT score used in school accountability reporting or calculations, although it is possible for a school to have a negative index. The lowest student index score would be the actual calculated index, based on the student's standard score.

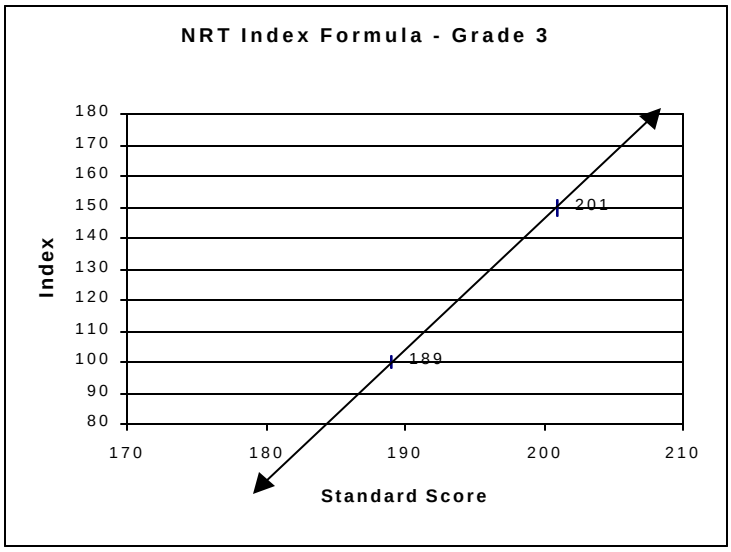
Highest score

The TAC recommends that no constraint be placed on the highest NRT index score.

Example

To illustrate, consider the NRT data for grade 3. The desired line will have the property of having an index equal to 100 when the average standard score for a school is 189 (the 55^{th} percentile, or 10-Year Goal for third graders) and an index of 150 when the average

standard score is 201 (the 75th percentile, or 20-Year Goal for third graders). This is shown in the graph below.



The line that meets the above conditions for grade 3 is:

$$Index_{Gr3} = (4.167 * SS) - 687.5$$

where SS is the average standard score.

When the formula relating average standard score to NRT accountability index has been determined, it is possible to calculate the needed information for a school's NRT accountability report.

An example is given below of calculating the NRT Index for grade 3 for a hypothetical school:

	Grade 3	
	SS	Index
Student 1	170	20.9
Student 2	179	58.4
Student 3…	175	41.7
Student n	168	12.6
Number Included	63	
Average Standard Score	183.5	
Index	**77.1**	
Percentile Rank of Average Standard Score	46	

Given the equation,

$$Index_{Gr3} = (4.167 * SS) - 687.5$$

it is a matter of simple substitution to calculate the index for each student. For example, substituting a standard score of 170 into the formula produces an index of 20.9 for Student 1. These student index scores can be averaged to produce a school index. (It is equivalent, within rounding error, to produce an average standard score for all the students, and then use the formula to produce an index score.)

The *Iowa Tests* standard score of 181.2 is associated with the 40th percentile rank in Grade 3, and now, by extension, so is the NRT Accountability Index of 67.6 for that same grade. The percentile rank is obtained from the publisher's standard score-to-percentile rank tables.

Discussion

Converting standard score to NRT accountability index score

These formulas are necessary to convert the standard score so that a performance at the 55th percentile (or its standard score equivalent in the grade) is equal to the 10-Year Goal of 100, and performance at the 75th percentile (or its standard score equivalent in the grade) is equal to the 20-Year Goal of 150. The same approach would be followed to establish formulas based on other NRT goals, e.g., 50th and 68th percentile ranks.

Lowest score

These formulas yield a zero index score at the 9th–14th percentile rank, depending on the grade. A chance response level on the NRT is approximately equal to the 10th percentile. That is, if a student responded by guessing on every question, that student would be expected to score at approximately the 10th percentile.

Highest score

These formulas yield an index score of 200 at approximately the 91st percentile rank. At the 99th percentile, the index scores range from about 235 – 265. The maximum index scores for standard scores above the 99th percentile are about 310 – 435.

Standard scores and percentile ranks

The TAC discussed whether to use standard scores or percentile ranks in making the calculations of a school's average score. While percentile ranks have an inherent meaning that standard scores do not, it is technically incorrect to average percentile ranks. Therefore, the TAC's recommendation is to use standard scores to calculate averages, but then to add meaning to the resultant average by always reporting the associated percentile rank. Thus, the

recommended approach is to report the average standard score and the percentile rank of that score alongside it. This additional statistic is referred to as the "percentile rank of the average standard score."

Alternatives

The TAC considered several other possible approaches for generating an index measure of NRT achievement and growth. These approaches included:

a. Dividing performance on the NRT into fifths (quintiles), quarters (quartiles), or stanines. These approaches were perceived as being potentially more familiar, simpler, and/or similar to the CRT performance levels. For example, using quintiles, it would be possible to make the cutscores of the five NRT categories match the initial percentage distributions on the performance categories of the CRT. These approaches were not recommended because they are less sensitive and reliable than the recommended method, and because the similarity to the CRT and previous uses might be confusing, especially if there were not parallel achievement and score changes on the NRT and CRT over time.
b. Converting the NRT percentile rank into an approximation of the CRT standard score (the NRT percentile rank ranges from 1 to 99, so NRT percentile rank x 2 would range from 1 to 198; CRT ranges from 0 to 200). The TAC viewed this approach as susceptible to misunderstanding and inferior for scaling purposes to the recommended approach.
c. Consolidating student performance by reporting the percentage of students in the school scoring above the 50^{th} national percentile on the NRT. This approach would be simple and easily understood in terms of achievement. However, the TAC did not recommend this approach because it is substantially less reliable and also less sensitive as an indicator of progress; i.e., it would only reflect growth at the 50^{th} percentile.

The TAC believes that the recommended method represents the best combination of simplicity, reliability, technical accuracy, and acceptable policy of all the approaches considered.

Standard scores and renorming of the NRT

The recommended approach of using standard scores also avoids another problem peculiar to percentile ranks. If the goals or the reporting statistic were based on percentile ranks, Louisiana would encounter confusion partway through the accountability process when the publisher renorms the test. Today, for example, the 55^{th} percentile of the composite score at grade 3 is a standard score of

189. That standard score of 189 is the 10-Year Goal for third grade classrooms in Louisiana. Suppose there is steady improvement in student achievement throughout the United States over the next decade, and the publisher renorms the NRT eight years from now. At that time, the 55th percentile might be a standard score of 192. If that happened, there would be confusion over what Louisiana's goal should be—the meaning of the 55th percentile would have changed. However, the meaning of standard scores does not change over time, so establishing goals in terms of that statistic provides for consistent interpretation over the years.

Student Index Scores The TAC recommends that NRT index scores be used for school accountability purposes. Student NRT index scores are needed for those calculations, but should not be used in student reports. For student reporting, the NRT reports from the publisher should be used.

3.2.3 Recommendation on Calculating NRT Index for Multiple Grades in a School

Issue *How shall an NRT Index be calculated when there are multiple grades that take the NRT in a school?*

Background Since there will be norm-referenced testing in grades 3, 5, 6, 7, and 9, it is likely that most elementary and middle schools will include two grades of NRT test results. A way must be specified to combine the data that will be fair, technically sound, and operationally feasible.

Recommendation The TAC recommends that a single NRT accountability index be calculated for each school. When a school has multiple grades that take the NRT (e.g., Grades 3 and 5 in a K-5 school), then the school NRT index should be calculated by calculating the NRT index for each student, then averaging the student indices across all students in all grades in the school.

Example The illustration shows how a school's total NRT Accountability Index can be calculated from NRT results from multiple grades in the school.

Sample K-5 School Results on NRT Composite

	Grade 3		Grade 5		School Total
	SS	Index	SS	Index	
Student 1	170	20.9	223	108.7	
Student 2	179	58.4	193	20.5	
Student 3…	175	41.7	207	61.7	
Student n	168	12.6	207	61.7	
Number Included	63		58		121
Average Standard Score	183.5		210.7		
Index	**77.1**		**72.6**		**75.0**
Percentile Rank of Average Standard Score	46		44		

Using the Grade 3 formula, an index is calculated for each student in Grade 3. For example, Student 1 has a standard score of 170, which is equivalent to an index of 20.9. Student 1 in Grade 5, using the Grade 5 formula, has an index of 108.7. These index scores are averaged to yield a school index of 75.0. Indices can also be calculated for each grade taking the NRT.

Discussion

Averaging multiple grades and nominal weight of NRT

Note that if the scores were not averaged over multiple grade levels, the NRT might count more than once, which would violate the Commission's recommendation that the NRT count for 30% of the school as a whole, not for each grade.

Recommended method of calculating NRT Index

The TAC recommends calculating student-level index scores, then averaging since it avoids rounding error. Using student indices is also more operationally feasible, since the total index can be directly derived from the student file, and any changes at the student level file can be directly reflected in the school results rather than referring to an intermediate calculated result.

Standard scores averaged across multiple grades and percentile ranks

Note that no percentile rank is calculated for the standard score averaged across grade levels. This is because the standard score-to-percentile rank is grade-specific. That is, a scaled score of 180 is associated with a certain percentile rank in grade 3 and another percentile rank in grade 5.

3.3 Attendance Formula

Issue

How shall student attendance be included in accountability system?

