



State Board of Education System Performance Accountability
Policy Framework Proposal
October 28, 2008

SUMMARY OF POLICY ISSUES/SBE STRATEGIC PLAN GOAL

Washington State statute¹ assigns the State Board of Education (SBE) the authority to create a statewide accountability system, which includes:

- Identifying objective, systematic criteria for successful schools and districts.
- Identifying objective systematic criteria for schools and districts in need of assistance or where significant numbers of students persistently fail to meet state standards.
- Identifying range of state intervention strategies for legislature to consider authorizing.

The Board has three strategic plan goals, which are the underpinnings of an effective statewide accountability system: 1) improve student achievement; 2) improve graduation rates; and 3) improve student preparation for success in post secondary education, 21st century world of work and citizenship.

BACKGROUND

Why has the Board engaged in this work? ***All students deserve to receive a quality education.***

The Board believes in continuous improvement for all schools and districts to ensure that they have the tools to do their work to help all their students. The Board wants to recognize schools that are doing an outstanding job for their students and many of them are. Like all states, Washington has a small number of schools where students persistently achieve at significantly lower levels than at peer schools. The Board estimates that 70,500 students² are enrolled (one out of 14 students in the K-12 system) in struggling schools (identified by its accountability index). There are no state incentives or significant consequences for making transformational changes in these schools and districts, thus the need for the Board's work, to help these students.

Also like all states, Washington has not been able to eliminate the large achievement gap between affluent and high-poverty students and schools. And finally, Washington's public schools are not yet broadly and successfully preparing most high school graduates with work-ready and or college-ready skills, after 15 years or more of standards-based reform.

¹ RCW 28A.305.130 (4)

² If alternative education students are included, the number is 83,000.

The Board has engaged in an extensive review of accountability issues through its work sessions with the Office of Superintendent of Public Instruction, advisors, stakeholder meetings, community surveys, and research over the last two years. The Board has examined other states' accountability systems, national studies on high performing high poverty schools and accountability issues, and the policy barriers to student achievement in Washington, as identified by policy makers and practitioners.

It has retained the services of consultants to assist with the development of the proposals described below in the framework. The consultants worked with Washington practitioners to develop their proposals. National and Washington-based research reveals a clear set of barriers that have undercut the impact of school reform efforts to date. They include insufficient and unstable resources, insufficient time, inflexibility in allocating resources to higher need areas to improve student achievement, lack of coherent systems to recruit and prepare quality educators, insufficient coordination among intrastate agencies, and insufficient focus (i.e., with funding) on schools serving high-challenge student populations.

The background information and consultant reports for this work can be found on the Board's Web site: www.sbe.wa.gov

DRAFT SYSTEM PERFORMANCE ACCOUNTABILITY POLICY FRAMEWORK

The System Performance Accountability Policy Framework operates under a central premise: all schools and their districts should be engaged in continuous improvement efforts to ensure that all students are reaching their highest potential.

Four suggested guiding principles to this overall accountability policy framework (based on feedback received):

- All students will have a quality education
- Basic Education will be redefined and funded
- A reciprocal relationship will be created between the state and local school district for student success
- The state will create one unified accountability system

The framework includes key and connected components to identify ways to focus on increasing student achievement:

1. An **accountability index**, which uses objective systematic criteria to identify successful schools and districts, as well as those in need of assistance or those where students persistently fail to meet state standards. Those in the latter category will be analyzed in greater detail after identification through the accountability index to develop a list of "Priority Schools" that clearly demonstrate a need for additional support.
2. A **preventative, proactive system** of support to help all schools and districts continue to improve, which would be voluntary for districts except where the accountability index indicates a clear need for support in specific areas, such as closing the achievement gap among certain subgroups (e.g., English Language Learners or African-American students), or in certain curriculum areas including math and science.
3. A range of **voluntary state and/or local district intensive assistance strategies** for districts with one or more "Priority Schools", to develop a systems approach for improving student academic performance including: a) the voluntary state/local Innovation Zone Program, which would allow local school boards (together with their superintendents, union leaders and other stakeholders) to create a systemic turnaround effort that directly and comprehensively addresses the barriers to reform identified by the research cited earlier, supported by state investment in resources and capacity; b) the

voluntary state/local Summit District Program (OSPI operates this currently under No Child Left Behind identification), a district-wide reform initiative focused on developing effective leadership, quality instruction, data analysis, needs assessment, and targeted strategies for improvement, supported by state investment in resources and capacity; or c) a voluntary local district program to develop and implement its own strategies to bring its schools out of Priority status with state approval of the district's plan and accompanying state resources and support.

4. A category of deeper state and local partnership, called **Academic Watch**, if after two full years of implementation there is insufficient progress under any of the forms of intensive assistance described above based on the accountability index and follow-up review of local district conditions and strategies.

I. THE PROPOSED ACCOUNTABILITY INDEX

SBE has developed a draft accountability index to sort schools and districts into different “tiers” based on multiple measures. It is expected that additional work will need to be done to refine this accountability index over the next six months.

Schools and districts in most need are given “Priority” status, making them eligible to receive more significant support as outlined under Table III below.

A set of principles has 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 and familiar concepts; (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) rely mainly on criterion-referenced measures instead of norm-referenced measures in order to create clear goals and encourage cooperation among educators; and (11) provide multiple ways to demonstrate and 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 below.

Table 1: Accountability Matrix

INDICATORS	OUTCOMES				
	Reading	Writing	Math	Science	Extended Grad Rate
Achievement					
Achievement of low-income					
Achievement 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
	70-85.9%		3		85-94.9%	3
ACHIEVEMENT (LOW INCOME)	55-69.9%		2		75-84.9%	2
	40-54.9%		1		65-74.9%	1
	< 40%		0		< 65%	0
ACHIEVEMENT VS. PEERS²	<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
	-.20 to -.10		1		-5.01 to -12	1
	< -.20		0		< 12	0
IMPROVEMENT³	<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.

¹This 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 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 in elementary, middle, and high schools.

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

INITIAL RESULTS

Table 3 shows the suggested 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% ¹	5%
Priority (eligible for Innovation Zone) ²	0.00 – 0.99	TBD	TBD

¹ About 40% of the schools in this tier were alternative schools or served other special populations. Schools in this tier had a total enrollment of about 83,000 students, with about 70,500 attending “regular” schools. About 78% of the schools in this tier had a 2-year index average below 1.00, and 55% of these were “regular” schools with a total enrollment of about 50,500 students. There were 22 districts that had at least two regular schools with a two-year index average of less than 1.00, and eight districts had at least four regular schools with a two-year index below 1.00.

² Those in this tier would be determined after an in-depth analysis of their data and local conditions.

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 falls into four categories:

1. Contextual Data:

- Type of school
- Changes in student population
- Programs served by the school
- Level of student mobility

2. 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

3. AYP Results:

- Distance from the annual goal.
- Type of cells not making AYP
- Percentage of cells not making AYP

4. 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,³ the schools must be screened to eliminate those that clearly should not fall into the Priority tier, which would 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 will remain in the “struggling” tier if the in-depth analysis provides good cause for why they should not be a part of the Priority Schools Tier.
- (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 will be sorted again into the same two categories. Those placed in the possible Priority tier will be notified of the possible designation and given the reasons why designation is possible. The district/school will be given a chance to avoid the Priority designation by providing more information, including what explains the low index results. Districts, with school board approval, could appeal to OSPI. OSPI would review the additional information, and then recommend a final Priority list to the State Board of Education for review and approval.

RECOGNITION

The Board intends to provide recognition based on sustained exemplary performance, and it will provide multiple ways to reward success using the results from the accountability matrix. The Board is considering three options: 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. A minimum rating of 3.00 is required to receive recognition in the 20 “inner” cells, and a minimum rating of 2.75 is needed to receive recognition for the “averaged” cells (see Table 4). 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. This system of recognition will supplement the federal and state awards currently in place.

³ The number will still be far fewer than those not making AYP or identified for “improvement” under NCLB.

Table 4: Minimum Requirements for Recognition**

	Reading	Writing	Math	Science	Extended 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.”

II. PREVENTIVE, PROACTIVE SYSTEM OF SUPPORT FOR ALL SCHOOLS AND DISTRICTS

Traditionally the state has not had a strong comprehensive program of general assistance for all schools and districts to improve student achievement. The State Board of Education requires that all schools have school improvement plans with specific elements, which the local school board approves and monitors. OSPI is in the process of developing more ways, in partnership with the Educational Service Districts and local districts, to provide general and more targeted assistance. Some of the areas that they are working on include: online professional learning and data collection tools; school and district plan management tools, regional training on specific areas such as English language learners, reading and mathematics. Additional areas could include training for school board members from the Washington State School Directors Association (WSSDA).

The accountability index will help districts and the state identify areas of particular challenge, and in these areas, districts meeting certain criteria for underperformance will be required to participate in a new set of state services designed specifically to help them meet these specific challenges. OSPI plans to create services focused on helping districts that are trying to close an achievement gap with one or more subgroups of students (e.g. English Language Learners, African- American or other groups) and may include services designed to assist with certain curriculum areas including math and science.

III. VOLUNTARY INTENSIVE ASSISTANCE STRATEGIES FOR DISTRICTS WITH PRIORITY SCHOOLS

Priority Schools designation reflects school-wide issues that go beyond achievement gaps for students facing certain challenges or within selected curriculum areas, and therefore need a more comprehensive solution as described in Table II above. Districts will be notified by OSPI that they have schools with a Priority designation. Districts will have two years of full implementation (plus six to nine months of planning time) to move their schools out of the Priority Schools designation using one of the following three strategies to work in greater intensity to improve student achievement:

a) the voluntary state/local Innovation Zone Program, which would allow local school boards (together with their superintendents, union leaders, and other stakeholders) to create a systemic turnaround effort that directly and comprehensively addresses the barriers to reform identified

by the research cited earlier, supported by state investment in resources and capacity;

b) the voluntary state/local Summit District Program (OSPI operates this currently using NCLB and primarily federal resources), a district-wide reform initiative focused on developing effective leadership, quality instruction, data analysis, needs assessment, and targeted strategies for improvement, supported by state investment in resources and capacity;

c) a voluntary local district program to develop and implement its own strategies in a plan to bring its schools out of Priority status with OSPI approving the district's plan and providing resources and support.

For the purposes of this memo, we will focus on the Innovation Zone as one of the three options open to districts under Table III, acknowledging that OSPI had provided, or is developing the programs in Table III b and c.

A. INNOVATION ZONE: AN OPPORTUNITY TO UNDERTAKE TRANSFORMATIONAL CHANGE

The Innovation Zone is provided to allow local districts, through their school board, to develop a performance contract with the state in exchange for state resources to assist them. The Innovation Zone concepts represent:

- **At the instructional level**, a chance for educators to ask fundamental questions about what it takes to help high-challenge, high-poverty students succeed, and to reshape their approach accordingly based on research conducted nationally and in Washington State.
- **At the systems level**, an opportunity for district and community leaders and their partners, supported by the state, to re-imagine and rebuild the structures and operating habits that shape the nature and quality of the education they offer.
- **At the policy level**, an effort to pilot the next generation of standards-based reform in Washington State – an approach marked by greater degrees of accountability by every stakeholder in the enterprise.

