

Executive Summary

Updated Recommendations to the State Board of Education for a New State Accountability Index

November 5, 2008

Pete Bylsma, EdD, MPA

CREATING THE ACCOUNTABILITY INDEX

The legislature requires the State Board of Education to develop a statewide accountability system that will help improve academic performance among all students in the state. Part of that requirement is to identify schools and districts for recognition and for receiving additional state support. To meet this requirement, the Board has developed an accountability index to sort schools and districts into different “tiers” based on multiple measures. Schools and districts in most need are given “Priority” status, making them eligible to receive more significant support. These Priority schools and districts would be required to participate in a state system of support if initial offers of more support are not accepted and substantial improvement does not occur after two years.

Several principles have guided the development of the system. The accountability system will (1) be transparent and simple to understand, (2) use existing data, (3) rely on multiple measures, (4) include assessment results from all grades and subjects tested statewide, (5) use concepts of the federal No Child Left Behind Act (NCLB) and its Adequate Yearly Progress (AYP) system when appropriate, (6) be fair, reasonable, and consistent, (7) be valid and accurate, (8) focus at both the school and district levels, (9) apply to as many schools and districts as possible, (10) use familiar concepts when possible, (11) rely mainly on criterion-referenced measures instead of norm-referenced measures; and (12) provide multiple ways to reward success.

The proposed index is based on how schools and districts perform on a set of five outcomes and four indicators. The five outcomes are the results of state assessments in four subjects (reading, writing, mathematics, science) and the “extended” graduation rate (for high schools and districts). These five outcomes are examined using four indicators: (1) achievement for all students, (2) achievement of low-income students, (3) achievement of all students compared to similar schools (controlling for the percentage of students who are learning English, have a disability, live in low-income homes, and are mobile), and (4) improvement. The results of the 20 measures form a matrix as shown in Table 1.

Table 1: Accountability Matrix

INDICATORS	OUTCOMES				
	Reading	Writing	Math	Science	Ext. Grad. Rate
Achievement					
Ach. of low-inc.					
Ach. vs. peers					
Improvement					

Each cell of the matrix is rated on a 5-point scale (from 0 to 4) using fixed benchmarks. Each of the four subjects is rated using the same set of benchmarks across the entire school (i.e., all subjects have the same set of benchmarks, and the assessment results are the aggregate totals for all the tested grades). **The index is the simple average of all 20 ratings.** The index ranges from 0.0 to 4.0 and is a number similar to a GPA where 4.0 is the highest score. Table 2 shows how each of the five outcomes are measured using the four indicators and the benchmarks that produce the ratings. Tier assignments are determined based on the index score. Schools and districts would fall into four tiers, with an in-depth analysis of the data and conditions of those in the lowest tier to see if they merit being placed in a fifth (Priority) tier.

Table 2: Benchmarks and Ratings for Outcomes and Indicators

	READING	WRITING	MATH	SCIENCE	EXT. GRAD. RATE ¹	
ACHIEVEMENT (ALL STUDENTS)	<u>% MET STANDARD</u>		<u>RATING</u>		<u>RATE</u>	<u>RATING</u>
	86-100%		4		> 95	4
ACHIEVEMENT (LOW INCOME)	70-85.9%		3		85-94.9%	3
	55-69.9%		2		75-84.9%	2
	40-54.9%		1		65-74.9%	1
ACHIEVEMENT VS. PEERS²	< 40%		0		< 65%	0
	<u>DIFFERENCE IN LEARNING INDEX</u>		<u>RATING</u>		<u>DIFFERENCE IN RATE</u>	<u>RATING</u>
	> .20		4		> 12	4
	.10 to .20		3		5.01 to 12	3
	-.099 to .099		2		-5 to 5	2
IMPROVEMENT³	-.20 to -.10		1		-5.01 to -12	1
	< -.20		0		< 12	0
	<u>CHANGE IN LEARNING INDEX</u>		<u>RATING</u>		<u>CHANGE IN RATE</u>	<u>RATING</u>
	> .12		4		> 6	4
	.05 to .12		3		3.01 to 6	3
	-.05 to .05		2		-3 to 3	2
	-.051 to -.12		1		-3.01 to -6	1
	< -.12		0		< -6	0

Note: Assessment results include both WASL and WAAS results.