Background

The Commission recommended including student attendance in the accountability system as one means of including more than achievement test scores. The Commission noted that Attendance was chosen because these data are collected in a consistent and accurate manner for every school, and because research has shown that attendance is related to appropriate classroom instruction and a supportive school environment. The Commission recommended that attendance count 10% of the total weight for Grades K-6 and 5% of the total weight for Grades 7-12. The Commission recommended the following 10-year and 20-year goals for attendance:

	10-year goal	20-year goal
Grades K-6	97%	98%
Grades 7 & 8	95%	96%
Grades 9 – 12	93%	94%

Recommendation

To generate an index for Attendance, the TAC made recommendations regarding the following:

1. Defining attendance goals
2. Defining an attendance index score

10-Year and 20-Year Attendance Goals

The TAC recommends the following attendance goals:

	10-year goal	20-year goal
Grades K-8	95%	98%
Grades 9-12	93%	96%

Attendance Formulas

Using the above goals, Attendance index scores would be calculated using the following formulas:

Grades K-8

$\text{Indicator}_{\text{ATTK-8}} = (16.667 * \text{ATT}) - 1483.3,$

Grades 9-12

$\text{Indicator}_{\text{ATT9-12}} = (16.667 * \text{ATT}) - 1450.0,$

where ATT is the attendance percentage, using the definition of attendance established by the Department of Education.

Lowest Attendance Index Score

The TAC recommends that zero be the lowest Attendance Index score for accountability calculations, although it would be possible to get a negative index score if the Attendance rate were low enough.

Example

Examples are given below of what the Attendance index scores for K-8 schools with different attendance rates would be, using the TAC's recommendation:

Example School - Grades K-8	Attendance Rate (%)	Attendance Index Score
School A	89	0.0
School B	91	33.3
School C	95	100.0
School D	98	150.0
School E	86	0.0 (-50.0)

A K-8 school with an attendance rate of 89% would receive an Attendance Index score of 0 (zero), as would a high school with an attendance rate of 87%. K-8 schools with attendance rates lower than 89% would receive an Attendance Index score of 0, although by the formula their scores would be a negative number. High schools with attendance rates lower than 87 would also be given Attendance Index scores of 0.

Discussion

Effective weight

The TAC was mostly concerned with creating a formula that would yield an effective weight for Attendance that would resemble the 10%/5% recommended by the Commission and adopted by BESE. Experience with other states' accountability systems has shown that typical methods of including attendance have resulted in a very low effective weight. That is, an increase in student attendance at a school would not be reflected in a meaningful increased accountability score unless the Attendance formula were constructed properly.

For example, if the attendance statistic used to create the School Performance Score were then attendance rate, a school that increased its attendance from 94% to 96% would get credit for just 2 points of gain on that indicator. Thus, when considered with the 10% weight given attendance for elementary schools, the School Performance Score would increase by only 0.2 points. With the TAC's proposed method, an increase of 2% in attendance rate would

translate into a 33.3-point gain in the Attendance Index. Multiplied by the weight of 10% for an elementary school, this would result in a change of 3.3 points on the School Performance Score.

The approach recommended by the TAC provides a realistic effective weight by establishing goals with more spread between the 10-year and 20-year goals, and by using the linear formula approach discussed previously with the NRT.

Goals

In making these recommendations, the TAC felt that the attendance goals are challenging for most schools in the state, yet it will be possible for most schools to show progress. The state means (and standard deviations) for 1997 Attendance were: Elementary: 95% (2); Middle: 93% (4), and High Schools: 91% (5). The TAC felt that the 20-year goal should not be higher than 98%, while the 10-year goal needed to be placed as low as feasible so that the spread between the 10-Year and 20-Year Goals be as large as feasible. This was accomplished by combining the K-6 and 7&8 goals into one goal, using the 7&8 10-year goal as the 10-year goal for both levels, and the 20-year goal for K-6 as the 20-year goal for both levels. For high school, the objective was accomplished by leaving the 10-year goal as originally recommended, and raising the 20-year goal to 96%.

Alternatives

The TAC discussed various methods of computing an index for attendance. Most alternatives did not involve rescaling. For example:

> Attendance Index = (school average attendance rate / state standard) x 100 x nominal weight

As discussed by the TAC, the main objection to this approach was that the effective weight for growth in attendance would be very small – most schools would see gains of much less than a point on the composite school index. For example, a high school that increased 5% in attendance (a very substantial improvement) would see an increase in its total School Performance Score of only one-quarter of one point under this alternative scheme.

The TAC also does not recommend increasing the nominal weight (e.g., make attendance count for 50% of the total index) in order to gain a larger effective weight. The TAC also considered other uses of attendance rates in accountability (e.g., a minimum attendance level could trigger an audit or be a precondition for rewards/level of assistance). These approaches, if used, should supplement, not replace, the method recommended by the Commission and adopted by BESE that makes attendance one of the four indicators.

3.4 Dropout Formula

Issue

How shall the Dropout indicator be included in the accountability system?

Background

The Commission recommended including dropout rate, along with student attendance, as a means of including more than only achievement test scores in the accountability system. The Commission noted that dropout rate was chosen because these data are collected in a consistent and accurate manner for every school, and because research has shown that declining dropout rates have been related to appropriate classroom instruction and a supportive school environment in other states. In addition, the meaning of other accountability indicators is questionable if dropout is not at least held constant. The Commission recommended that dropout count 5% of the total weight for Grades 7-12. The Commission recommended the following 10-year and 20-year goals for dropout rates:

	10-year goal	20-year goal
Grades 7 & 8	4%	3%
Grades 9 – 12	8%	6%

Recommendation

To generate a Dropout Index score, the TAC made recommendations regarding the following:

1. Defining dropout rate
2. Dropout Goals for various grade levels
3. Defining dropout formula

Definition of Dropout Rate

The TAC recommends using the definition of dropout and dropout rate established by the Department of Education, inverted so it is consistent with the other indicator scales (e.g., higher is better). This inversion would be calculated as:

> Non-Dropout Rate = 100 – Dropout Rate (expressed as a percentage).

Dropout Goals

The TAC concurs with the Commission on the 10-Year Goals for Dropout in Grades 7 & 8 and 9 – 12. However, for technical and empirical reasons the TAC recommends reducing the dropout rate even more for the 20-Year Goals. The TAC recommends the following goals for the Dropout Rate:

Dropout Goals Recommended by TAC (%)		
	10-Year Goal	20-Year Goal
Grades7 & 8	4	2
Grades 9-12	8	4

Dropout Formulas for Grades 7 & 8 and Grades 9 - 12

The formulas for calculating the Dropout Index are given below.

The Dropout Index for Grades 7 & 8 would be calculated:

$$Indicator_{DOGr7\text{-}8} = (25 * NDO) - 2300$$

The Dropout Index for Grades 9 – 12 wold be calculated:

$$Indicator_{DOGr9\text{-}12} = (12.5 * NDO) - 1050,$$

where NDO is the Non-Dropout Rate expressed as a percentage.

Terminology

Although it is based more directly on the non-dropout rate, the TAC recommends that this be referred to as the Dropout Index.

Low scores

The TAC recommends that zero be the lowest Dropout Index score for accountability calculations, although would be possible to get a negative score if a school's dropout rate were high enough.

Example

Examples are given below of the Dropout Index for high schools with different dropout rates, based on the TAC's recommendation:

Dropout Example

Grade 9 - 12	(%)	Index
School A	15.0	12.5
School B	10.0	75.0
School C	8.0	100.0
School D	6.0	125.0
School E	4.0	150.0
School F	17.0	0.0 (-12.5)

Discussion

Effective weight

The formulas recommended by the TAC are necessary to put the Dropout Index on the same scale as the other indicators, with the 10-year goal being 100 and the 20-year goal being 150, with the recommended goals. As was discussed with Attendance, other approaches to calculating a Dropout Index would lead to changes in

the dropout rate having very low impact on the total School Performance Score.

Different 20-year goals from Commission's recommendations

The TAC recommends the same 10-Year Goals as did the Commission. However, examining the actual dropout rates for schools in the state led the TAC to recommend higher 20-Year Goals both for grades 7 & 8 and for grades 9 –12. One reason was that spreading the 10- and 20-Year Goals further apart would yield Dropout Index formulas that would provide an effective weight for the Dropout Index that is closer to the nominal weight of 5% of the total composite index desired by the Commission. A second reason was that if the Commission's recommendations were used, about half of the schools in the state would have a zero score on the Dropout Index. That is, with 10-Year and 20-Year Goals of 92% and 94% for grades 9-12, any school with a dropout rate of 12% or greater would have an index score of zero. With the TAC's recommendation, schools with a dropout rate of 14% or greater would get a zero. In 1997 the state average dropout rate in high school was 13% (S.D. of 11). The average dropout rate in middle school was 6% (S.D. of 5).

Middle and High School Goals

The TAC recommends that different dropout goals be set for middle and high schools, based on empirical information about large differences in rates at these two levels.

4. COMBINING COMPONENTS: EXAMPLE OF CALCULATING A SCHOOL PERFORMANCE SCORE

Using the recommendations made thus far, it is possible to calculate a School Performance Score. We provide an example using an elementary school.

School A – Grade K-5

Calculating the CRT Index

The students in grade 4 take the CRT (criterion-referenced test). Assume the school has 50 students in grade 4, and their CRT scores were as follows:

Sample School – CRT Index example		
	Initial Achievement	
Performance Level	Number of Students	Index Points
Advanced	1	200
Proficient	9	1350
Basic	10	1000
Approaching Basic	15	750
Unsatisfactory	15	0
Total	50	3300
CRT Index	66.0 (3300/50 students)	

The school's CRT Index can be calculated using the values for Advanced (200 points), Proficient (150), etc., and the number of students at each performance level. For example, one Advanced student times 200 points equals 200 points; nine Proficient students times 150 points equals 1350 points.

The sample school's CRT Index is 66.0.

Calculating the NRT Index

The students in grades 3 and 5 take the NRT (norm-referenced test). The NRT Index is calculated by taking the standard score of each student, and using the formula for that grade level to convert it to an NRT Index score,

$$\text{Index}_{Gr3} = (4.167 * SS) - 687.5$$

$$\text{Index}_{Gr5} = (2.941 * SS) - 547.1$$

where SS = the average standard score.

These student index scores are averaged to give the school's NRT Index score.

Although not all of the student scores are shown in the example, assume that the sample school's average standard score is 183.5 for grade 3 and 210.7 for grade 5.