- **Key Elements of the Innovation Zone:**

- **Making the reforms systemic and “scale-able.”** Districts with Priority Schools as determined by the state's Accountability Index will be encouraged to apply to the Innovation Zone on behalf of a small cluster of schools – including their Priority School(s) – organized intentionally by feeder pattern or school type (within or across district lines), so that the reforms are systemic and scale-able, rather than being limited to a focus on individual schools.
- **Focusing on those districts best positioned to achieve success.** Districts will be selected to develop a comprehensive Innovation Zone plan after careful vetting by OSPI and the SBE for readiness (i.e., strong signals of commitment to transformative change; evidence that it will be a collaborative effort among district leaders, including the school board, superintendent, teachers' union, and community officeholders; and a strong preliminary plan).
- **Establishing demanding criteria and encouraging districts to enlist a highly capable lead partner.** Districts will be provided with resources to develop their Innovation Zone plan. The SBE recognizes that in most cases, districts will need outside support to produce a plan that meets the rigorous criteria the SBE will establish for Innovation Zone plans. The SBE will instruct OSPI to assist with this

process and to facilitate the development of partnerships between districts and lead turnaround providers, both for the planning cycle and for implementation.

- **Incorporating changes in operating conditions into the Zone criteria.** Through a collaborative, local process involving all key stakeholders (district administrators, school board, union, community, and parents), districts with Priority Schools that want to apply to the Innovation Zone will need to develop more flexible operating conditions that research shows are required for transformational – not simply incremental – change and to serve high-challenge, high-poverty students successfully.

The four critical elements of the Innovation Zone are contrasted with Traditional School Improvement in the following chart to understand the differences:

Traditional School Improvement Approach	Transformative, Comprehensive Turnaround Approach (Innovation Zone)
CRITICAL ELEMENT #1: PEOPLE	
Help current staff perform at a higher level through training, coaching, and leadership development.	Turnaround leaders have all necessary tools and authority to fulfill the turnaround plan including: recruiting incentives; flexibility on staff hiring, allocation; and time to make staff development coherent.
CRITICAL ELEMENT #2: TIME	
Tweak existing schedule, while maintaining same-length school day and year.	Strategic assessment to determine if expanding school day, school year, and/or significant change to the schedule is necessary to fulfill the plan; resources to help fulfill those requirements.
CRITICAL ELEMENT #3: MONEY	
Minimal impact on budgetary authority. Sometimes includes additional resources generally for staff development.	Strategic re-allocation of the budget is allowed. Additional resources are provided to support the implementation including: pay for extra time, incentives, and partner support.
CRITICAL ELEMENT #4: PROGRAM	
Improve quality of existing strategies through evaluation of curriculum, instruction, and assessment tools.	Development of a coherent, whole-school plan that integrates strategies to address impacts of poverty and other strategies shown to succeed in high-challenge schools. Also provides relief from compliance burden to allow focus on instruction.

IV. ACADEMIC WATCH IF NO IMPROVEMENT

It is intended that Academic Watch would be used only after all other intervention efforts fail to improve the academic performance of students in the District’s Priority Schools. The Accountability Index will be used to determine initially if the district’s schools have moved out of Priority Status. OSPI will verify this status based on additional analysis.

Based on this verification, OSPI will notify the district that it is on Academic Watch if the district has been unable to bring its Priority Schools out of Priority status after two full implementation years. The district will be required to undergo a performance or academic audit managed by one of a number of Peer Review Teams convened by OSPI. The Peer Review Teams will be composed of educators and experts with knowledge of school district processes and improvement strategies. The district will then take the performance audit and strategies and develop an implementation plan for approval by OSPI.

There are two options for the Board to consider under Academic Watch:

Option A: Authority for Implementation Remains with the Local District

After the corrective action and implementation plan is approved by OSPI, the local school board would be responsible for implementation of that plan and the state would provide needed resources to assist the district. OSPI would continue to monitor the district's progress with periodic updates to the State Board of Education.

Option B: Authority for Implementation Requires State-Specified Binding Conditions

OSPI may determine that the district requires a deeper level of state partnership to implement their plan successfully. In that event, after the corrective action and implementation plan is approved by OSPI, the Superintendent of Public Instruction may recommend to the State Board of Education that OSPI place the local school board under a set of binding conditions to carry out the corrective action and implementation plan. The State Board of Education could approve, disapprove, or modify the binding conditions. If the plan is not being carried out successfully after one year, OSPI and the Peer Review Team may recommend to the State Board of Education a new corrective plan of action for implementation for that district.

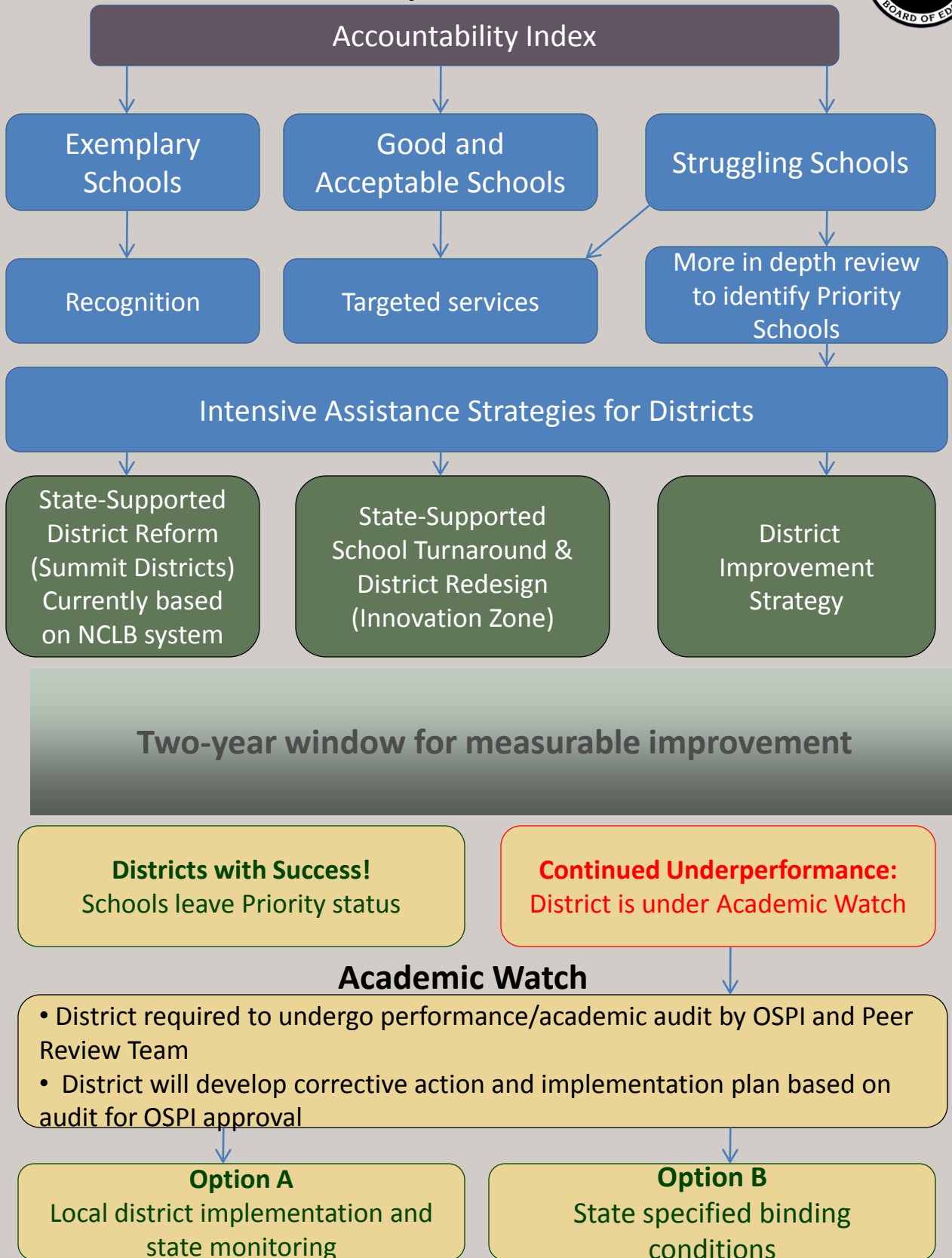
PROPOSED BOARD ACTIONS

In the interests of ensuring that all students in our state have access to a quality education, the Board will adopt the System Performance Accountability Policy Framework as follows:

1. Motion to adopt the general concepts of a system performance accountability policy framework, per this document, consisting of the guiding principles, found on the second page of this document, and: a) the accountability index; b) preventive, proactive assistance to all districts and schools; c) intensive assistance strategies for districts with one or more Priority Schools, and d) an "Academic Watch" for those districts with Priority Schools that continue not to improve student achievement using Option A or Option B (pick one) as outlined under "Academic Watch."
2. Direct SBE staff to work with OSPI on:
 - a. Refining the overall accountability index through:
 - i. A unified accountability system which creates a coherent system between the current NCLB system and the proposed SBE accountability index.
 - ii. Continued refinement of the draft accountability index that includes different weights assigned to indicators, additional data items such as, but not limited to, the number of cells a school misses Annual Yearly Progress and the percent of college/work ready courses high school students take.
 - iii. A request for national experts to review the SBE proposed accountability index to determine if it measures the achievement and improvement the Board intends.
 - iv. A review of achievement for different subgroups in the non-struggling tiers and recommending ways to address those gaps.
 - v. A request that the Federal government replace its current NCLB system with the state's proposed statewide accountability system.

- b. Refining the recognition system to:
 - i. Define a clear purpose for recognition that encourages schools to continue to make significant and sustained improvements that meet certain achievement levels.
 - ii. Coordinate the SBE proposed accountability index, with the current OSPI schools of distinction process, to determine and ensure a coherent system to identify districts and schools for recognition.
- c. Creating the necessary administrative structures between OSPI and the SBE to carry out the concepts for this new statewide accountability system to:
 - i. Ensure a system of program supports for continuous improvement of student achievement for all schools by identifying practices, policies, and tools necessary to assist and hold districts accountable for closing the achievement gap for students of poverty and color.
 - ii. Refine the continuum of programs available to schools and districts for targeted and intensive assistance, including the SBE-proposed Innovation Zone.
 - iii. Define the specific processes needed to move forward under the Academic Watch under Option A or Option B or an additional Option to be defined by Board members (pick one).
 - iv. Examine ways to address the needs of alternative education schools under the new accountability system.
- d. Defining the resources needed to implement the new statewide accountability system for the state and local districts.
- e. Providing periodic updates to the Board with a final report on the Accountability Index by June 30, 2009, and a final report on the recognition system and administrative structures and resources needed by October 15, 2009.

All Students Deserve a High Quality Education Continuous Improvement for All Schools



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

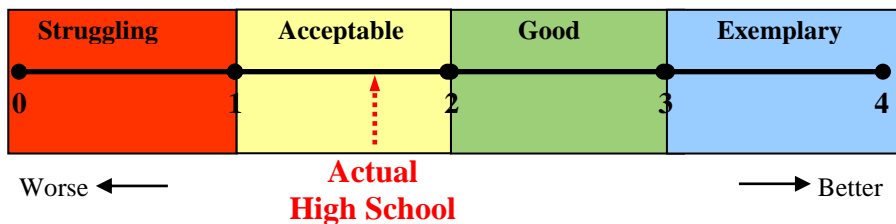
¹ Those in this tier would be determined after an in-depth analysis of their data and local conditions.