¹ This outcome only applies to schools and districts that are authorized to graduate students.

² This indicator adjusts the outcomes using statistical methods (multiple regression) to control for four student characteristics beyond a school’s control: the percentage of low-income, ELL, special education, and mobile students. (Mobile students are those who are not continuously enrolled from October 1 through the testing period.) Scores are the difference between the actual level and the predicted level. Scores above 0 are “beating the odds” and negative scores are below the predicted level. Separate analyses are conducted for each of the four assessments for each type of school (elementary, middle, high).

³ Measured in terms of the change from the previous year.

INITIAL RESULTS

Table 3 shows the ranges for the tier assignments and the number of schools and districts that would have been placed in each tier in 2007 using the above criteria.

Table 3: Tier Ranges and Preliminary Results (2007)

Tier	Index Range	Percent of Schools	Percent of Districts
Exemplary	3.00 – 4.00	4%	1%
Good	2.00 – 2.99	32%	35%
Acceptable	1.00 – 1.99	51%	59%
Struggling	0.00 – 0.99	13%	5%
Priority (eligible for Innovation Zone) ¹	0.00 – 0.99	TBD	TBD

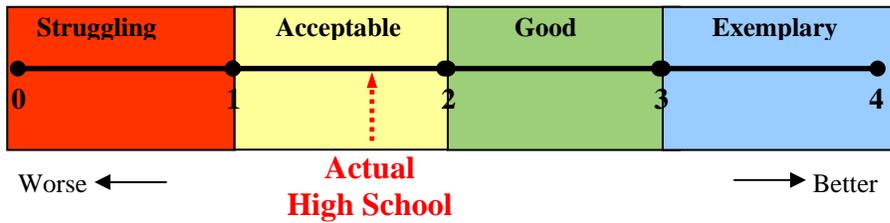
¹ Those in this tier would be determined after an in-depth analysis of their data and local conditions.

About 40% of the schools in “struggling” tier were alternative schools or served other special populations. Schools in this tier had a total enrollment of about 83,000 students, with about 70,000 attending “regular” schools. About 10% of the schools in the state had a 2-year average index below 1.00; about 5% of the schools statewide were “regular” schools with a 2-year average index below 1.00 (total enrollment was about 50,500 students). Fewer districts were in the exemplary and struggling tiers compared to the school results. However, 22 districts had at least two regular schools with a 2-year index average below 1.00, and eight districts had at least four regular schools with a 2-year index average below 1.00.

Table 4 provides an example of the ratings for an actual high school and how the average of the individual ratings generates the index/tier assignment. The school’s average rating of 1.65 is the index score, which puts the school in the middle of the “acceptable” tier. The index is shown graphically relative to the entire continuum. Tiers and average ratings are color-coded to correspond with the colors used for the WASL levels shown on the OSPI Web site. A set of “stars” indicate the rating so the overall results can be seen at a glance. These types of results could be made public on the Web site (the format for presenting the results must still be determined). Results presented in this “dashboard” give policymakers, educators, and the public a quick snapshot of where a school is strong and weak, its overall rating, and where it falls within the tier. It also provides transparency about how the index number is determined.

Table 4: “Actual” High School, 2007

Indicator	Reading	Writing	Math	Science	Grad Rate	Average
Achievement	3	3	1	0	3	2.00
Low-inc. ach.	2	2	0	0	4	1.60
Ach. vs. peers	1	1	1	1	3	1.40
Improvement	0	2	0	2	4	1.60
Average	1.50	2.00	0.50	0.75	3.50	1.65
Achievement	***	***	*		***	
Low-inc. ach.	**	**			****	
Ach. vs. peers	*	*	*		***	
Improvement		*		**	****	