Sample K-5 School Results on NRT Composite

	Grade 3		Grade 5		School Total
	SS	Index	SS	Index	
Student 1	170	20.9	223	108.7	
Student 2	179	58.4	193	20.5	
Student 3…	175	41.7	207	61.7	
Student n	168	12.6	207	61.7	
Number Included	63		58		121
Average Standard Score	183.5		210.7		
Index	**77.1**		**72.6**		**75.0**
Percentile Rank of Average Standard Score	46		44		

Attendance

The Attendance Index can be calculated using the formula for attendance. The Attendance formula for K-8 schools is:

Attendance $\text{Index}_{K\text{-}8}$ = (16.667 * ATT) – 1483.3,
where ATT is the attendance percentage

Assume that the sample school has an attendance rate of 92%. Substituting 92% into the formula would yield an Attendance Index for the school of 50.0.

Dropout Rate

Dropout only applies to Grades 7-12, so a Dropout Index is not calculated for the sample elementary school.

School Performance Score

The School Performance Score for the sample school can be calculated by multiplying the index values for each indicator by the weight given to that indicator and adding the total scores. In the example, [(66.0 * 60%) + (75.0 * 30%) + (50.0 * 10%)] = 67.1.

Indicator	**Indicator Points**	**Weight**
CRT	**66.0**	60%
NRT	**75.0**	30%
Attendance	**50.0**	10%
Dropout	**NA**	0%
School Performance Score = 67.1		

The school in this example would have a School Performance Score of 67.1.

5. SCHOOL GROWTH TARGETS

5.1 Formula for School Growth Targets

Issue *How shall expected school progress be set?*

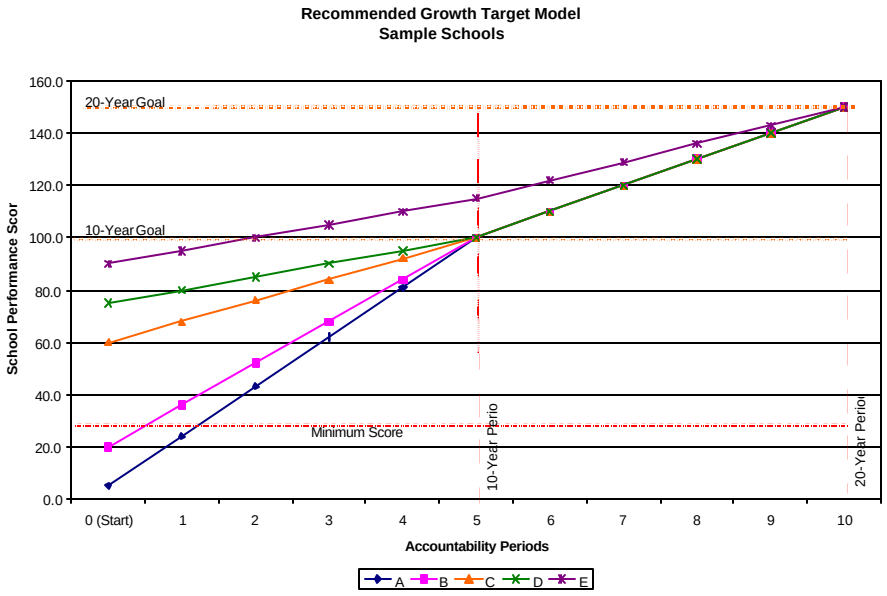
Background The Commission recommended that individual Growth Targets for each school be set every two years. The Commission recommended that the Growth Target for each school reflect how far the school had to go, and how much time it had. Specifically, the Commission recommended that a school's Growth Target be calculated by subtracting its School Performance Score from 100 (e.g., the 10-Year Goal) and dividing this difference by the number of two-year intervals remaining in the ten-year period.

Recommendation The TAC concurs with the recommendations of the Commission regarding the setting of Growth Targets for schools, with one modification. The TAC believes any change less than 5 points should be considered no real change.[1] The TAC recommends the following:

1. During the first 10 years, the formula is:
   Growth Target = (100 – School Performance Score) / number of two-year intervals remaining in the ten-year cycle, *or 5 points, whichever is greater*
2. During the second 10 years, the formula is:
   Growth Target = (150 – School Performance Score) / number of two-year intervals remaining in the ten-year cycle, or 5 points, whichever is greater
3. Schools above 150 are expected not to decline below 150.

Example Growth Targets are shown in the chart below for several hypothetical schools.

[1] The specific number of 5 points could be subjected to additional analysis. However, the principle is the same: since every measurement entails error, decisions about change should acknowledge the probability of error. The amount of acceptable error is largely a function of the decisions to be made, and the consequences.



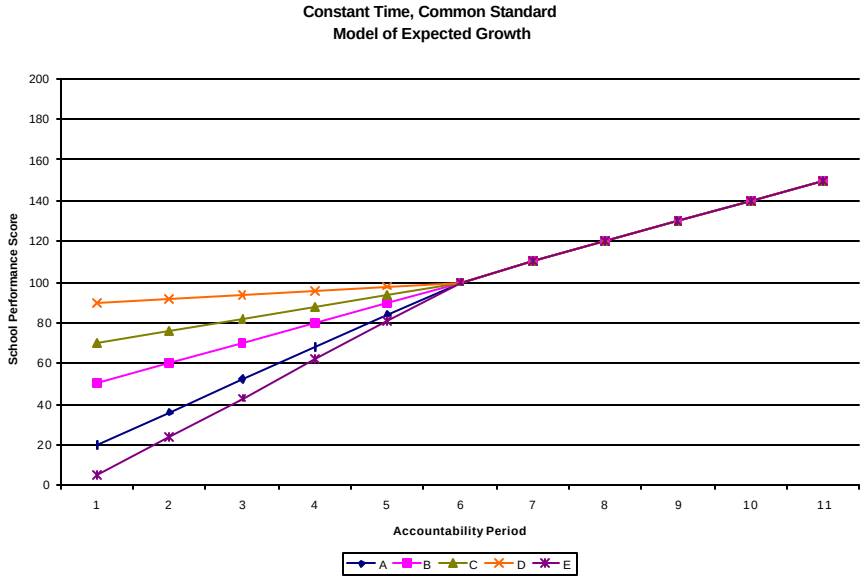
Discussion

Alternate models of Growth Targets

The TAC considered several alternate models of how growth expectations could be established. These alternate growth expectation models included two basic groups of models. The models differed in whether time or level of achievement would be held constant for all schools. The TAC's recommendation combines the two models for setting growth targets.

Constant time, common standard, different Growth Targets for different schools

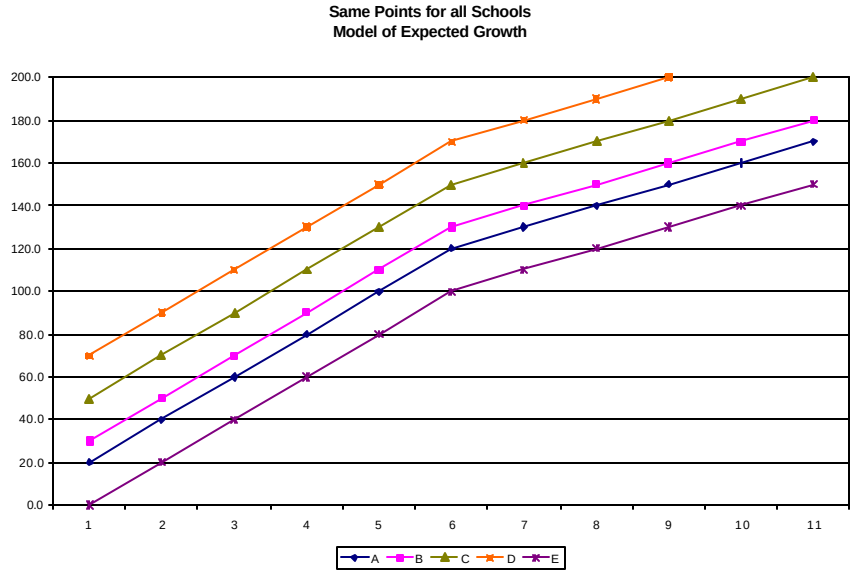
One model expects all schools to reach a common standard in a certain number of years. The Commission's recommendation was one variation of this model. All schools would have a set period of time (e.g., 10 years) to achieve a certain goal (e.g., a score of 100). An implication of this model is that schools will be required to progress at different rates from each other. A school that starts lower must improve more points each cycle to reach the same goal in the same amount of time as a school that starts higher. Another implication of this model is that it leads to convergence of expected performance over time; i.e., after a certain time period all schools will be expected to have the same score and same expected growth. This type of model reflects a commitment to common standards and equity in terms of requiring the same level of performance and rate of improvement eventually. The Commission's recommended way of setting growth targets so all schools reach 100 by the 10-year goal is an example of this type of model, and is shown in the chart below.



Different time, common standard, common Growth Targets for each school

A second model expects all schools to progress a certain amount each accountability period. If the expected growth is the same number of points for all schools, then all schools will eventually meet the standard (e.g., a score of 100), but will do so at different times. Those schools that started higher will reach the standard more quickly than those schools that started lower. One implication of this model is that schools will reach the goal at different times. Another implication is that relative differences between schools will be preserved; i.e., schools that start higher will be expected always to score higher than schools that start lower.

An extreme example of this model would be to require all schools to improve the same amount as the lowest-scoring school is required to improve. The lowest a school can score is 0 points. To be at 100 after 10 years (5 accountability cycles), it would be expected to gain 20 points each accountability cycle; thereafter the school must gain 10 points each accountability cycle to meet the 20-Year Goal of 150. The chart below depicts this growth model where all schools are expected to gain 20 points for the first 10 years, and 10 points for the second 10 years.