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		*		**	****	



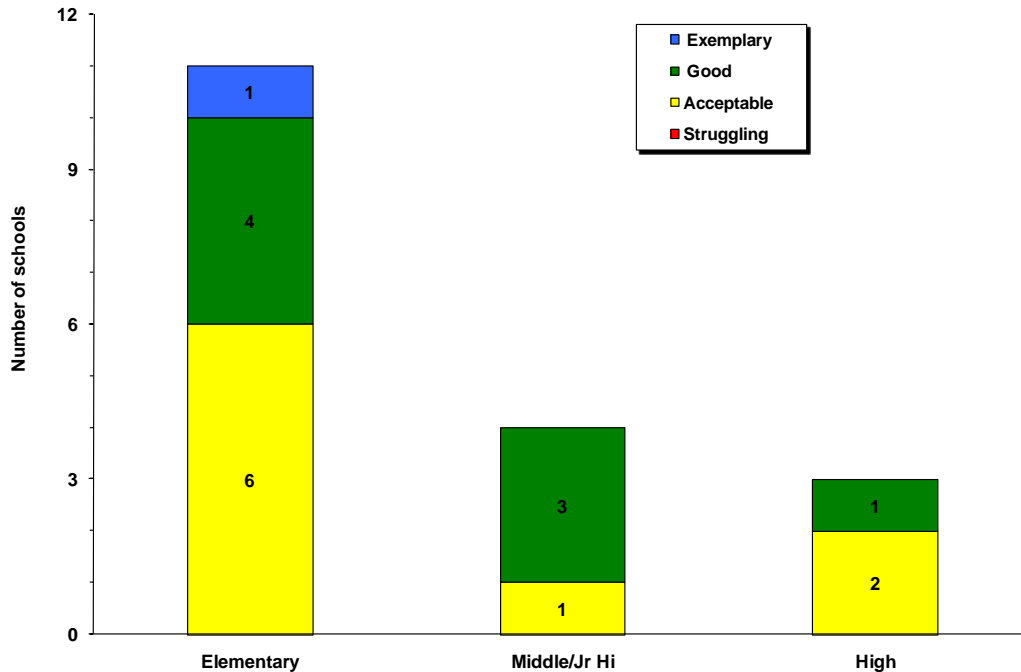
The proposed system holds *districts* accountable using the same indicators, outcomes, and criteria that are used for schools. The results are based on districtwide data for all grades rather than being disaggregated by grade bands (elementary, middle, high). In addition, financial data are used in the “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.

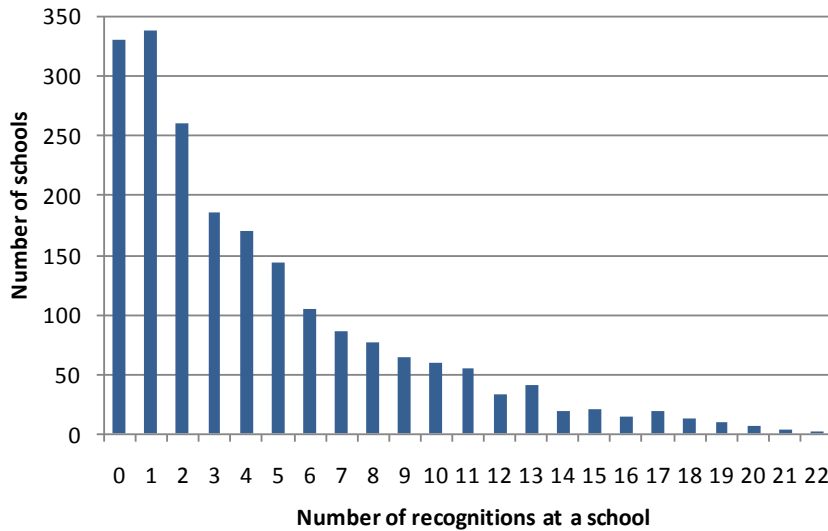
Table 6: Minimum Requirements for Recognition**

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.

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

November 5, 2008

Pete Bylsma, EdD, MPA

TABLE OF CONTENTS

Background	2
Accountability Index	3
District Accountability	7
Identifying Priority Schools and Districts (Lowest Tier)	14
Integrating the Systems	16
Recognition	16
Appendix A – Indicators and Outcomes	20
Appendix B – Identifying Priority Schools and Districts	37
Appendix C – Current State Assistance Program	40
Appendix D – Advisory Group Members	43
Appendix E – Relevant Legislative Mandate	44

BACKGROUND

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 is developing an accountability index to sort schools and districts into different “tiers” based on multiple factors. Various stakeholders and advisors have provided input and feedback about the proposed index and the data that can be used to help identify “Priority” schools and districts in most need. (Mass Insight is designing a system to support the schools and districts in most need, and this system will be aligned with the system of support that OSPI offers.) This document provides the initial recommendations for the index and information about identifying Priority schools and districts.

A set of principles guided the development of the accountability index. Specifically, the index will:

- Be transparent and simple to understand;
- Use existing data;
- Rely on multiple measures;
- Include assessment results from all grades (3-8, 10) and subjects tested statewide (reading, writing, mathematics, science);
- Incorporate concepts of the federal No Child Left Behind (NCLB) Act and its Adequate Yearly Progress (AYP) system when appropriate;
- Be fair, reasonable, and consistent;
- Be valid and accurate;
- Focus at both the school and district levels;
- Apply to as many schools and districts as possible;
- Use familiar concepts when possible;
- Rely mainly on criterion-referenced measures instead of norm-referenced measures; and
- Provide multiple ways to reward success.

Three assumptions were made during the development of the index.

- Priority schools and districts should be those that are the most challenged in the state – they should meet a “common sense” test as those needing the most support.
- Priority schools and districts would be eligible to receive additional resources to make dramatic improvement in student outcomes through an initiative such as that being developed by Mass Insight. Criteria to be met to receive this support will be specified by the State Board of Education.
- Priority schools and districts would be required to participate in a state-supported initiative, as described by the system being designed by Mass Insight, if offers of additional support are not accepted and substantial improvement does not occur after two years.

ACCOUNTABILITY INDEX

The proposed index is based on how schools and districts perform on a set of outcomes and indicators. Specifically, the recommended index uses a matrix 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 measured using four indicators: (1) achievement, (2) achievement of students from low-income families, (3) achievement compared to peers (the predicted level controlling for four student characteristics—special education, ELL, low income, and mobility), and (4) improvement. This results in 20 measures, forming the matrix in Table 1.

Table 1: Matrix of Accountability Measures

	OUTCOMES				
INDICATORS	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 (0-4) using a set of fixed benchmarks. These benchmarks reflect the performance in each cell, with 4 being the best outcome. 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 higher the index, the better the level of performance of the school/district.

Table 2 shows the four indicators, the five outcomes, and the benchmarks that produce the various 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. This numbering scheme also reflects the same system used to describe the levels of performance on the WASL (Levels 0-4). The Learning Index is used to measure the assessment outcome for two indicators: *achievement compared to peer schools* and *improvement*. This index (not to be confused with the accountability index) takes into consideration the percentage of students performing at the five different WASL levels, not just those meeting standard. The Learning Index ranges from 0 to 4, with 4.00 the highest score (similar to a grade point average). This index is explained in detail in Appendix A.

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
	< 40%		0		< 65%	0
ACHIEVEMENT VS. PEERS²	<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
	-.20 to -.10		1		-5.01 to -12	1
	< -.20		0		< 12	0
IMPROVEMENT³	<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-related results are the combined results of both the WASL and WAAS from all grades.

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.

The proposed system does not include AYP results generated for NCLB. Feedback from the advisors, members of the Board, and other stakeholders showed a lack of confidence in the validity of AYP results for accountability purposes. The proposed system 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. Nevertheless, various stakeholders believe AYP results still have a role in the state accountability system because (1) the law will likely remain in effect for several more years and AYP results must be calculated, (2) the disaggregation of results by subgroups provides additional details that provide deeper insights into the level of student learning taking place in schools and districts and at individual grade levels, and (3) federal law requires a single accountability system, which means AYP results need to be included in some way. As a result, the proposed system uses AYP results as one source of data to identify Priority schools and districts once initial index numbers are computed.

Tier assignments are determined based on the index score. Schools and districts would initially fall into four tiers based on their accountability index score, with an in-depth analysis of the data and conditions of those in the “struggling” tier to determine if they merit being placed in a fifth (Priority) tier and be eligible to receive more intensive support. The 5-tier system provides sufficient differentiation among schools and districts.

Table 3 shows the suggested ranges for the 5-tier system, along with a descriptive name. The index and tier can be made available in a “report card” for use by policymakers and the public, with a set of “stars” indicating the rating so the overall results can be seen at a glance. This intuitive rating symbolism is used in other settings (e.g., rating movies, restaurants, athletes, tourist attractions) and does not require much interpretation. Table 3 also shows the distribution of schools using the criteria shown in Table 2 and data from 2007. A total of 2,046 schools had an index score.

Table 3: Tier Ranges and 2007 Results (N=2,046)

Tier	Index Range	Number of Schools	Percent of Schools
Exemplary	3.00 – 4.00	72	3.5%
Good	2.00 – 2.99	664	32.5%
Acceptable	1.00 – 1.99	1,043	51.0%
Struggling	0.00 – 0.99	267	13.0%
Priority (eligible for Innovation Zone) ¹	0.00 – 0.99	TBD	TBD

¹Schools and districts in the lowest tier would be determined after an in-depth analysis of quantitative and qualitative information.

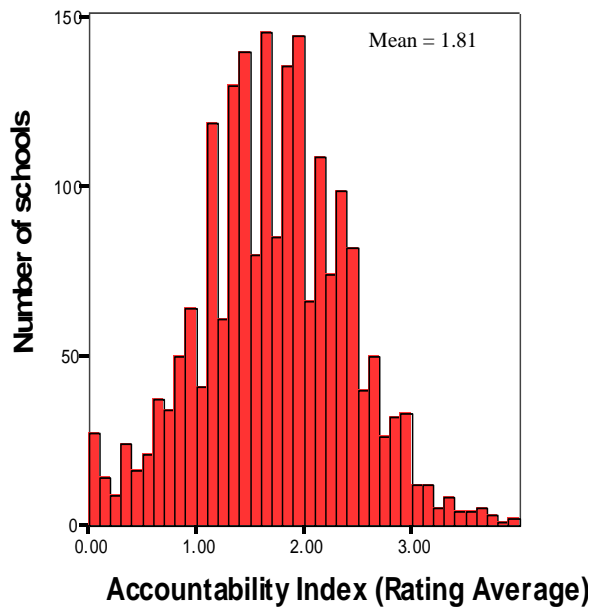
About 83,000 students were enrolled in the 267 schools in the struggling tier in 2007 (about 8.3% of all students statewide). Of these 267 schools, 103 (39% of this group) were alternative schools or served other special student populations. About 70,500 students attended the 164 “regular” schools that were in this tier.

Of the 267 schools in the struggling tier, 209 had a 2-year index average of less than 1.00. These 209 schools enrolled approximately 60,200 students. Of the 209 schools, 114 (55%) were “regular” schools that enrolled approximately 50,500 students (about 5% of the statewide student population). There were 22 districts that had at least two regular schools with a 2-year index average of less than 1.00, and eight districts had at least four regular schools with a 2-year index below 1.00.

Figure 1 shows the index distribution for the 2,046 schools in the analysis based on data in school year 2006-07. There was little difference in the distribution of schools based on their grades served (i.e., elementary, middle, high).¹

¹ Only one regular high school that had a 2-year average index of less than 1.00.

Figure 1: Distribution of Schools by Index Score, 2007**



** All the schools with an index of 0.00 served special populations (correctional facilities, alternative schools, dropout recovery programs), and most had fewer than 10 assessed students so their results would not be reported. The lowest index for a regular school was 0.13, but this school made substantial gains in 2008.