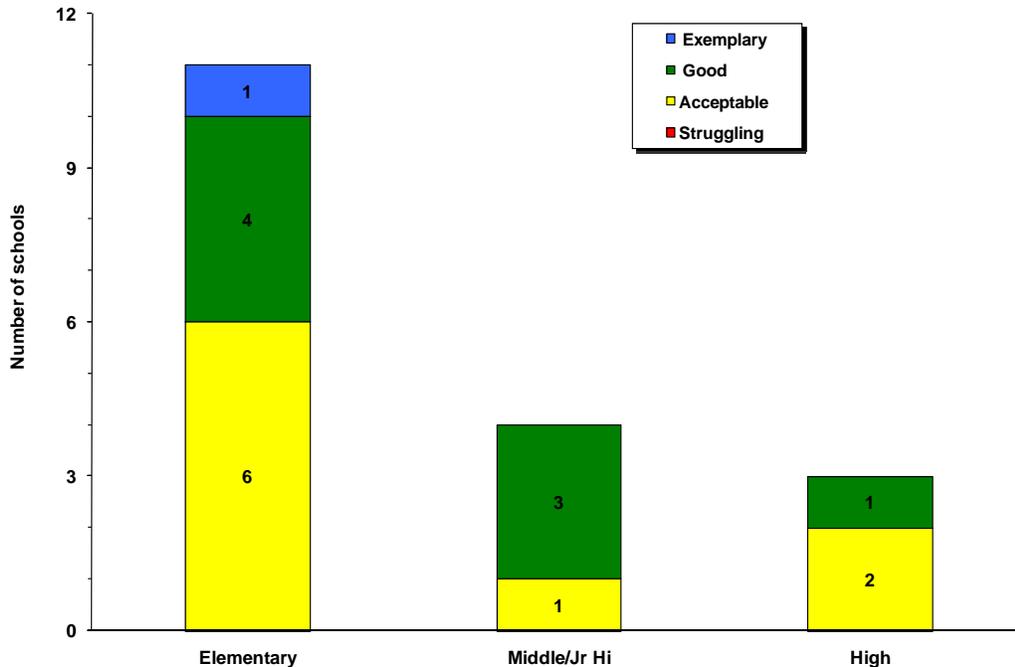
The proposed system holds *districts* accountable using the same indicators, outcomes, and criteria that are used for schools. The results are based on districtwide data for all grades rather than being disaggregated by grade bands (elementary, middle, high). In addition, financial data are used in the “peers” analysis to control for the amount of total operating expenditures per pupil (adjusted for student need). A deeper analyses would also occur for districts that have an index number in the “struggling” tier to determine if they merit receiving extra support.

Other tables and charts can illustrate school and district results as well. Table 5 shows how all the results can be shown across multiple years for a hypothetical district (data in shaded cells are not available). In addition, Figure 1 shows the distribution of the *number of schools* by tier for an actual district. These are examples of how results could be displayed. The actual methods for displaying the results must still be determined.

Table 5: Showing Longitudinal District Results (All Grades)

	YEAR			
<i>Indicator/Outcome</i>	2004	2005	2006	2007
Achievement	1.25	1.25	1.60	1.60
Reading	**	***	***	***
Writing	**	**	**	***
Math	*	*	*	**
Science				
Grad. rate	NA	**	**	**
Low-income ach.	0.50	0.80	1.00	1.20
Reading	*	**	**	**
Writing	*	*	**	**
Math				*
Science				
Grad. rate	NA	*	*	*
Ach. vs. peers	2.00	2.00	2.00	2.00
Reading	**	**	**	**
Writing	**	**	**	**
Math	**	**	**	**
Science	**	**	**	**
Grad. rate	NA	**	**	**
Improvement	3.67	3.25	2.60	1.80
Reading	****	****	**	**
Writing	NA	***	****	**
Math	****	***	**	**
Science	***	***	**	**
Grad. rate	NA	NA	***	*
INDEX	1.73	1.84	1.80	1.75

Figure 1: Distribution of Schools by Grade Level and Tier in “Actual” District



The proposed system does not include AYP results generated for NCLB. Feedback from all the stakeholders revealed a lack of confidence in the validity of AYP results for accountability purposes. The proposed system is not only more valid and transparent for accountability purposes, but it is more inclusive than the federal system because it includes both writing and science, uses a smaller minimum number for reporting (10 students across the entire school/district), and includes the results of all students, regardless of how long they have been attending school or district. It also combines results across all grades, which reduces the volatility of the results over time.

IDENTIFYING “PRIORITY” SCHOOLS AND DISTRICTS (LOWEST TIER)

Various quantitative and qualitative data will be used to determine which schools and districts that fall in the “struggling” tier should be placed in the “Priority” tier and be eligible to receive significant support. The data fall in four categories.