Specific alternate models

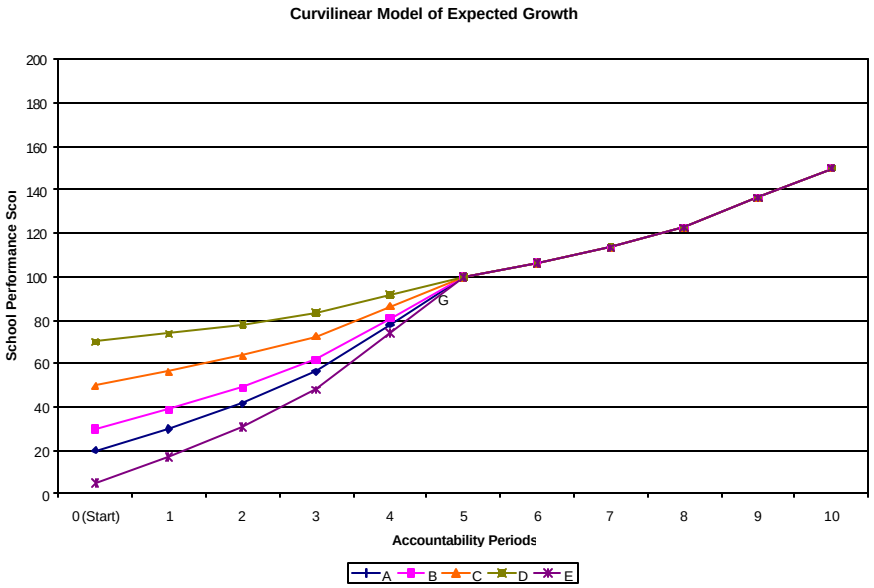
The TAC discussed several of the many variations possible within these two main models, including whether growth rates are linear or curvilinear, set once every ten years or recalculated each accountability period, and how to handle exceptionally low- or high-scoring schools.

Constant time variants

A. The Commission proposed a linear model of growth targets. It would be possible to set curvilinear growth expectations, either so faster progress is expected at the beginning, or slower. Arguments for curvilinear growth expectations reflect beliefs about probable school improvement. A model that requires relatively faster improvement at the beginning, for example, might reflect a belief that schools should be able to improve by doing simple and easily implemented things, such as working longer and using textbook funds appropriately. Subsequent growth might be expected to be slower as schools tackle difficult problems such as extending appropriate instructional strategies to help all students achieve higher standards.

   On the other hand, some researchers and educators might argue for a model that anticipates slower improvement at the beginning, since schools need to get used to the system and start the change process. Later progress might be expected to be relatively faster since schools would be able to learn from each other, and would have the necessary infrastructure (e.g., technology, school culture) in place.

   A chart of a curvilinear growth model is shown below.



B. Calculate improvement goals *once in the 10-year period* for each school, based on each school's individual initial score and 10-year state goal. This would provide the school a Growth Target for five cycles. Strengths of this model include promoting longer-term planning by schools—a school would know what its exact growth targets are for 2, 4, 6 years, etc., rather than having them recalculated after every two-year cycle. This approach also dampens changes in school Growth Label classifications due in part to non-linear school improvement. For example, a school might improve dramatically the first two years, and then somewhat more slowly the next two years, but over the four years still be on track to achieve the 10-year goal. This model would recognize the progress of the school over the four years, for example, rather than rewarding the school for the first two years and then possibly sanctioning the school for the next two years.

<u>Constant point variants</u>

A. A variant model would be that schools with scores lower than some "top bar" (less than 10-year goal, such as 80) follow the designated expected growth model, but schools above a top bar have less time to meet the 10-year goal. Strengths of this model include the perception of fairness that schools that start lower are not expected to progress faster than schools that start higher (in terms of number of points), simple, and that higher scoring schools can't coast for 10 years. One drawback of this model is the fact that schools with higher scores will be identified as

deficient if the goals are too high. Another is that this requirement institutionalizes expectations that existing differences in achievement between schools will be maintained in the future (all schools meet the minimum standard, but some more quickly).

Discussion of recommended model – variation for high-scoring schools

The TAC recommends the model proposed by the Commission with the change that the minimum growth expected be at least 5 points a cycle. The 5-point minimum is to ensure that the decision is reliable and reflects more than measurement error. (The specific number of points is subject to confirmation.) The TAC's recommendation takes the Commission's model and adds a constant Growth Target for higher-scoring schools. The Growth Targets calculated by both the TAC and the Commission recommendations would be identical for schools that start with a School Performance Score above 75. Lower scoring schools will have the full amount of time to reach the 10-year goal. However, schools that begin with higher scores—over 75, in the TAC's recommendation—also ought to have a reasonable challenge. This combination balances simplicity with technical soundness and realistic expected growth for all schools.

Recommendation for transition between 10-year and 20-Year Goals

The TAC discussed whether the Growth Targets should change at one point in time (at the transition between the first 10-years and the second 10-year period), or individually for each school as it reached the 10-year goal of 100 or the 20-year goal of 150. The TAC recommends that the Growth Targets be calculated in a consistent manner for 10 years for all schools. At the end of the first 10-year period, the Growth Targets for all schools would be reset, using the 20-year Goals and five remaining accountability cycles. This would allow the same rules to apply to all schools for each of the 10-year periods. This recommendation simplifies the system operationally, and allows schools to plan over long periods of time. These considerations apply primarily to high-scoring schools—defined as those that start off at 75 or higher. Since the TAC recommends that these schools be required to make a minimum of 5 points progress per cycle, schools that start off higher than 75 can be anticipated to end the 10-year period higher than 100. The maximum would be that a school that started at 100 would be expected to improve to 125 over the first 10 years. This is consistent with the Commission's recommendation to require that all schools progress, and with the TAC's recommendation that Growth Targets be set large enough to reduce the probability of misclassification.

5.2 Examples of Expected Growth for a School

Based on the recommends of the TAC and the Commission, the expected growth for a school will be that the school should increase as calculated by the formula:

Growth Target = (100 – School Performance Score) / number of accountability periods remaining in 10-year cycle,

with the additional provision that the Growth Target will be at least 10 points higher than the previous School Performance Score each cycle.

Two example schools are given in the following examples. The examples show how Growth Targets are calculated. Then the examples illustrate one possible way for the schools to achieve their Growth Targets. The examples use the same proportion of gain (points) for each indicator although a school could compensate for a lower score on one indicator with a higher score on another indicator. One example shows a lower scoring school that would begin below the "bottom bar." The other example shows a higher scoring school.

School A – K-5

Calculating a Growth Target

Assume School A, an elementary school, has an initial School Performance Score of 67.1 in 1999 at the beginning of the accountability period.

School A's Growth Target is calculated using the formula:

Growth Target = (100 – School Performance Score) / number of two-year intervals remaining in the ten-year cycle

Substituting the initial School Performance Score of 67.1 into the formula shows that School A needs to improve 6.6 points over the two-year period,

Growth Target = (100 – 67.1) / 5 = 6.6

School A's target School Performance Score for the next accountability cycle is 73.7.

Deciding on Areas of Improvement

Assume that the school decides to try to improve 6.6 points on each of the indicators: CRT, NRT, and Attendance. Because it is an elementary school, Dropout does not apply. The school can calculate the goals for each indicator in both the accountability metric and in the metric of the NRT standard scores, attendance percentages, (and dropout rate). This is shown in the table below.

Indicator	**Start (in index points)**	**Start (in indicator metric)**	**Goal (in index points)**	**Goal (in indicator metric)**
CRT	**66.0**	66.0	**72.6**	72.6
NRT	**75.0**	183.5/210.7	**81.6**	184.6/213.8
Attendance	**50.0**	92.0%	**56.6**	92.4%
Dropout	**NA**	NA	**NA**	NA
Total Weighted Index	**67.1**		**73.7**	

Using the formulas discussed previously, the school can calculate how to achieve its progress goals:

1. On the CRT, increasing from an index of 66.0 to 72.6, by changing the proportions of students achieving Advanced, Proficient, Basic, etc., as shown in the table below.
2. On the NRT, increasing the average standard score about 2 points. Because the school has multiple NRT grades, the school could improve from 75 to 81.6 on the NRT Index by increasing its average grade 3 standard score from 183.5 to 184.6 (or from the 46th percentile rank to the 48th percentile rank), and increasing its average grade 5 standard score from 210.7 to 213.8 (equivalent to increasing from the 44th to the 47th percentile rank).
3. In Attendance, increasing from 92.0% to 92.4%.

Meeting the CRT Goal

The following table illustrates how the example school could meet its improvement goal on the CRT test, i.e., by increasing the number of students at the Advanced level from 1 to 2, at the Proficient level from 9 to 10, Basic from 10 to 14, etc.

Sample School – CRT Index example				
	Initial Achievement		Goal Achievement	
Performance Level	Number of Students	Index Points	Number of Students	Index Points
Advanced	1	200	1	200
Proficient	9	1350	10	1500
Basic	10	1000	14	1400
Approaching Basic	15	750	15	750
Unsatisfactory	15	0	13	0
Total	50	3300	53	3850
CRT Index	**66.0** (3300/50 students)		**72.6** (3850/53 students)	

Note that this example, for convenience, uses the same proportions of improvement for each of the indicators—6.6 points. A school would likely examine its reports, consider its programs, resources, and students, and set specific goals not only for each indicator, but for subareas within the indicators, e.g., "increase mathematics performance by 8 points on the CRT, and science by 6 points." A school could even decide to set goals with specific

subpopulations, such as "improve the performance of minority boys 10 index points on the mathematics section of the CRT by reducing the proportion scoring Unsatisfactory and increasing the proportion scoring Basic." The school obviously would need to plan what it would do specifically to accomplish this increase in student performance over the next year or two. In this way, the accountability results can be directly linked to needs analysis and curriculum planning for school improvement.

School B – Grades 9 - 12

Calculating a Growth Target

Assume that School B has an initial School Performance Score of 10.

The Growth Target for the school is 18 additional points in the next accountability period. This can be calculated using the formula:

Growth Target = (100 – School Performance Score) / number of accountability periods

Growth Target = (100 – 10) / 5 = 90 / 5 = 18 points on the School Performance Score.