Tables 4 and 5 give examples of how the individual ratings generate the index/tier assignment for two actual schools using results available from 2007. The schools' final index is shown graphically relative to the entire continuum. The tiers and average ratings are noted in colors that correspond to the colors used for the WASL levels on the OSPI Report Card. The results could be made public as part of the OSPI Report Card (the format of the presentation must still be determined). Results presented in this type of "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.

- The *high school* described in Table 4 is located in a medium-sized suburb of a large city with fewer low-income students than the typical high school in the state. Its WASL scores had been about the state average in most subjects but both reading and math scores dropped dramatically from 2006 levels. Like many high schools, it has low math and science scores. It also has lower scores than high schools serving similar students, and the performance of its low-income students was below that of "all" students in three subjects. Its graduation rate is fairly high, even when compared to its peers, the rate improved substantially from the previous year, and surprisingly, low-income students had a higher rate than the "all" students rate. Its index of 1.65 puts it close to the middle of the "acceptable" tier, which is probably worse than educators and community members expected.
- The *elementary school* described in Table 5 is located in a medium-sized city with above-average levels of low-income, ELL, and mobile students. Its WASL scores are well above the state average in several grades but below the state average in one grade. It had sharp declines from very high WASL scores the previous year, resulting in low improvement ratings in 3 subjects. Its reading and writing scores are still quite high and its scores are

very high compared to schools serving similar students. Low-income students had the same rating as “all” students in three subjects but were lower in writing. The graduation rate does not apply. Its index of 2.13 is slightly above the middle of the index scale and in the lower end of the “good” tier.

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		*		**	****	

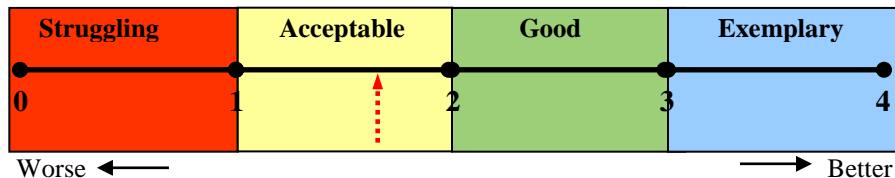
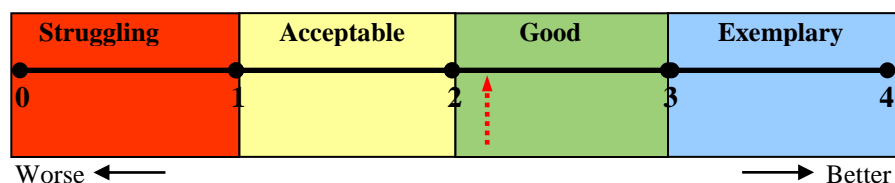


Table 5: “Actual” Elementary School, 2007

Indicator	Reading	Writing	Math	Science	Grad Rate	Average
Achievement	3	3	2	0		2.00
Low-inc. ach.	3	2	2	0		1.75
Ach. vs. peers	4	4	4	3		3.75
Improvement	0	2	1	1		1.00
Average	2.50	2.75	2.25	1.00		2.13
Achievement	***	***	**			
Low-inc. ach.	***	**	**			
Ach. vs. peers	****	****	****	**		
Improvement		*	**	*		



DISTRICT ACCOUNTABILITY

The proposed system would hold districts accountable using the same rules, indicators, and outcomes that are used for school accountability. The results would be based on districtwide data for all grades rather than being disaggregated by grade bands (elementary, middle, high). District results are more likely to be made public when using the combined results for all

grades—only five extremely small districts, with a combined total of 34 students, had fewer than 10 students in their tested grades in 2007. Financial data, which is available only at the district level on a consistent basis, is used the district-level “peer” analysis to control for the amount of total operating expenditures per pupil. The same type of deeper analyses would occur for districts that have an index number in the lowest tier in order to determine if they merit receiving extra support, just like the process used for schools. This closer look would also include examining the percentage of schools and number of students that are found in the lowest tier and the consistency of problems in a particular set of grade bands or subjects. Since more information is available at the district level, district accountability could include additional measures besides the 20 in the matrix. Moreover, other data could be used when analyzing districts and their peers, such as unemployment rates, crime rates, per capita income, and tax base if this information is available at the district level.

Various tables and charts can illustrate the district results. Table 6 and Figure 2 show how all the results for a district can be shown across multiple years to show trends over time. (State results are used, and the data in shaded cells of the table are not available.) Figure 3 shows the distribution of the *number of schools* by tier for an actual district. Figure 4 shows the *percentage of students* enrolled at those schools. (One alternative high school has relatively few students.)

Table 6: Showing Results Over Time (All Grades)

<i>Indicator/Outcome</i>	<i>YEAR</i>			
	2004	2005	2006	2007
Achievement	1.25	1.25	1.60	1.60
Reading	**	***	***	***
Writing	**	**	**	***
Math	*	*	*	**
Science				
Ext. grad. rate	NA	**	**	**
Low-income ach.	0.50	0.80	1.00	1.20
Reading	*	**	**	**
Writing	*	*	**	**
Math				*
Science				
Ext. grad. rate	NA	*	*	*
Ach. vs. peers¹	2.00	2.00	2.00	2.00
Reading	**	**	**	**
Writing	**	**	**	**
Math	**	**	**	**
Science	**	**	**	**
Ext. grad. rate	NA	**	**	**
Improvement	3.67	3.25	2.60	1.80
Reading	****	****	**	**
Writing	NA	***	****	**
Math	****	***	**	**
Science	***	***	**	**
Ext. grad. rate	NA	NA	***	*
INDEX	1.73	1.84	1.80	1.75

¹This indicator does not apply in this example (the state has no peer); a middle rating is given for all outcomes.

Figure 2: Average Ratings, 2003-2007

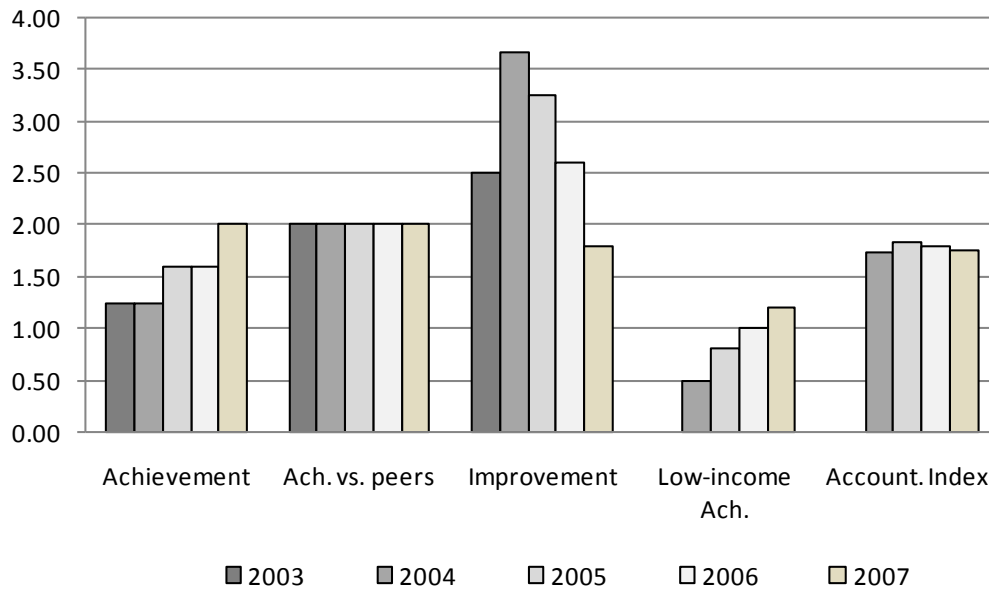


Figure 3: Distribution of Schools by Grade Level and Tier in “Actual” District

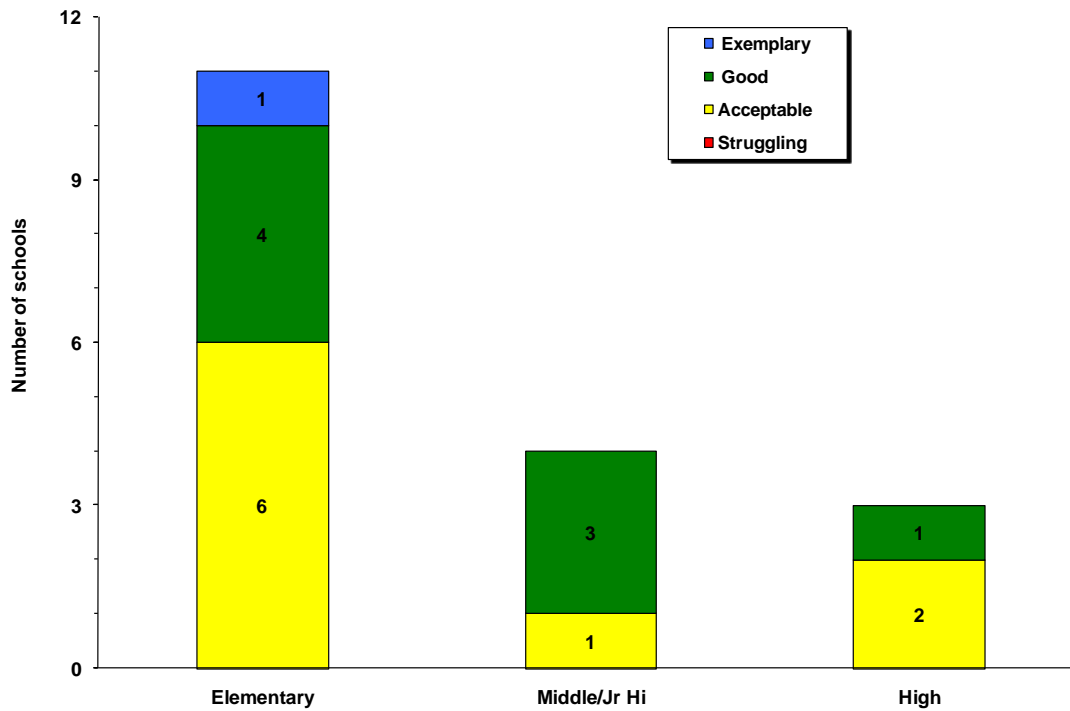


Figure 4: Distribution of *Students* by School Tiers and Grade Level in “Actual” District