- **Contextual Data:**
 - Type of school
 - Changes in student population
 - Programs served by the school
 - Level of student mobility
- **Assessment Results (WASL/WAAS/WLPT)**
 - Trends over multiple years for each subject area
 - Subgroup trends
 - Results for students who have been enrolled for at least two years
- **AYP Results:**
 - Distance from the annual goal
 - Type of cells not making AYP
 - Percentage of cells not making AYP

- **Other Data:**

- Graduation and dropout rates for subgroups
- Student/teacher ratio
- Teacher education and experience levels
- Funding from local levies/bonds and outside sources
- Recent changes in leadership (key central office staff and principals) and teachers

Each year, the process would begin when OSPI computes the index using the most recent data and prepares a set of preliminary results. Given the relatively large number of schools that may fall into the “struggling” tier,¹ the schools must be screened to eliminate those that clearly should not fall into the Priority tier. This will reduce the number of schools and districts that require a deeper analysis. OSPI staff would review the index results for each school and district in the “struggling” tier and sort them into two categories:

- (1) Schools/districts that *remain in the struggling tier* are those that have not been in this tier in the past two years or have obvious data problems that affected their results (e.g., errors in reporting the number of graduates, missing data for ELL, special education, and low income students that can affect the results of the “peers”).
- (2) The remaining schools/districts are placed in a *possible Priority tier* category pending a deeper analysis.

OSPI staff will conduct a deeper analysis using available data for the schools and districts placed in the possible Priority tier category. This may require contacting the district and/or local ESD to get more information. Based on this review, the schools and districts are sorted again into the same two categories. Those placed in the possible Priority tier are notified of the possible designation and given the reasons why designation is possible. The district/school is given a chance to avoid the Priority designation by providing more information, including what explains the low index results. Appeal would then be made to OSPI with local school board approval. OSPI would review the additional information, and then recommend a final Priority list to the State Board of Education for review and approval.

INTEGRATING THE SYSTEMS

Federal law requires states to have a single accountability system. Many states combine their state accountability system with the federal NCLB system. Washington state can pursue two options to meet this requirement.

1. The preferred approach is to request that the state system be used in place of the current federal system. A new administration may provide more flexibility to states that design alternative systems. The proposed system has many desirable features that could make it a viable alternative to the current rules used to measure AYP.
2. If Washington is not allowed to use the proposed system to replace the current AYP system, the results of the index calculations will still be used to help determine the type of assistance the state provides. Those in “improvement” status under AYP would still face the federally required sanctions. Schools with relatively favorable index results that do not make AYP and fall into school improvement will receive minimal assistance from the state. In addition, some schools will make AYP and not be in school improvement, but they still have relatively low

¹ The number will still be far fewer than those not making AYP or identified for “improvement” under NCLB.

index results. (This happens most often in small schools that have less than 30 continuously enrolled students in a grade band.) In these cases, state funds can be used to focus assistance in the areas of greatest need.

If two systems coexist, the state will clarify what happens when schools and districts fall into the various AYP categories and state tiers in order to minimize any confusion that could occur about the two ways for measuring accountability.

RECOGNITION

The Board intends to provide recognition based on sustained exemplary performance. The accountability system will provide multiple ways to reward success and will rely on criterion-referenced measures using the results from the accountability matrix. Three options should be considered: providing recognition for (1) each of the 30 cells of the matrix, (2) the 20 “inner” cells of the matrix, and (3) the 10 “average” cells of the matrix. Advisors recommended providing recognition in all 30 cells based on the belief that people are motivated to improve the most when they can experience success. A minimum rating of 3.00 is required to receive recognition in the 20 “inner” cells, and a minimum of 2.75 rating is needed to receive recognition for the “averaged” cells (see Table 6). Any cell with a 3.5 or above would receive recognition “with honors.” The ratings will be calculated every year, and recognition is given when the two-year average rating meets the minimum requirement.