The school's target School Performance Score is 28.0.

Deciding on Areas of Improvement

Again, assume that the school has equal performance on each indicator, as indicated by an accountability index score of 10. The school decides to try to improve 18 points on each of the indicators: CRT, NRT, Attendance, and Dropout. The school's start and goals are shown in the table below.

Indicator	**Start (in index points)**	**Start (in indicator metric)**	**Goal (in index points)**	**Goal (in indicator metric)**
CRT	**10.0**	10.0	**28.0**	76.0
NRT	**10.0**	222.9	**28.0**	231.5
Attendance	**10.0**	87.6%	**28.0**	88.7%
Dropout	**10.0**	15.2%	**28.0**	13.8%
Total Weighted Index	**10.0**		**28.0**	

Using the formulas discussed previously, the school can calculate how to achieve its progress goals:

1. On the CRT, increasing from an index of 10.0 to 28.0, by changing the proportions of students achieving Advanced, Proficient, Basic, etc., as shown in the table below.
2. On the NRT, increasing the average standard score from 222.9 to 231.5 in Grade 9.
3. In Attendance, increasing from 87.6% to 88.7%.
4. In Dropout, decreasing the dropout rate from 15.2% to 13.8%.

Meeting the CRT Goal

The following table illustrates how the example school could meet its improvement goal on the CRT test, i.e., by increasing the number of students at the Advanced level to 2 from 1, at the Proficient level to 11 from 9, Basic to 13 from 10, etc.

Sample School – CRT Index example				
	Initial Achievement		Goal Achievement	
Performance Level	Number of Students	Index Points	Number of Students	Index Points
Advanced	0	0	0	0
Proficient	0	0	0	0
Basic	0	0	6	600
Approaching Basic	20	1000	44	2200
Unsatisfactory	80	0	50	0
Total	100	1000	100	2800
CRT Index	**10.0** (1000/100 students)		**28.0** (2800/100 students)	

If this school continues to make progress at this rate on all the indicators, it will meet the 10-Year Goal of a School Performance Score of 100.

Growth Targets of 5 Points

Only actual experience in Louisiana after the accountability system is in place will tell whether these Growth Targets can be achieved. The TAC has proposed that schools' Growth Targets be at least 5 points per cycle on the School Performance Score. This is equivalent to the school having to do the following every two years:

a) increase 10% of the students one performance level on the CRT; and
b) increase the average standard score on the NRT about 1.5 points, or about 3 percentile ranks (depending on the specific grade level); and
c) increase attendance $3/10^{th}$ of a percent (0.3%); and
d) decrease dropouts $4/10^{th}$ of a percent (0.4%) for high schools, or $2/10^{th}$ of a percent (0.2%) for middle schools.

Schools with initial School Performance Scores below 75 will need to make proportionally greater growth in order to achieve the 10-year goal.

6. GENERATING ACCOUNTABILITY JUDGMENTS

6.1 Growth Labels

Issue

How should accountability Growth Labels be assigned?

Background

The Commission proposed a four-tier approach to assigning school growth labels. These categories reflected whether a school exceeded its Growth Target by 10% or more (*Exemplary Academic Growth)*, exceeded its Growth Target by less than 10% (*Recognized Academic Growth)*, improved some but did not meet its Growth Target (*Minimal Academic Growth)*, or made no growth or declined (*School in Decline).*

Recommendation

The TAC endorses the Commission's recommendations regarding accountability Growth Labels, with one modification. The TAC recommends including the provision for growth by a minimum of 5 points per cycle. That is, the criteria (and *labels*) would be:

- exceeded its Growth Target by 5 points or more (*Exemplary Academic Growth)*,
- exceeded its Growth Target by less than 5 points (*Recognized Academic Growth)*,
- improved some but did not meet its Growth Target (*Minimal Academic Growth)*,
- made no growth or declined (*School in Decline).*

Discussion

Performance bands and measurement error

The TAC recognizes the importance of having criteria for separating schools in terms of acceptable and unacceptable growth. The names of the categories are a policy decision. However, the criteria essentially establish bands around the School Performance Score and the Growth Target. The TAC recommends that those bands be at least the standard error of measurement, to minimize misclassification due to measurement error. The TAC's recommendation of 5 points should exceed the standard error for the School Performance Score. This recommendation of 5 points is subject to verification upon availability of more complete data.

6.2 Accountability Consequences

Issue

How should accountability consequences (e.g., corrective actions) be determined?

Background

The Commission proposed a three-tier approach to dealing with schools that were above the Minimum Score and below the 10-year goal. The categories reflected whether a school obtained its Growth Target (exits Corrective Actions); obtains 25% of less of its Growth Target (moves to next level of Corrective Actions); obtains more than 25% of its Growth Target (Level I Corrective Actions). In addition, schools would enter more intensive levels of Corrective Actions if they failed to improve in specified ways. For example, a school in Level III Corrective Actions must obtain at least 40% of its Growth Target by the end of the first year of Level III, or have an approved Reconstitution Plan, or the school will lose state approval.

Recommendation

The TAC endorses the Commission's recommendations regarding determining accountability consequences (e.g., Corrective Actions) and criteria for exiting Corrective Actions. However, the TAC recommends including the provision that the minimum points used be 5 points per cycle for invoking corrective actions. The TAC's recommendations for criteria *(*and *actions*) be as follows if a school:

- meets or exceeds its Growth Target (*exits Corrective Actions*);
- increases less than its Growth Target but more than 5 points (*Level I Corrective Actions*);
- increases less than 5 points (*moves to next more severe level of Corrective Actions*).

Discussion

Accountability actions and measurement error

There is always some error in any test measurement. How much measurement error is acceptable largely depends upon the types of decisions to be made, and the nature of the consequences. The TAC recommends that measurement error certainly be taken into account for strong Corrective Actions. Other decisions may be such that more error is acceptable.

Monetary rewards as incentives

The TAC did not discuss the issue of whether rewards should be monetary, and if so, how much makes a difference under what considerations. This remains a policy issue, not a technical issue. And it is the opinion of the TAC that while there is a growing body

of research associated with monetary rewards in school accountability systems, it is a field of study still in development. Monetary rewards and other positive and negative incentives are very specific to the individual state, school, and even teacher.

Socio-economic differences and goals

The TAC concurs with the recommendation of the Commission that no distinction be made in terms of ultimate standards or expected growth for either students or schools on the basis of socio-economic status. The TAC notes that the recommended growth model does provide some additional time for schools that begin lower relative to other schools. However, this is on the basis of student achievement, not on the basis of student or school demographics. The TAC concurs with the Commission that common standards of achievement should be established for all schools, and additional resources provided on the basis of need.

Additional conditions

The Commission recommended that schools be required that "the performance of high poverty students increase" in order for the school to be eligible for rewards. BESE may wish to consider this and other conditions for a school to receive rewards or assistance. For example, some states have required that growth be across some spectrum of students in order for a school to receive rewards. That is, a school that met its Growth Target but "left behind" some subgroup would not be eligible for rewards. At least one state requires that all ethnic subgroups within a school meet the progress goals. Another state requires that the proportion of students in the lowest CRT category (e.g., Unsatisfactory) be decreased by 10%. These additional conditions may powerfully communicate values. Whether they make a difference in actual school behavior and accountability decisions depends largely on the conditions in the state's schools and the specific formulation of the goals.

6.3 Setting a Minimum Score ("Bottom Bar")

Issue

What minimum school performance score should be set for accountability purposes?

Background

The Commission recommended that BESE set a minimum level of performance for all schools called the *Minimum Score*. Schools with a School Performance Score below the Minimum Score will be identified as Academically Unacceptable Schools. These schools

will immediately be eligible for Level I Corrective Actions. In accepting the recommendation of the Commission, BESE indicated that it desires that the Minimum Score reflect pilot data, and that the Minimum Score be "fixed" and not be raised over time.

Recommendation

Use of a "Bottom Bar"

The TAC concurs with the Commission that a "bottom bar," or minimum accountability threshold, should be established. The purpose of a bottom bar is to set a minimum standard. Schools below the minimum standard would receive additional assistance.

The TAC recommends a bottom bar that would serve to identify schools that need dramatic assistance from the district, state, and community. The bottom bar should signal dire need. It is an absolute performance standard, not a relative standard. That is, it is not the case that the "bottom 10 percent" of the schools would always be identified, no matter how high their scores were.

Standard for "Bottom Bar"

The TAC recommends that the bottom bar be set at 30, on the School Performance Score which ranges from 0 to 100 as the 10-year goal. A school could be below 30 on individual indicators (e.g., Attendance), as long as the total School Performance Score was above 30.

Corrective Actions

The TAC supports the Corrective Actions recommended by the Commission regarding schools below the Bottom Bar: schools would be immediately eligible for Level I Corrective Actions, and would need to improve at least 40% of their Growth Target or 5 points, whichever is less, by the end of the first year or face additional sanctions.

Discussion

In making this recommendation, the TAC considered several issues. A School Performance Score below 30 indicates, in the opinion of the TAC members, a low school performance that warrants significant external assistance if the school is eventually to meet the school's long-term goals established by the state.

What does a school with a School Performance Index of 30 look like in terms of the indicators? An illustrative set of indicator scores is given below for an elementary school. On the CRT, 50% of the students would be at Unsatisfactory, 40% would be at Approaching Basic, and 10% Basic. No students would be at the Proficient or Advanced levels. On the NRT, the average standard score (and associated percentile rank) would be 172.2 (24^{th} PR) for grade 3,

196.2 (25th PR) for grade 5. The attendance rate would be 90.8%.

Note that no school could meet the expected growth targets and be under the bottom bar more than two cycles. That is, any school with an initial School Performance Score of 1 (one) would be expected to increase 19.8 points [(100 – 1) / 5] to 20.8 the next cycle, and have a School Performance Score of 40.6 the next cycle. A school with an initial score of 12.6 or higher would be expected to be above the Minimum Score of 30.0 in one cycle.