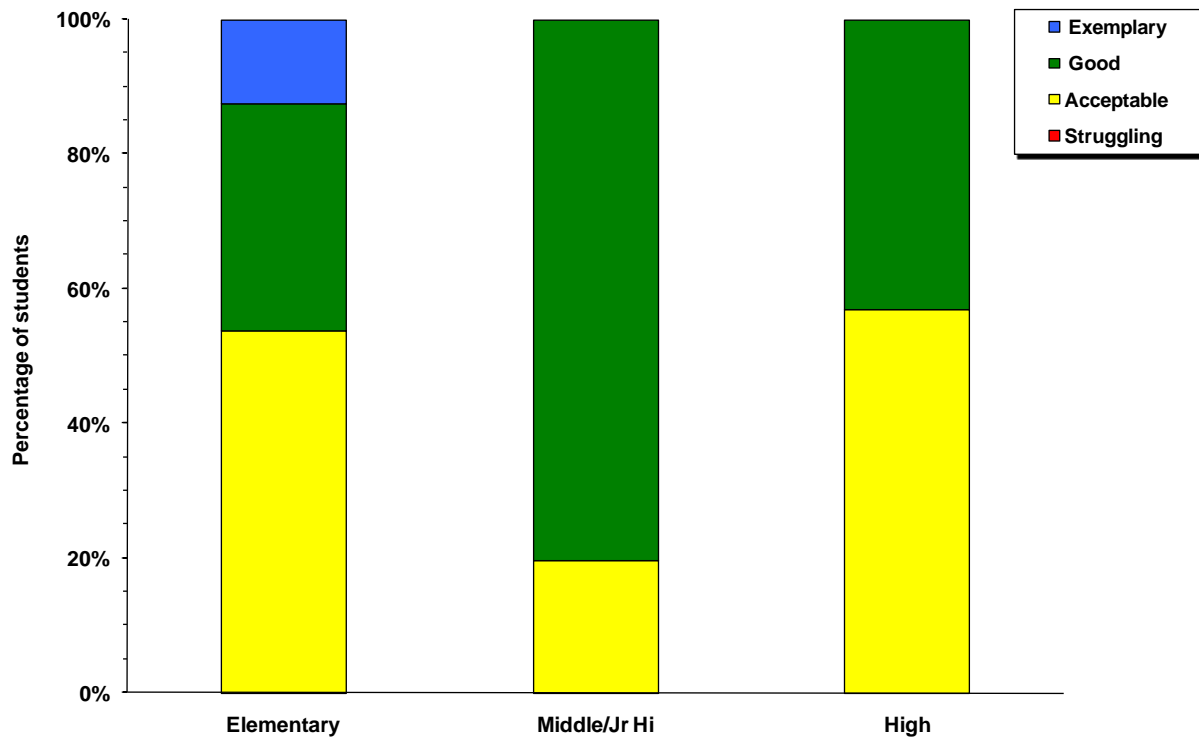


Table 7 shows the district results using the same criteria and rating system used for schools. Districts are more tightly clustered in the distribution than schools, with fewer districts in the top and bottom tiers (see Figure 5).² Figure 6 provides the distribution of all the district index results. Of the 16 districts in the “struggling” tier, the average size was about 1,000 students (the median was slightly more than 400 students). Half of the 16 districts made AYP in part because the AYP targets were relatively low in 2007, the margin of error is large for small districts, and many of the student groups in the smaller districts had fewer students than the required minimum to make an AYP determination. Figure 7 shows the results by county.

Table 7: Tier Ranges and 2007 Results for Districts (N=296)

Tier/Suggested Name	Index Range	Number of Districts	Percent of Districts
Exemplary	3.00 – 4.00	3	1.0%
Good	2.00 – 2.99	102	34.5%
Acceptable	1.00 – 1.99	175	59.1%
Struggling	0.00 – 0.99	16 ¹	5.4%
Priority (eligible for Innovation Zone)	0.00 – 0.99	TBD	TBD

¹About 16,100 students were enrolled in these districts (less than 2% of all students statewide). Only five districts had a 2-year average index below 1.00. These five districts had a total enrollment of 691 students (an average of 138 students).

² District results do not include correctional institutions, tribal schools, contract schools, and schools serving more than 50% of students outside the district boundary. The aggregation rules using in these calculations are the same as those used by OSPI when calculating district results. Results would not be published when the combined number of students assessed is less than 10.

Figure 5: Distribution of Districts by Index Score

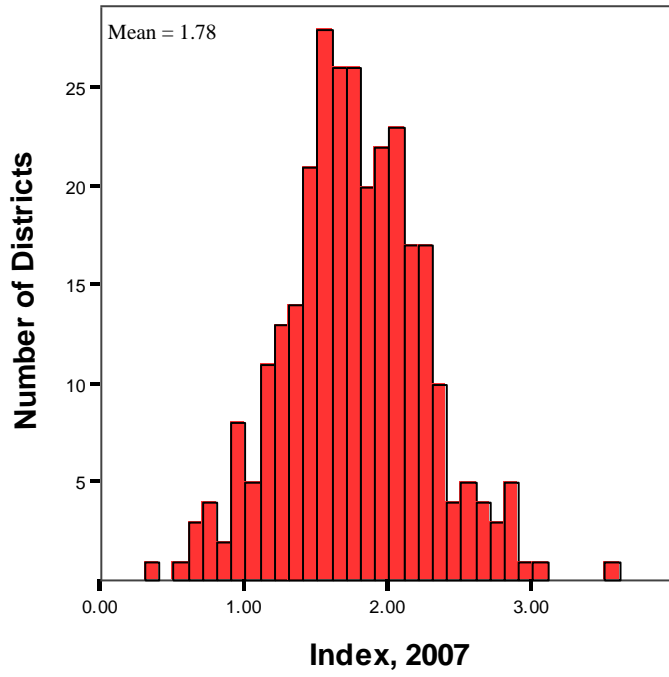


Figure 6: Distribution of Index Score by District, 2007

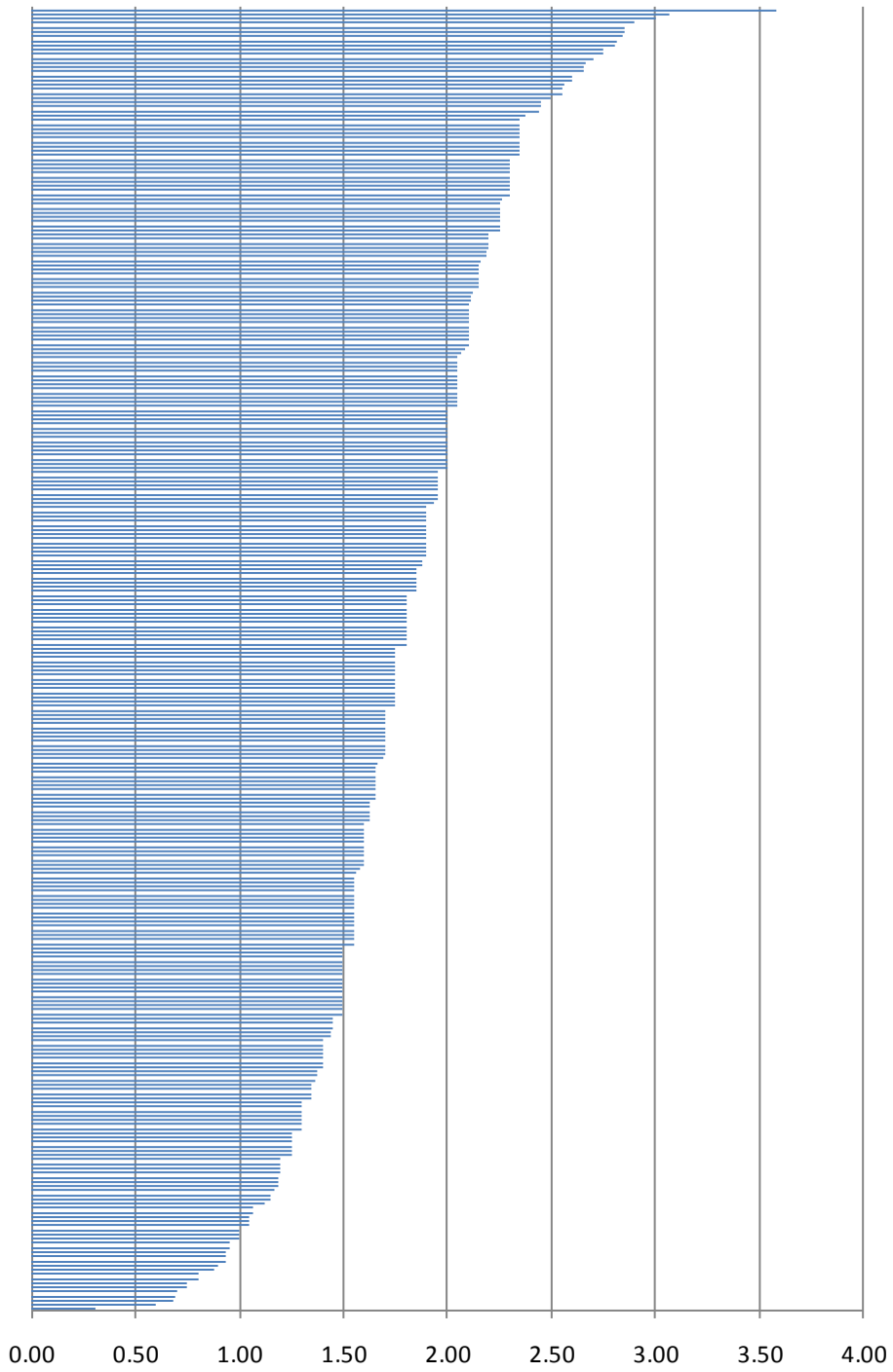
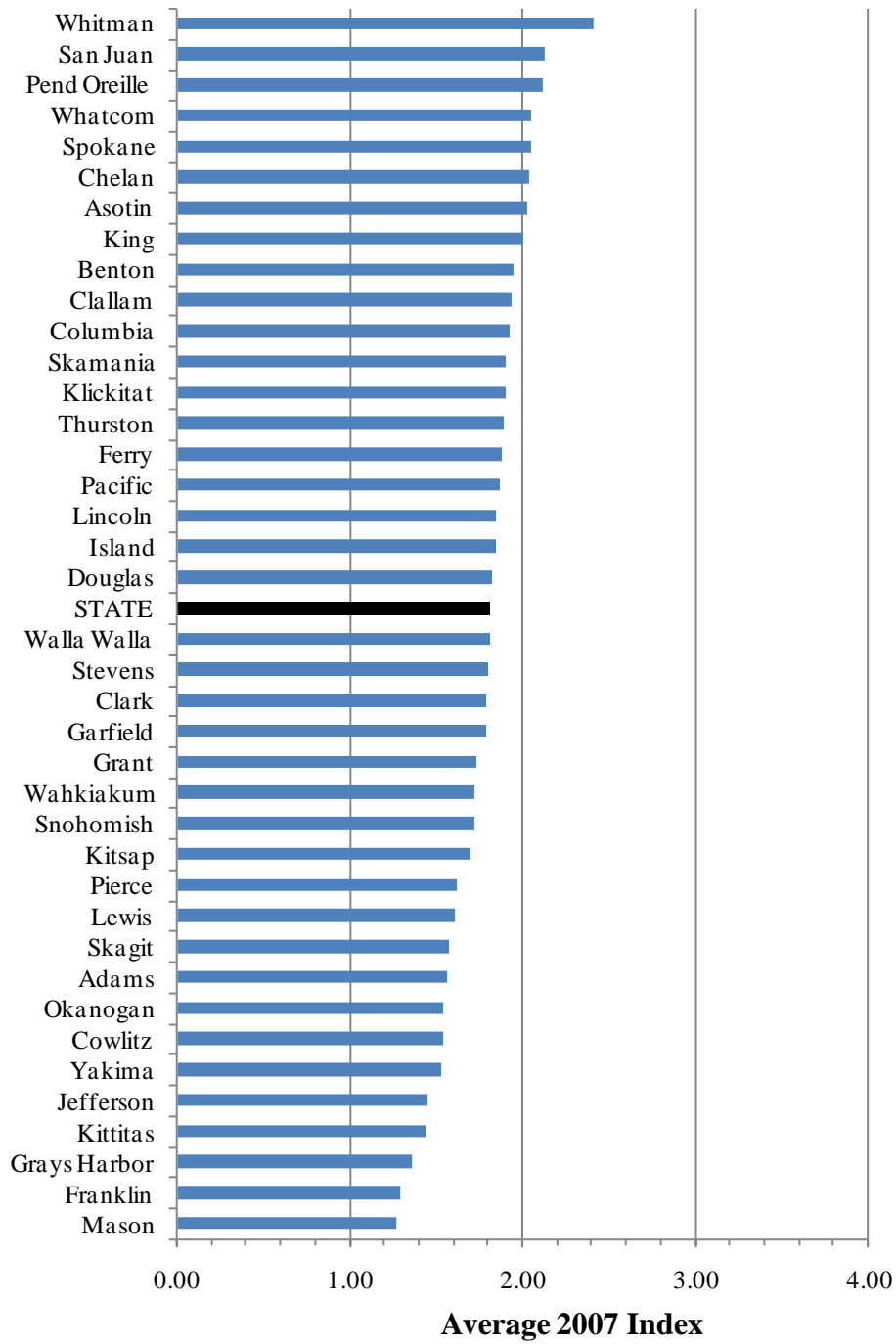


Figure 7: Distribution of Index Results by County, 2007



IDENTIFYING PRIORITY SCHOOLS AND DISTRICTS (LOWEST TIER)

Each year, the process for identifying Priority schools and districts will begin when OSPI computes the index in mid-August 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,³ 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. When OSPI and SBE staff are confident the index has been calculated correctly, OSPI staff will review the index results for each school and district that falls in the “struggling” tier, and then sort them into two categories:

- (1) Schools/districts that will *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 further analysis.