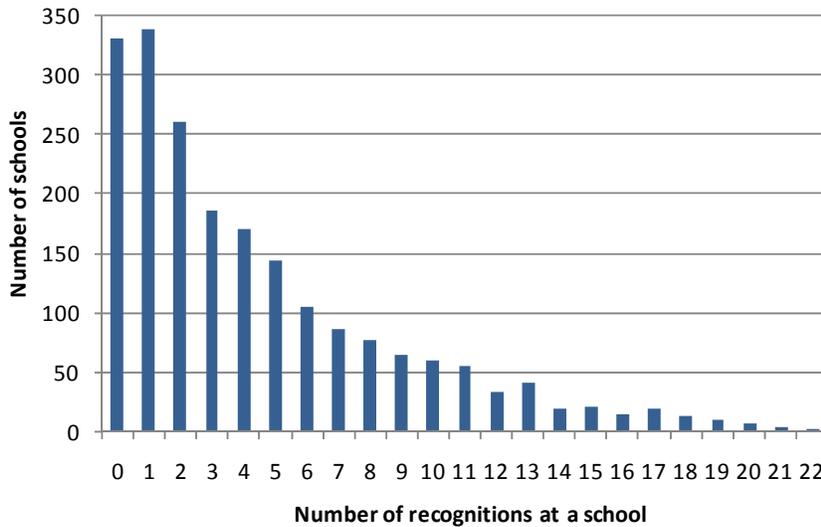
Table 6: Minimum Requirements for Recognition**

Indicator	Reading	Writing	Math	Science	Grad Rate	Average
Achievement	3.00	3.00	3.00	3.00	3.00	2.75
Ach. vs. peers	3.00	3.00	3.00	3.00	3.00	2.75
Improvement	3.00	3.00	3.00	3.00	3.00	2.75
Low-inc. ach.	3.00	3.00	3.00	3.00	3.00	2.75
Average	2.75	2.75	2.75	2.75	2.75	2.75

**Any cell of the matrix with a 2-year average rating of 3.50 or above would be recognized “with honor.”

Figure 2 shows the number of schools that would have received awards if the proposed system was in place in 2007 and all 30 cells were eligible to receive recognition. The largest number of schools would have received recognition in just one or two of the 30 areas, and 16% would not have received any recognition. At the other extreme, about 14% of schools would have received recognition in 10 or more areas, and 2 schools would have received recognition in 22 of the 30 cells of the matrix. The largest number of schools (52% of 2,046 schools) met the criteria for reading achievement. Achievement in math, science, and among low-income students had fewer schools meeting the criteria. Only 4% had an overall average of 2.75 on the accountability index over the 2-year period. Of the schools that had a 2-year index average of less than 1.00 (the “struggling” tier), 64% would not have received any recognition in any of the 30 cells, and the remaining schools averaged only one area of recognition among the 30 possible cells (it was nearly always an “improvement” cell that had a 2-year average that met the minimum criteria).

Figure 2: Number of Schools of Distinction, by Number of Recognitions (2007)



This system of recognition will supplement and could replace some types of recognition currently in place. The federal government provides funding for three awards, primarily for schools receiving Title I funds. OSPI also provides awards for improvement but no extra funding as part of its recognition. Schools and districts that receive recognition in the proposed system will not be compensated monetarily, although exceptions could be made. In its compensation proposal to the Basic Education Finance Task Force, OSPI recommended that schoolwide financial rewards be given each year when a school reaches a certain sustained level of improvement. The improvement dimension of the proposed recognition system could be used as a basis for these rewards. For example, schools that have an average of at least 3.0 for overall improvement could be given a schoolwide financial bonus. In 2007, about 8% of the schools statewide would have qualified for this bonus.

* * * * *

The proposed accountability system will need to remain flexible to adapt to changes in NCLB and graduation requirements, the assessment system, and other factors that may impact the results. Moreover, a number of issues must still be resolved before the index can be implemented effectively. For example, further review of the results should occur to ensure the index measures the achievement and improvement the Board intends. Various OSPI and State Board activities need to be integrated and aligned with one another to avoid duplication and confusion (e.g., how the index relates to NCLB requirements, how to use the index to identify Priority schools and districts, how and when assistance and recognition occur, how index results are represented and made available to the public). Further study is needed to ensure alternative schools and other “buildings” that serve populations with special needs are held accountable in appropriate ways. Finally, the method for measuring improvement needs to be reviewed, particularly when a school is already achieving at very high levels or far above its peers.