In addition, based on analyses of available data, the number of schools likely to fall below this minimum standard is both credible and yet manageable for the external resources likely to be made available. Using the available pilot data from the Math CRT and full ITBS data for grades 4 and 8, approximately 15% of the schools would initially be below a bottom bar of 30. The TAC anticipates that School Performance Scores would be higher if the CRT English were included and after the accountability system officially begins.

6.4 Setting a "Top Bar"

Issue

Should provisions be made for high-scoring schools?

Background

The notion of a top bar was brought up in the context of setting a threshold that would recognize high scoring schools and provide them appropriate accountability incentives. The Commission recommended that schools scoring above the 10-Year Goal before 10 years be designated as *Academically Distinguished Schools*. They would not be subject to Corrective Actions during the first 10 years so long as their School Performance Scores were above 100, but they would have to show growth to receive rewards.

A top bar was also discussed by the TAC in the context of setting growth expectations for higher scoring schools (e.g., expected growth would be 5 points for all schools initially scoring over 75 on the School Performance Score) that would be simple and not allow these schools to "coast."

Recommendation

Use of a "Top Bar"

The TAC supports the Commission's recommendation regarding providing appropriate incentives for high-achieving schools.

Setting a "Top Bar"

The TAC concurs with the recommendation that schools above 100 in the first 10 years not be subject to Corrective Actions. The TAC recommends that all schools be eligible for rewards upon meeting the growth requirements discussed previously.

Discussion

Most issues involved in a top bar are intended to address are handled in the recommendations made by the TAC regarding growth targets. This has the advantage of keeping the system simple, and focusing on progress rather than different rewards or sanctions.

6.5 Accountability Cycles

Issue

What should be included in an accountability cycle?

Background

The Commission recommended that each school be expected to grow and that its progress be assessed every two years. While the Commission was not explicit in its report, there was a clear implication that only the testing done in the odd years (e.g., 1999, 2001) would be included in the accountability system. However, students will be tested annually. Therefore, several other possibilities exist for accountability cycles.

Recommendation

The TAC recommends that an accountability cycle incorporate the following:

1) Test (CRT and NRT) every year in every content area (English/Reading, mathematics, science, social studies) in the designated grades.
2) Make accountability decisions every two years, on the odd years.
3) Make accountability decisions on the basis of two years' "baseline" and two years' "growth" data (except the first cycle).
4) Average the two years of baseline data to yield an Initial or baseline score, and the two years of growth data to yield a growth School Performance Score.
5) Use the maximum available data to make accountability decisions during the initial start up of the system when tests are being phased in.

Discussion

Using results from every year in accountability

These recommendations differ somewhat from what the Commission proposed and what has been publicly discussed as a likely design by Department staff. These discussions proposed using a single year's results every other year for accountability purposes, and using the alternating years' results for informational purposes only (for example, comparing 1999 to 2001 for accountability purposes, and reporting 2000 for informational purposes only).

Averaging produces more reliable scores

The TAC recommends using all the data from each year for two main reasons.

a) Averaging data from multiple years (i.e., multiple grades of students) will reduce measurement error due to student cohort differences, resulting in more reliable school scores. It may be anticipated that differences between student cohorts year-to-year will be the largest source of measurement error in the system, based on other states' experiences.

Averaging avoids identifying certain students with certain types of testing

b) Averaging multiple years will avoid cohorts of students taking predominately one type of test for accountability purposes (i.e., NRT or CRT). If only every other year's data were used for accountability purposes, it would lead to the pattern illustrated in the following table. The table shows that the student cohort that starts in Grade 3 in 1999 would have only their NRT results count for accountability through 2005 when they are in Grade 9. In contrast, the student class that starts in Grade 4 in 1999 would have only CRT results count in three of their four accountability cycles until they graduate from high school. To the extent that there are method differences between NRT and CRT, it would be desirable to average these across different student cohorts. Also, if there are instructional effects associated with preparing students for the CRT or NRT, these should also be spread across cohorts of students. The TAC's recommendation of making all the years count by averaging their results should lessen these undesirable effects.

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007
Cohort – Grade	3	4	5	6	7	8	9	10	11
Test	NRT	**CRT**	NRT	NRT	NRT	**CRT**	NRT	**CRT**	**CRT**
Cohort – Grade	4	5	6	7	8	9	10	11	12
Test	**CRT**	NRT	NRT	NRT	**CRT**	NRT	**CRT**	**CRT**	none

Shaded = accountability year under Commission's proposal
Non-shaded = non-accountability year under Commission's proposal

7. DEALING WITH SPECIAL CASES

7.1 Start-up of Accountability System

Issue

How should incomplete testing data be handled during the start-up of the accountability system?

Background

The TAC notes that there will be issues associated with the start-up of a new accountability system in terms of what testing results will be available for which grade. Notably, since the content area tests are being phased in for both elementary and secondary schools, a plan must be put in place of how to construct accountability indices for the initial two accountability periods.

For K-8, accountability testing will begin in 1999 and accountability results will be declared after the tests in spring 2001. English and mathematics in the CRT and the full NRT battery (Reading/ Language Arts, mathematics, science, and social studies) will be tested beginning 1999. CRT testing for science and social studies will start in 2000.

For grades 9-12, CRT accountability testing will begin in 2001 in English and mathematics in grade 10. Science and social studies will be tested in Grade 11, beginning spring 2002. The NRT full battery will be administered in grade 9 beginning 1999. According to the Commission, prior to 2003, high school reports will be for informational purposes only and will not include accountability-like calculations or declarations.

Recommendation

K-8

The TAC recommends that for K-8:

1) Accountability results for K-8 schools be declared in 2001, based on English and mathematics only from the CRT, and the NRT battery results. The results from 1999 will serve as the baseline (pretest), and the averaged results from 2000 and 2001 will serve as the growth index (posttest). Only English and math will be included in the CRT in all three years, since they are the only content areas from the CRT that will have been included in the baseline of 1999.
2) Accountability results for 2003 for K-8 schools will include

English, mathematics, science, and social studies from the CRT, and the full battery for the NRT. The CRT baseline will consist of the average of two years each in English and mathematics (2000, 2001), and one year of science and social studies (2000). The CRT growth index will consist of two years' results in each of the four content areas (2002, 2003).

Grades 9-12

The TAC recommends for Grades 9-12:

1) Accountability results for 9-12 schools will be declared in 2003, using available results from Grades 9, 10, and 11. The 2003 accountability results will be based on:
   - One year's data in English and Math (spring 2001), compared with two years' data (spring 2002, 2003);
   - Two years' data on the NRT (spring 2000, 2001) compared with two years' data (spring 2002, 2003).
2) Accountability results in 2005 will be based on:
   - Two years' of data in all CRT subject areas (Spring 2002, 2003) compared with two years' of data (spring 2004, 2005)
   - Two years' of data in the NRT battery (spring 2002, 2003) compared with two years' of data (spring 2004, 2005)

Discussion

Note that the usual pattern for accountability recommended by the TAC consists of pairs of years: an even-odd pair compared to a subsequent even-odd pair. Because of the phasing in of the tests, incomplete sets of test results are available in the initial years of the system. One alternative is exclude the content areas for which there are only partial results until the next accountability cycle. In the TAC's experience, most policy makers find such a lag unacceptable. While not as stable, using one year's results is usually technically acceptable, given the understanding that at the beginning of the accountability system there are many unknowns.

Generally, when tests that are introduced in an odd year, only one year's data are available for a baseline. Tests that are introduced in an even year must wait until the next full cycle to be included in accountability system calculations, although they may be reported.

7.2 Special Cases of Student Accountability

Issue *What students should be included in the accountability system?*

Background The state will need to determine what students should be included in the accountability system. The state will also need to decide what to do with a student who should be included but did not take the test.

Special education students The Commission recommended that the performance of special education students be part of a district accountability system, where performance of special education students will be used in calculating scores for the district rather than an individual school. The Commission recommended that all 504 students who are classified as special education students with disabilities will be included in the district's accountability system. All other students (e.g., Attention Deficit Disordered, etc.) who have been classified as 504 students will be included in the calculation of the school's School Performance Score. The Department is working with advisory groups to refine existing concepts for the special education accountability system. The Commission recommended that these issues include how to establish a 10-Year Goal for special education students, appropriate Growth Targets, and a process to integrate the special education goals into a district system, to begin by 2001.

Special purpose school students There are many types of special purpose schools (e.g., alternative schools, schools in correctional facilities, adult education facilities), and many reasons a student may be assigned to a special purpose school. The Commission recommended that these special purpose schools have an alternative accountability system to be implemented no later than 2001. The Department is working with advisory groups on the issue of accountability for special purpose schools.

Mobile students The Commission recommended that a student be attributed to the school for accountability purposes if the student was enrolled in the school on October 1 and on the day the tests were administered. If the student moved into the school after October 1, then the student's performance would be included in the district accountability system, but would not be included in the School Performance Score.

Students who do not take the test The Commission recognized that it is likely that it will be necessary to establish a "score" for students for whom a school is accountable but who do not take the CRT and/or NRT. The Commission recommended that a score of "0" be entered for any eligible student

who does not take the test, including those who were absent.

Recommendation

All students, with few exceptions included in school accountability

The TAC endorses the Commission's stance that all students should be included in the assessment and accountability program, with very few exceptions. The Department of Education should define the permissible exceptions, such as for medical reasons. The TAC recommends that all students, to extent possible, be included in school accountability.

Special education and 504 students

The TAC recommends that special education and 504 students be included in the school accountability system. The schools that make decisions about the provision of instructional services should be accountable. In most cases this will be the school where the student is enrolled. In cases where the decision is made, with the school's consent, that the student is better served elsewhere, the student's score should be attributed to the sending school. This recommendation does not restrict the construction of other accountability mechanisms for districts or special purpose schools.