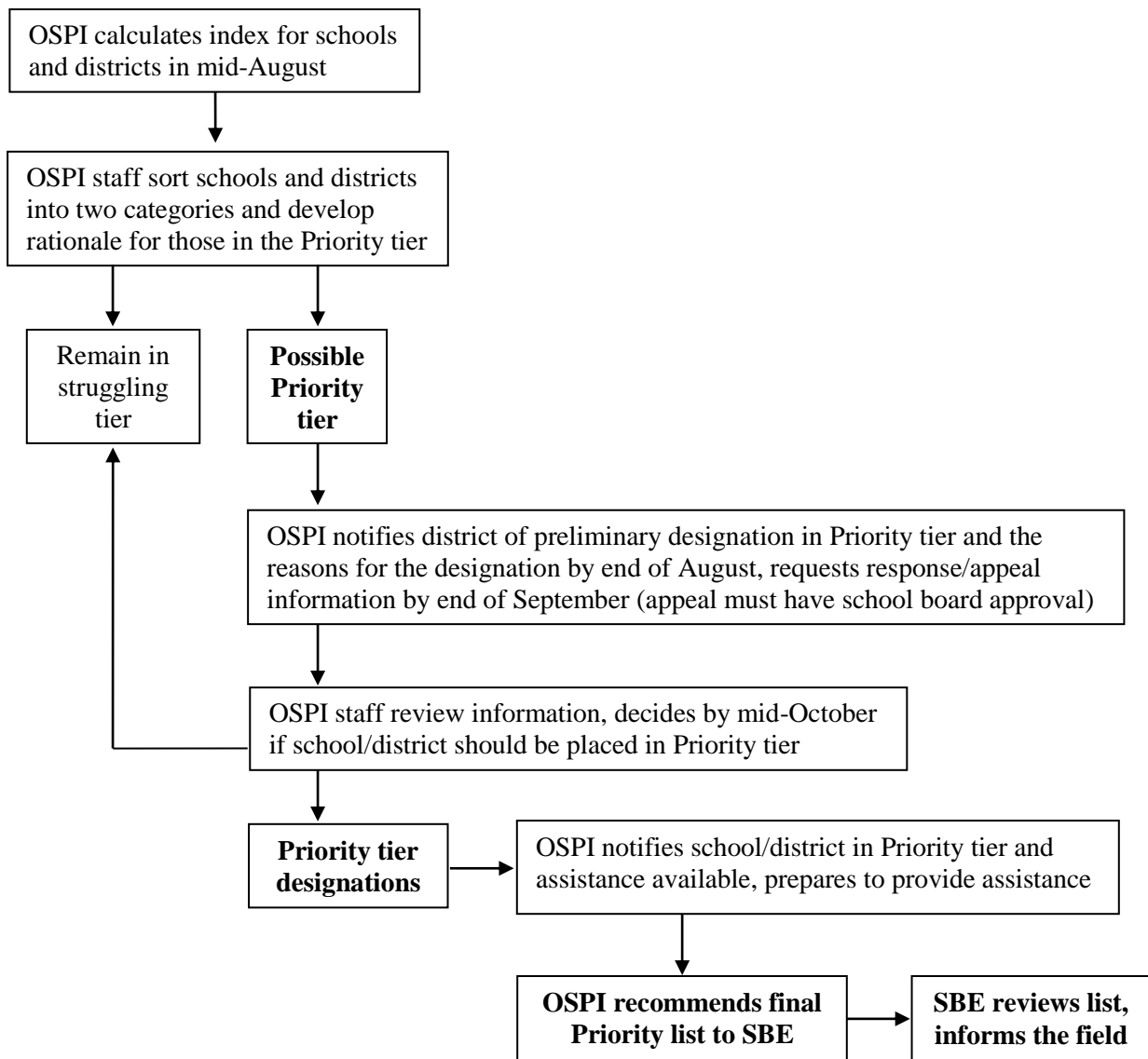
OSPI staff will conduct a deeper analysis using available data for the schools and districts placed in the *possible Priority tier* category. This may require contacting the district and/or local ESD to get more information. A comprehensive list of quantitative and qualitative data was developed that could be used to help determine which schools in the “struggling” tier will fall into the “Priority schools” tier (see Appendix B). Given the comprehensive nature of the list and the limited capacity to analyze all the data for every school and district in the “struggling” tier, the list was refined to determine which factors were the most important to analyze. The data that will be initially reviewed at this exploratory phase fall into four general areas:

- **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

³ The number will still be far fewer than those not making AYP or identified for “improvement” under NCLB.

Based on this review, the schools and districts will be sorted again into the same two categories—those that remain in the struggling tier and those in the *possible Priority tier*. By the end of August, districts and schools placed in the possible Priority tier are notified of the possible designation and the reasons why this designation is possible. If required by federal law, this initial list will be made public. During the month of September, the district/school is given a chance to avoid the Priority designation by providing more information that would explain the low index results, and it could provide more favorable results (e.g., feeder school information, results of district assessments, personnel changes, type of interventions made to date) and any plans being made for the future. Any appeal to OSPI will need to have local school board approval. OSPI will review the additional information, and by mid-October, it recommends to the State Board of Education the schools and districts that should be placed in the Priority tier. The State Board will review the list, receive comments, finalize the list, and inform the priority schools and districts about how they can respond to the designation. Figure 8 provides a flow chart of this process.

Figure 8: Process for Identifying Priority Schools and Districts



INTEGRATING THE SYSTEMS

Federal law requires states to have a single accountability system. Many states combine their state accountability system with the federal system described by NCLB. The details for integrating the federal and state system must still be determined. Washington can pursue two options to meet this requirement.

1. The preferred approach is to request that the proposed system, once adopted, be used in place of the current system. A new administration may provide more flexibility to states to design alternative accountability systems and approve them if they meet certain requirements. 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 from the accountability matrix will still be used when determining the type and level of assistance the state provides. Those that fall into “improvement” status under AYP will still face the federally required sanctions. However, schools that do not make AYP and fall into school improvement may achieve relatively favorable index results. In these cases, the amount of extra assistance the state provides will be minimal. On the other hand, some schools will make AYP and not be in school improvement, but they may have relatively low index results. In these cases, available state funds can be used to focus assistance in areas of greatest need. Regardless of the results from the two systems, the state will clarify what happens when schools and districts fall into the various AYP categories and state tiers and make every effort to minimize confusion that could occur about the two ways for measuring accountability. Appendix C provides an overview of the current assistance system being used by OSPI to help schools and districts that are in “improvement” status.

RECOGNITION

Two guiding principles apply to recognition system – it should provide multiple ways to reward success and rely mainly on criterion-referenced measures. The proposed recognition system is consistent with these principles and is based on a “theory of change” that people are motivated more by success than by blame or guilt. Recognition is based on results from the accountability matrix. The State Board should consider at least three options: provide recognition in each of the 30 cells of the matrix, or in each of the 20 “inner” cells of the matrix, or in the 10 “average” cells. The advisors recommend providing recognition in all 30 cells. The recommended minimum ratings are 3.00 for the 20 “inner” cells and 2.75 for the “averaged” cells (see Table 8). Any cell with a 3.5 or above will 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. This ensures that recognition is given for sustained exemplary performance.

Table 8: Recommended Minimum Requirements for Recognition

Indicator	Reading	Writing	Math	Science	Ext. 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

Figure 9 shows how many of the 2,046 schools would have received recognition if the proposed system were in place in 2007 and all 30 cells were able to receive recognition. The largest number of schools would have received recognition in just one or two of the 30 areas, and 330 schools (16%) would not have received any recognition. At the other extreme, 291 schools (14%) 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. Of the 209 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 9: Number of Schools with Recognition, by Number of Recognitions (2007)

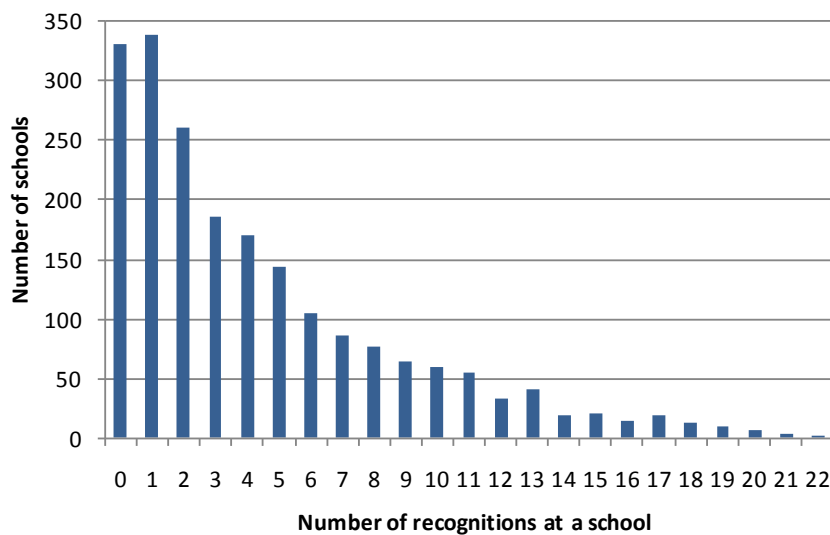
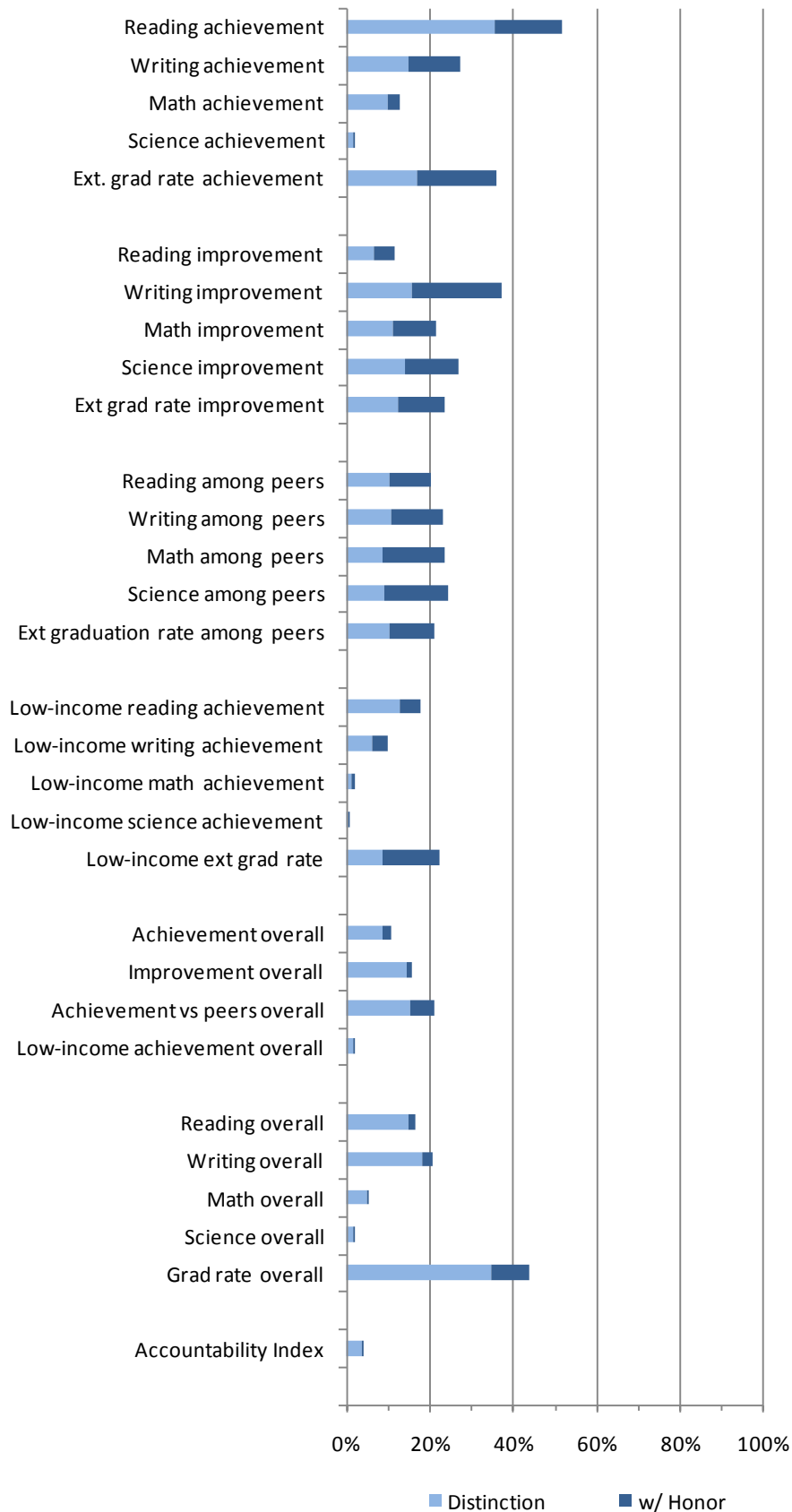


Figure 10 shows the percentage of all schools that met the recognition criteria in each of the 30 cells in 2007. The largest number of schools (52%) met the criteria for reading achievement. Achievement in math, science, and among low-income students had fewer schools meeting the criteria. Only 4% had an overall average of 2.75 on the accountability index over the 2-year period. Although schools would have received recognition in a total of 9,082 areas, this represents less than 15% of the maximum number of areas (30 cells x 2,046 schools).

Figure 10: Percentage of Schools with Recognition, 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 types of awards, primarily for schools receiving Title I funds. OSPI also provides awards but no funding as part of the recognition (Appendix A provides more information on these awards). 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 has recommended that schoolwide financial rewards be given each year when a school reaches a certain level of improvement. 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 (15% of the *districts* averaged 3.0 or better in the improvement cells). The amount of the bonus suggested by OSPI was \$20 to \$50 per student FTE. Other types of recognition, with or without financial awards, could be developed. These could be available to all that meet certain criteria and/or be competitive in nature.

* * * * *

The proposed accountability system needs to be flexible. Changes in NCLB requirements, graduation requirements, the graduation rate formulas, the assessment system, and content standards may have an impact on some measures, which may require changes to the system. As data systems improve statewide and more information becomes available, other indicators could be added to the system and other more sophisticated analyses could be used (e.g., growth models). These changes could be in the form of additional columns in the matrix (e.g., college eligible rates) or additional factors outside the matrix that could be included when calculating the index (e.g., funding amount of local levies).

A number of issues must still be resolved before the index can be implemented effectively. Further review of the cut points and results generated by those cut points 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.

Appendix A provides more details about how the index is calculated. Appendix B provides a list of possible data that could be used to identify Priority schools. Appendix C gives an overview of the current state assistance system that is funded primarily by the federal government. Appendix D lists the names of those who provided advice and feedback during the development of this proposal.

APPENDIX A

INDICATORS AND OUTCOMES

This appendix provides more detailed information about the proposed accountability index. It also includes information about how the indicators and outcomes were selected and how the index number is calculated.

SELECTION OF INDICATORS AND OUTCOMES

One of the guiding principles for the accountability system is the use of multiple measures. The advisors (see Appendix D) recommended using four indicators and five outcomes, resulting in a 4x5 matrix with 20 measures. Other indicators and outcomes were discussed besides the WASL and graduation rates, and the advisors wanted to include other outcome data in order to have multiple measures. However, no other reliable and accurate data are available statewide that is collected in a consistent manner. Moreover, using more indicators (e.g., results for separate student groups such as ELL or each race/ethnic group) would make the system much more complicated.

The index is achieved by using the simple average of the ratings across the 20 outcomes. The graduation rate is not applicable for elementary and middle schools, but these types of schools have multiple grades with WASL results that generate the ratings. By using averages, schools without data for some indicators are still included in the system and a separate system is not needed for different types of schools to generate the index.

The advisors preferred a system that uses fixed criteria rather than norm-referenced measures in order to keep the measures simple and to avoid changing goals over time and the use of measures (e.g., standard deviations) that vary by subject. This means that recognition would be given when schools meet certain criteria, and there would not be a limit to how many schools can be recognized (unlike the Schools of Distinction which only recognizes the top 5% based on improvement). With fixed criteria in place, a school and district would know in advance what it needed to do to receive recognition, regardless of how others perform. It would also encourage cooperation among educators because they would not be in competition with one another for recognition.

The advisors discussed other types of analyses that could provide more accurate results (e.g., hierarchical linear modeling, value-added growth models). However, these methods were not selected because they lack transparency, are overly complex, and are not calculated easily at the school and district levels due to capacity and software limitations.

All stakeholder groups believed the federal AYP system is not a valid way to identify schools and districts for recognition and additional support. The advisors felt the current system is too complex, has too many adjustments, and is neither transparent nor fair in its accountability determinations. For example, AYP has different goals for reading and math at three different grade levels, the goals change over time, performance is adjusted with margins of error, some students are not counted (those enrolled after October 1), and schools and districts have different minimum numbers (N) for counting the results. Moreover, AYP is almost entirely punitive in nature and has unrealistic goals. Schools must meet up to 37

goals, and districts must meet as many as 111 goals. Not meeting just one goal leads to negative consequences and labeling. The consequences are the same regardless of how many goals are missed and by how much. If a school “needs improvement,” students in groups that meet the goals must be allowed to transfer to another school, with transportation costs paid by the district. This can reduce the school’s academic performance even further. In addition, AYP does not include two subjects (writing and science) that are assessed in a standardized manner statewide, which has resulted in a narrowing of the curriculum. Finally, AYP’s narrow emphasis on students who meet standard has often resulted in more focused help being given to students that perform near that cut point (sometimes called the “bubble kids”) at the expense of students who are farther above and below that level of performance.

The proposed system is preferred because it is more inclusive and less complex than the federal AYP system. The ratings are based on the results for all students, including those who are not “continuously enrolled” since October 1. No margin of error is used, and the minimum N is 10 across the entire school/district (rather than a grade). This increases the chance that very small schools and districts (e.g., those with less than 10 students in a grade) are included in the accountability system. For example, a K-6 school that has only 4 students in each tested grade (grades 3-6) would have a total of 16 students with assessment results and would therefore be included in the system. (Grade-level results are not reported when there are fewer than 10 students in a grade in order to keep the results confidential.) Grade configurations are not an issue when calculating the results because the same benchmarks are used for each grade and subject. (AYP uses grade bands of 3-5, 6-8, and 10 with separate benchmarks and results generated for each grade band, regardless of the school’s grade configuration. The calculations to compute school results can become very complex and therefore lack transparency.) The current AYP system for holding *districts* accountable is even more complex than the school accountability system. It has different rules and sometimes produces results that are confusing and at odds with its school-level results (e.g., a district might not make AYP but all its schools do and vice versa). A district’s size is the major determinant in its AYP results—only two districts with fewer than 1,000 students are in improvement status. The proposed district accountability system is essentially the same as the system for schools, which makes it relatively easier to understand and compute.

USING THE INDEX

The results from the 20 ratings create an index number for each school and district based on the average rating. Schools and districts are assigned to a “tier” based on their index number.

- Those with the highest index numbers, from 3.00 to 4.00, are in the “exemplary” tier.
- Those with an index of 2.00 to 2.99 are in the “good” tier.
- Those with an index of 1.00 to 1.99 are in the “acceptable” tier.
- Those with an index below 1.00 are in the “struggling” tier.

Schools should not be compared and judgments should not be made about school quality based solely on their overall index score. Even though the index uses multiple measures, some schools have missing data that can affect their index number. Moreover, schools that administer assessments with lower scores overall (e.g., science and math) will tend to have a lower index score than those that do not. For example, schools serving grades 5, 8, and 10 give the science WASL, and these results tend to be very low compared to the other subjects. So a K-4 school will likely have a higher index score than a K-5 or K-8 school. As a result,

the index is only comparable across schools that serve the same grades. In addition, the index does not reflect how close a school may be to the benchmarks—small differences in results could still generate different ratings (e.g., 85%=3 and 86%=4). Moreover, schools serving very few students may have more volatile ratings from year to year. Finally, the lack of vertical alignment of the assessments presents another complicating factor when making comparisons across schools that serve different grade levels.

Given the different types of schools being rating, school results should be reported for similar types of schools. The six suggested categories for reporting the results are as follows:

- *elementary schools* (those serving from kindergarten up to grade 6)
- *middle/junior high schools* (those serving only 6,7 or 8)
- *high schools* (grades 9 or 10 to 12)
- *comprehensive schools* (e.g., K-8, K-12)
- *schools serving special populations* (alternative schools, correctional facilities, those primarily serving ELL students and those with disabilities, private schools on contract)
- *small schools* (those which have their results suppressed because the y have fewer than 10 assessed students).

Many districts have only one school, so the district and school index, tier, and recognition would be the same. This has implications for how the state structures the consequences of the accountability system (either with assistance or recognition).

The accountability system will need to remain flexible. Changes in NCLB requirements (e.g., number of tested grades), graduation requirements, the method for calculating the graduation rates, the assessment system (e.g., moving to end-of-course exams in math, adjustments to cut scores), and content standards (e.g., science) may have an impact on some measures, which may require adjustments to the accountability system. Moreover, as data systems improve statewide and more information becomes available, other indicators can be added to the system⁴ and other more sophisticated analyses could be used (e.g., growth models). Other measures of improvement could be used (computing expected change, percent increases). Changes could also be in the form of additional columns in the matrix (e.g., college eligible rates) or additional factors outside the matrix that could be included when calculating the index or peer results (e.g., funding amount of local levies).