Students who attend special purpose schools

The TAC recommends that schools that make or approve decisions to send students to special purpose schools be held accountable for those students. The student scores should be attributed to the sending schools. Where the student's enrollment in a special purpose school is a function of legal proceedings or some other circumstance beyond the school's control (e.g., the student is committed to a correctional facility through a court order), then the TAC recommends that the school not be held accountable for that student. This recommendation does not restrict the construction of accountability mechanisms for districts or special purpose schools.

Mobile students

The TAC recommends that all students who are in the school at the time the tests are administered be attributed to the school for accountability purposes. Schools with high proportions of mobile students—that is, students who were not enrolled at that school as of October 1 of the school year—would be able to appeal.

Assign a score for students who miss the test

The TAC makes the following recommendations, elaborating on the Commission's recommendation that students who do not take the test be assigned a score of "0."

Missing CRT Scores

- students who do not take the CRT test will be assigned the lowest score, Unsatisfactory, unless validly excluded. If a student does not participate in a section of the CRT (e.g., mathematics), the student will be assigned a score of Unsatisfactory for that content area of the test. The

Unsatisfactory level is given a score of zero points when calculating a CRT Index, as has been discussed.

Missing NRT Scores

- students who do not take the NRT test will be assigned an NRT Index score of 0 (zero), unless validly excluded.

Discussion

Every student, school accountability

The TAC endorses the Commission's thoughtful attention to the issues of assigning scores for missing students, which reflects the beliefs that schools and districts should be accountable for every student, and that all students can and must learn at significantly higher levels. This ensures that accountability supports instruction, and lessens the probability that accountability will act as a perverse incentive to exclude or not serve certain students. There are three main reasons the TAC recommends inclusion at the school level of special education students and students enrolled in most special purpose schools, contrary to the recommendation of the Commission that these students be held accountable only at the district level or in their special purpose schools.

Reasons for including at school level

Schools have instructional control most of the time

1. Accountability is a policy tool to encourage improved educational services for students. Accountability should be at the level where educational services are provided and decisions about educational resources are made. For the very large majority of special education students, day-to-day instructional decisions are made at the school level. While districts may have a special role in being accountable for services to these children, that should be in addition to, not instead of school accountability.

Schools are responsible for decisions about services

2. Accountability should provide incentives to be responsible for the student. A school that decides that a student should be enrolled in a special purpose school or other program should do so on the basis of what is best for the student, not on the basis of excluding the student to maximize the school's test scores. Holding the regular school accountable reduces the incentive to "push" students where they will not count or will count very little. (The special education population is small enough in most districts that their impact on district scores would probably be negligible.)

Federal requirements on inclusion and reporting

3. The inclusion of all students at the school level also is in keeping with federal law, as recently expressed in IDEA (Individuals with Disabilities Education Act) and Title 1. Federal law requires that reporting for students with disabilities be done "*at the same level* and with the same frequency" as reporting for the general population. Since Louisiana's main assessment and accountability reporting mechanism is at the school level, it might be argued that inclusion of special education students in the school accountability system is most in keeping with federal

requirements.

The TAC recognizes that this is a thorny issue. Again, the TAC encourages the Commission, BESE, and the Department to move forward with development of accountability plans for districts and special purpose schools. Such plans can complement a strong school accountability scheme that is an inclusive and comprehensive as possible.

Mobile students

The TAC recommends including as many students in accountability at the school level as possible. The numbers of transient students in the state was reported to the TAC as being relatively low. Another consideration was minimizing the potential abuse of excluding students from the test by claiming they had not been enrolled in time. In addition, the TAC was told that it was practically infeasible to produce the accountability reports in a timely manner if it were required to match the students at two points in time.

Missing CRT scores

The TAC recommends assigning an Unsatisfactory score to students who do not take the CRT because as the lowest score it is equivalent to "zero points," as recommended by the Commission. In addition, there are empirical and logical reasons that support this policy. Preliminary results indicate that a large proportion of students will be at the Unsatisfactory performance level, so assigning a score of Unsatisfactory for missing students will not be excessively punitive to the schools. More importantly, since the worst any student could do if tested is to score at the Unsatisfactory level, it acts as an incentive for school not to intentionally exclude students from testing.

Missing NRT scores

The TAC's recommendation that students who miss the NRT be assigned an NRT Index score of zero follows the Commission's recommendation, and has a similar rationale as the CRT presented above. As mentioned in the NRT section, a level of performance on the NRT equivalent to all guessing is about equal to a standard score that would yield an index of zero. Thus, there is little incentive in terms of seeking higher scores for a school to push a student out of taking the NRT: it would be better for the school for the student to take the test than take a zero, if the student could do better than random guessing.

7.3 Special Cases of School Accountability

Issue

How shall accountability be implemented for schools with special circumstances, such as non-standard grade-level configurations?

How shall accountability be implemented for schools that have significant changes in population demographics?

Background

Missing test data due to grade configurations

There will be schools in the state that will lack some testing data because of their grade configuration. For example, some schools will not have any CRT and/or NRT test data (e.g., K, K-1, K-3 schools). The Commission recommended that schools that serve only students in K-2 be paired with schools in the district that receive its students. Schools with only K-2 students would then be judged based upon the performance of paired schools. The Commission recommended that local school boards determine how schools be paired.

More grades tested due to grade configurations

In addition, there will be schools that have more grades tested than others (e.g., K-8 schools). The Commission did not address this issue.

Significant changes in school situations

During an accountability period a school may have significant changes in the demographics of the student body it serves. This could be because the school is newly formed, because of revision in the sending area, or because of a change in the community (e.g., change in a major employer that leads to changes in school enrollment).

A decision must be made how to handle test results for these schools.

Recommendation

The TAC supports the recommendation of the Commission that schools without test data be paired with other schools in the district and held accountable through the results of that school.

The TAC recommends the following for special cases of school accountability.

School units for accountability should have results from at

Where a school building does not include the standard accountability grades:

1. Pair a non-standard school with at least one other school so that

least one grade of CRT and one grade of NRT testing

across the pair, at least one grade of NRT and one grade of CRT data are included. For attendance and dropout data, use the data unique to each school.

2. Have the local school district make the final decision of how to pair schools. There should be a consistent district policy for such decisions, such as feed-forward or feed-backward.

Newly created schools

The TAC recommends that schools be included in the accountability system as long and as rapidly as possible. The Department of Education should establish a policy for dealing with schools that are newly established during an accountability period, or that are extensively changed, such as due to reconfigured attendance areas. The policy may establish that at least one year of pre-test and one year of post-test data be available for accountability purposes.

Changes in populations served, school reconfiguration, mobile populations

While the TAC recommends that the accountability system be as inclusive of schools as possible, it recognizes that schools may have unusual changes. The Department of Education should establish a policy for dealing with schools that have changes beyond their control, such as large changes in school demographics. This could be through individual negotiations with the Department, appeals, or specific policies for identifiable situations. The TAC recommends that a change of more than 20% of a school's population be the basis for consideration of an appeal.

Discussion

There will always be the need to deal with special cases of school accountability. However, the principle should be clearly and firmly established that the accountability system applies to all schools, as much as possible. The purpose of a school accountability system is primarily to benefit the community and students through increased academic performance of students over time. Therefore, exceptions for schools should be considered in terms of the long-term effect they will have on individual students and on the credibility and soundness of the system as a whole over time.

The primary reason the TAC recommends that school accountability be based on at least one grade of NRT and one grade of CRT testing is that the results will be more valid and reliable. Both types of tests are needed to adequately address measurement of students' and schools' progress towards the standards and goals established by the state. In addition, inclusion of both types of tests is in keeping with the recommendation of the Commission and the action of BESE to include the CRT and NRT indicators at 60% and 30% weights, respectively.

NEXT STEPS

The design of a sound school accountability system is a major accomplishment. It is a crucial step towards fostering significant improvement of student academic achievement—but only one step. An operational accountability system involves more than administering tests and collecting data. It is also critical to ensure that the data are valid and that the results can be communicated effectively to the public and used powerfully by schools to improve student learning. As such, we encourage the Department of Education and the Board of Elementary and Secondary Education to vigorously pursue at least the following activities:

1. Conduct additional studies as needed to determine, as much as possible, the appropriate values for minimum Growth Targets (Is 5 points the most appropriate amount?), percent population change for reconfigured schools and mobile students, and impact of special education and special-purpose school students;
2. Design reports and develop the data systems and procedures necessary to produce the reports in a timely manner;
3. Develop the necessary administrative policies, including statements of appropriate testing practices, consequences for violations of test security, appeals procedures, and other administrative guidelines;
4. Plan for effective communication to the public, policymakers, and other key constituencies of the standards, available assessment information, and accountability results;
5. Design and disseminate materials to help educators use the assessment and accountability reports effectively for school improvement and focused teaching;
6. Train educators about the accountability system and how to use the reports and other information effectively;
7. Design district accountability and alternate school accountability programs;
8. Plan the necessary data collection and analyses to establish the validity of the assessment and accountability system.

APPENDIX: SUMMARY

The recommendations of the TAC regarding the assessment and accountability design are summarized in this appendix.

Goals Recommended by the TAC

School Goals and Goal Scores Recommended by TAC				
Student Indicators	**10-Year Goal**	**10-Year Goal Score (100)**	**20-Year Goal**	**20-Year Goal Score (150)**
CRT	Average student performance at BASIC proficiency level	100	Average student performance at PROFICIENT level	150
NRT	Average Composite standard score corresponding to the 55th percentile rank in the grade level	100	Average Composite standard score corresponding to the 75th percentile rank in the grade level	150
Attendance	95% (Grades K-8) 93% (Grades 9-12)	100	98% (Grades K-8) 96% (Grades 9-12)	150
(Non) Drop-out	96% (Grades 7 & 8) 92% (Grades 9 – 12)	100	98% (Grades 7 & 8) 96% (Grades 9 – 12)	150

Indicators and Weights

Indicators	**Grades Administered**	**Weights**
Criterion-Referenced Tests (CRT)	Grades 4, 8, 10, 11	60%
Norm-Referenced Tests (NRT)	Grades 3, 5, 6, 7, 9	30%
Student Attendance	Grades K-6 Grades 7-12	10% 5%
Dropout Rates	Grades 7-12	5%

CRT Performance Levels and Values

Performance Level	**Values**
Advanced	200
Proficient	150
Basic	100
Approaching Basic	50
Unsatisfactory	0

Calculating a CRT Index for a School

1. Calculate the total number of points by multiplying the number of students at each performance level times the points for those respective performance levels, for all content areas.
2. Divide by the total number of students tested times the number of content area tests.

Sample School		English/Lang. Arts		Mathematics		
Performance Level	Points/ Student	Number of Students	Total Points	Number of Students	Total Points	Total
Advanced	200	3	600	1	200	
Proficient	150	6	900	6	900	
Basic	100	9	900	9	900	
Approaching Basic	50	17	850	12	600	
Unsatisfactory	0	8	0	15	0	
Total		43	3250	43	2600	5850
Index			**75.6**		**60.5**	**68.0**

NRT Goals and Equivalent Standard Scores

ITBS Composite Standard Scores Equivalent to Louisiana's 10- and 20-Year Goals, by Grade Level[1]						
		Grade				
Goals	Percentile Rank	3	5	6	7	9
10-year Goal	55th	189	220	232	245	266
20-year Goal	75th	201	237	253	268	290

[1] Source of percentile rank-to-standard score conversions: *Iowa Tests of Basic Skills, Norms and Score Conversions, Form M* (1996) and *Iowa Tests of Educational Development, Norms and Score Conversions, with Technical Information, Form M* (1996), Chicago, IL: Riverside Publishing Co.

NRT Formulas Relating Standard Scores to NRT Index

When the 10-year and 20-year goals are the 55th and 75th percentile ranks, respectively:

Grade 3: $Index_{Gr3} = (4.167 * SS) - 687.5$ where SS = the average standard score
$SS = (Index_{Gr3} + 687.5) / 4.167$

Grade 5: $Index_{Gr5} = (2.941 * SS) - 547.1$
$SS = (Index_{Gr5} + 547.1) / 2.941$

Grade 6: $Index_{Gr6} = (2.381 * SS) - 452.4$
$SS = (Index_{Gr6} + 452.4) / 2.381$

Grade 7: $Index_{Gr7} = (2.174 * SS) – 432.6$
$SS = (Index_{Gr7} + 432.6) / 2.174$

Grade 9: $Index_{Gr9} = (2.083 * SS) – 454.2$
$SS = (Index_{Gr9} + 454.2) / 2.083$

Lowest NRT Index

The lowest NRT Index score for a school will be zero.
The lowest NRT Index score for a student will be the index score calculated using the formula. The standard scores and percentile ranks corresponding to an index score of zero are shown below, using the 55th and 75th percentile ranks as the 10- and 20-year goals.

NRT Index score of zero, using 55th and 75th percentile ranks as goals					
	Grade 3	Grade 5	Grade 6	Grade 7	Grade 9
Standard score	165	186	190	199	218
Percentile rank	14th	14th	9th	10th	14th

Highest NRT Index

There will be no constraint imposed on the highest NRT Index scores for a student or school calculated using the formula. The ITBS Composite standard scores and percentile ranks corresponding to an index score of 200 are shown below, using the 55th and 75th percentile ranks as the 10- and 20-year goals.

NRT Index score of 200, using 55th and 75th percentile ranks as goals					
	Grade 3	Grade 5	Grade 6	Grade 7	Grade 9
Standard score	213	254	274	291	314
Percentile rank	90th	89th	91st	92nd	92nd

The ITBS Composite standard scores and NRT index scores are shown below for the maximum and the 99th percentile rank, using the 55th and 75th percentile ranks as the 10- and 20-year goals.

	Grade 3	Grade 5	Grade 6	Grade 7	Grade 9
Maximum					
Standard Score	269	309	329	349	369
NRT Index score	433.3	361.8	330.9	326.1	314.5
99th PR					
Standard Score	228	274	293	308	332
NRT Index score	262.5	279.2	245.2	237.0	237.5

Calculating a School's NRT Index for Multiple Grades

1. Calculate the index for each student, using the grade-appropriate formula relating standard score to NRT Index.
2. Average the student indices across all students in all grades in the school.

Attendance Goals

	10-year goal	20-year goal
Grades K-8	95%	98%
Grades 9-12	93%	96%

Attendance Index Formulas

Grades K-8

$\text{Indicator}_{ATTK\text{-}8} = (16.667 * ATT) - 1483.3,$

Grades 9-12

$\text{Indicator}_{ATT9\text{-}12} = (16.667 * ATT) - 1450.0,$

where ATT is the attendance percentage, using the definition of attendance established by the Department of Education.

Lowest Attendance Index Score

The TAC recommends that zero be the lowest Attendance Index score for accountability calculations.

Dropout Goals

Dropout Goals Recommended by TAC (%)		
	10-Year Goal	20-Year Goal
Grades7 & 8	4	2
Grades 9-12	8	4

Dropout Index Formulas

Grades 7 & 8 $Indicator_{DOGr7\text{-}8} = (25 * NDO) - 2300$
$NDO = (Indicator_{DOGr7\text{-}8} + 2300) / 25$

Grades 9 – 12 $Indicator_{DOGr9\text{-}12} = (12.5 * NDO) - 1050$,
$NDO = (Indicator_{DOGr9\text{-}12} + 1050) / 12.5$
where NDO is the Non-Dropout Rate expressed as a percentage.

Lowest Dropout Index Score

The TAC recommends that zero be the lowest Dropout Index score for accountability calculations.

Calculating a School Performance Score

The School Performance Score for the sample school is calculated by multiplying the index values for each indicator by the weight given to that indicator and adding the total scores. In the example, [(66.0 * 60%) + (75.0 * 30%) + (50.0 * 10%)] = 67.1.

Indicator	Indicator Points	Weight
CRT	**66.0**	60%
NRT	**75.0**	30%
Attendance	**50.0**	10%
Dropout	**NA**	0%
School Performance Score = 67.1		

Recommended Growth Targets

1. During the first 10 years, the formula is:
   Growth Target = (100 – School Performance Score) / number of two-year intervals remaining in the ten-year cycle, *or 5 points, whichever is greater*
2. During the second 10 years, the formula is:
   Growth Target = (150 – School Performance Score) / number of two-year intervals remaining in the ten-year cycle, or 5 points, whichever is greater
3. Schools above 150 are expected not to decline below 150.

Growth Labels

The TAC recommends these criteria (and *Growth Labels*) for a school:

- exceeded its Growth Target by 5 points or more (*Exemplary Academic Growth)*,
- exceeded its Growth Target by less than 5 points (*Recognized Academic Growth)*,
- improved some but did not meet its Growth Target (*Minimal Academic Growth)*,
- made no growth or declined (*School in Decline)*.

Accountability Actions

The TAC recommends the criteria *(*and accountability *Corrective Actions*) be as follows if a school:

- meets or exceeds its Growth Target (*exits Corrective Actions*);
- increases less than its Growth Target but more than 5 points (*Level I Corrective Actions*);
- increases less than 5 points (*moves to next more severe level of Corrective Actions*).

Minimum Score

The TAC recommends that the bottom bar be set as a School Performance Score of 30. A school could be below 30 on individual indicators (e.g., Attendance), as long as the total School Performance Score was above 30.

Top Bar

The TAC recommends:

- schools with a School Performance Score of at least 150 would not be subject to any corrective actions;
- schools above 100 in the first 10 years would be subject to Level 1 Corrective Actions at most, depending on their growth (see previous section on Accountability Actions);
- all schools be eligible for rewards upon meeting the growth requirements discussed previously

Accountability Cycle

The TAC recommends that an accountability cycle incorporate the following:

1. Test (CRT and NRT) every year in every content area (English/Reading, mathematics, science, social studies) in the designated grades.
2. Make accountability decisions every two years, on the odd years.
3. Make accountability decisions on the basis of two years' "baseline" data and two years' "growth" data (except for the first cycle).
4. Average the two years of baseline data to yield an Initial or baseline score, and the two years of growth data to yield a growth School Performance Score.
5. Use the maximum available data to make accountability decisions during the initial start up of the system when tests are being phased in.

Special Cases of Student Accountability

The TAC recommends:

1. Include special education students in school accountability.
2. Include students attending special purpose schools in school accountability by assigning the student's scores to the school that consented to send the student. Where students are placed without the school's decision-making consent, do not hold the school accountable for the student.
3. Include in the school's accountability all students enrolled at the school when accountability tests are administered; schools with high proportions of students who moved in after October 1 of that school year would be able to appeal.
4. A student who does not take the CRT test will be assigned the lowest score, Unsatisfactory, unless validly excluded. A student who does not participate in a section of the CRT (e.g., mathematics) will be assigned a score of Unsatisfactory for that content area of the test.
5. A student who does not take the NRT test will be assigned an NRT Index of zero.

Special Cases of School Accountability

Where a school building does not include the standard accountability grades:

1. Pair a non-standard school with at least one other school so that across the pair, at least one grade of NRT and one grade of CRT data are included. For attendance and dropout data, use the data unique to each school.
2. Have the local school district make the final decision of how to pair schools. There should be a consistent district policy for such decisions, such as feed-forward or feed-backward.

Newly formed schools should be included in school accountability if they have at least one year of pretest data and one year of posttest data.

Schools with more than 20% population change (e.g., through reconfiguration) would be able to appeal.

Highest Possible Index Scores

CRT	200
NRT	Approximately 360 (200 is approximately the 91st percentile, and the 99th percentile ranges between approximately 240 – 280, depending on grade level, using the 55th and 75th percentile ranks as the 10-year and 20-year goals)
Attendance	183 for Grades K-8 217 for Grades 9-12
Dropout	200