CALCULATION METHODS

To calculate the assessment-related measures, student-level data were used and aggregated to the school and districts levels. This provides more accurate results than using aggregated school and district results. Moreover, using student-level data allows for the aggregation of results from the grade level that would be suppressed because the number of students assessed was less than 10. Results are only suppressed when there are fewer than 10 students

⁴ Most of the other outcomes relate to high schools and the transition to higher education. Some data require transcript information, such as AP enrollment, dual enrollment, and college-ready rates. Other data sources could provide information about college entrance exams, college going rates, and remediation rates in higher education institutions.

assessed in the combined grades.⁵ Students who took the alternate assessments (WAAS) were included in the calculations, as were students who previously passed (this relates mainly to high school students that met standard while in grade 9, but it also applies to students that are retained). Students who met standard in a previous year did not have their level included in the student-level database, so they were considered to have performed at Level 3. Students who were exempted from taking the assessments (i.e., those with excused absences and medical exemptions, first-year ELL students, home-based and private school students) were not included in the calculations.

When computing the index, all the ratings are counted equally (i.e., they are not weighted). Averages are computed only for cells of the matrix that had data (e.g., an elementary school has no graduation data, so the averages for the indicators used only the assessment results). District results are based on OSPI's aggregation rules, so the district results do not include results from correctional institutions, tribal schools, private schools or agencies providing services, vocational schools/skill centers, schools that enroll more than 50% of their students from another district, and schools operated by a college or university that are not affiliated with a district. Finally, the results are those for a single year rather than averages over multiple years for simplicity and to avoid the distortions when change takes place over time (e.g., when averaging, schools that have dramatic declines have better outcomes and schools with dramatic increases have worse outcomes).⁶

ACHIEVEMENT INDICATOR

This indicator has five outcomes: the four subjects tested by the WASL/WAAS statewide (reading, writing, math, and science) and the extended graduation rate (see explanation on how the rate is calculated below). The measure reflects the percentage of "all" students meeting standard. Unlike the AYP measure, this indicator is what is shown on OSPI's online Report Card and does not reflect any adjustments (i.e., margin of error, continuous enrollment). The percent meeting standard includes both the results of the WASL and the WAAS, which is given to students with disabilities. These results are the combined total of the WASL and WAAS results found on the Report Card and are used when calculating AYP (without the margin of error and including students not continuously enrolled). For grade 10, only the first grade 10 attempt as reported in June of the tested year is used (this includes results for students who met standard in grade 9). Results from August assessments and retakes will be considered when looking at the "struggling" schools and districts to determine if they should be included in the Priority tier. This will recognize the districts that go to extra effort to help students who are in danger of not graduating unless they pass the required assessments. Subgroups results (for the various race/ethnicity groups, ELL, students with disabilities, gender) are used when examining the "struggling" schools and districts to determine if they should be included in the Priority tier. Results for low-income students are used in aggregate in a separate indicator described below.

⁵ Very small schools (those with fewer than 10 assessed students) will have their index calculated but it will not be made public. However, the index will be viewed by state officials, and if the index is in the struggling tier on a consistent basis, the school could be placed in the Priority tier.

⁶ In small schools, a single student could cause large changes in the index from year to year. However, analyses found relatively little difference in the amount of change in small schools compared to larger schools from one year to the next.

The Achievement benchmarks and ratings for each of the four assessed subjects and the extended graduation rate are as follows:

- Achievement on *assessments* is rated based on the following percentage of students meeting standard:

86-100%	4
70-85.9%	3
55-69.9%	2
40-54.9%	1
< 40%	0

- Achievement on the *graduation rate* is rated on the extended graduation rate from the previous year (see below for more information on how the graduation rate is calculated):

≥ 95%	4
85-94.9%	3
75-84.9%	2
65-74.9%	1
< 65%	0

Students from all tested grades in a school are combined for each subject, and the percentage of these students that meet standard on their respective tests is the school's percent meeting standard for that subject. This means the index can be calculated easily, regardless of a school's grade configuration (although grade configurations influence the results due to differences in the tests given). The same scoring benchmarks are used for all subjects. This gives equal importance to each subject.⁷ It also encourages the vertical alignment of the state assessments.

A school/district must have at least 10 students for it to be included in the accountability system. The minimum number used by OSPI is 10, but this policy is applied at the test and grade level. Using an N of 10 for a *school* means that very small schools will now be included in the accountability system because they will likely have at least 10 students assessed across the entire school. Combining all the test results together and using an N at the school level increases the overall N so a single student in a small school has less impact on the results and causes less of a change in the results from year to year. By using this system, scores in many schools that are currently suppressed at the grade level when there is less than 10 students assessed will become known in their aggregate form. This N policy means the state accountability system is more inclusive than the current AYP system, where the N is 30

⁷ The advisors did not have consensus about how to include science results in the index. Some felt that science should not be included at all because of changing standards and that it is not being taken seriously in many cases, which results in low scores across the state and relatively little improvement over time. As a result, it has little ability to differentiate school performance. Some suggested using lower cut points and raising them over time or including science but giving it less weight. After much discussion, a majority of the advisors concluded that since science will be a graduation requirement relatively soon, the only way to have science taken seriously was to treat it like the other subjects. Keeping the same rating system as the other subjects also keeps the system consistent and less complex and provides the opportunity to receive high ratings for improvement. Moreover, science achievement affects only two of the 20 cells of the matrix. Finally, not including science with equal weight penalizes those who work hard in this subject, and it would send the wrong message about the importance of students learning science concepts.

and applies only students who are continuously enrolled. The advisors felt that the education system has a moral responsibility to serve all students, and having a small minimum N and counting students who have not been in class all year helps hold schools accountable for meeting the needs of *all* their students.

ACHIEVEMENT VS. PEERS INDICATOR

This indicator uses the Learning Index (described below) level and controls for student characteristics beyond a school's control. Scores are the difference between the school's adjusted level and the average level among the school's peers. Specifically, the school/district score is the unstandardized residuals generated by a multiple regression. Those with scores above 0 are performing better than those with the same student characteristics, and those with scores below 0 are performing below those with the same student characteristics. Separate analyses are run for elementary, middle, high, and comprehensive (e.g., K-12) schools because of the variation of the variables at each grade level. Schools serving specialized student populations (e.g., alternative schools, ELL and special education centers, private schools on contract, institutions) are not included in the regressions. Excluding these schools provides a better predicted level for the remaining regular schools in the analysis and better data for use when determining the cut scores for the various ratings. Since the specialized schools have such different characteristics, results for this indicator are not computed and their index is based on an average of their remaining ratings.

For schools, four student characteristics are the independent variables in the multiple regression: the percentage of (1) low-income students (percent eligible for free or reduced-price lunch⁸), (2) English language learners, (3) students with disabilities, and (4) mobile students (not continuously enrolled). A school's Learning Index from each of the four assessments (using WASL and WAAS results) as well as the extended graduation rate for high schools are the dependent variables. The regressions are weighted by the number of students assessed to prevent a small "outlier" school from distorting the regression (predicted) line. Although there is a high correlation between all the independent variables except special education, the regressions showed that all four variables helped improve the quality of the predicted levels, regardless of the regression method used.

For districts, three of the four student characteristics used in the school analysis were the independent variables in the multiple regression: the percentage of (1) low-income students (percent eligible for free or reduced-price lunch), (2) students with disabilities, and (3) mobile students (not continuously enrolled in a school). The percentage of English language learners was not used because the initial analyses using this variable did not provide meaningful results. The same five dependent variables from the school-level analyses were used in the district analyses (the Learning Index for the four subjects and the extended graduation rate).

Financial information was also used as an independent variable in the district analysis. This information is only available at the district level, and some communities are able to raise higher levels of funding. The financial variable used is the total amount of operating expenditures per weighted pupil (funding used for capital purposes is not included). This

⁸ The percentage of students in high schools who are eligible is often higher than what is reported, but this proxy for socioeconomic status is still the best available.

variable controls for the level of funds spent in the district. Weighting the student count “inflates” the enrollment figure because certain students require more resources to educate. The extra weight for ELL and low-income students was .20, which is the typical amount used in school finance studies (although the actual number is likely to be much higher). The weight for students with disabilities was .93, which is consistent with both the national research and the level of funding provided by the state.

- Achievement vs. Peers on the *assessments* is rated based on the difference between the actual and predicted Learning Index levels:

> .204
.10 to .203
-.099 to .0992
-.20 to -.101
< -.200

- Achievement vs. Peers on the *extended graduation rate* is rated based on the difference between the actual and predicted extended graduation rate:

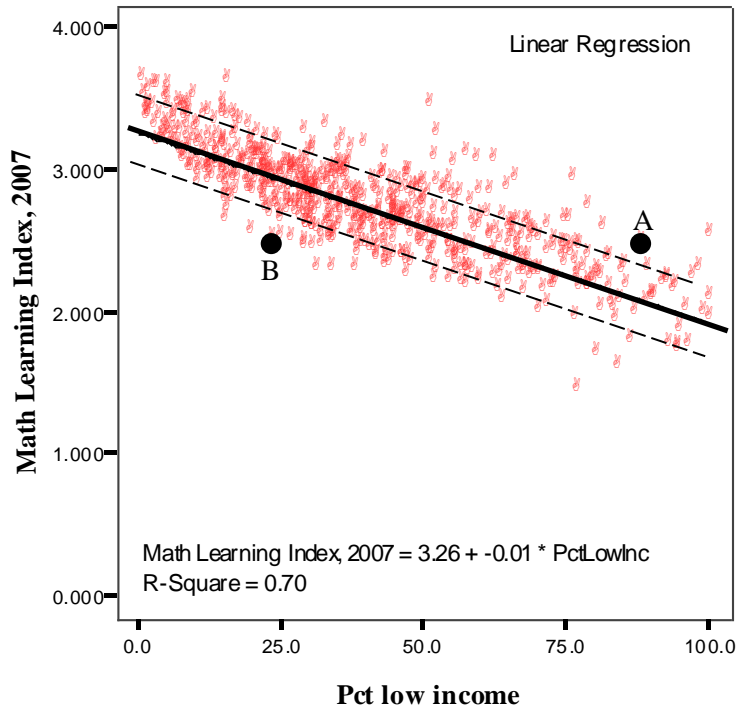
> 124
5.01 to 123
-5 to 52
-5.01 to -121
< -120

The mobility measure may need to be refined after further discussion takes place. Currently there is no common definition of mobility, and migrant student data does not include many students who are mobile. OSPI’s student data system includes information about students who are/are not continuously enrolled from October 1 through the end of the testing period in May as part of the AYP system. Using this measure, the average state mobility rate is less than 6%. Most schools with mobility rates above 15% are alternative schools, and very few districts (mainly those in Pierce County close to military bases) have many of their schools with this high of a rate. However, the proposed measure may not identify students who move in and out of a school or district multiple times during the school year and are considered continuously enrolled, nor does it identify students that are new to the district and are still enrolled during the entire year. The proposed measure, the percentage of non-continuously enrolled students, can be used until a better measure is identified.

The scatterplot in Figure 9 illustrates how this indicator works. It shows just one of the independent variables (percent low income students) in relation to one outcome (K-6 math results). Each dot represents a school. The dark line is the average (predicted) level for a given Learning Index and low-income percentage. The distance between the school and the line is the difference from the predicted level. In this example, schools A and B have almost identical Learning Index results, but A falls well above the line while B falls well below the line. The dashed lines running parallel to the trend line represent the highest and lowest cut points used for the ratings (.20 above and .20 below the trend line). When this kind of analysis is done factoring in the other variables (ELL, special education, mobility) at the same time in a multiple regression calculation, the distance from the predicted line is the school’s score, which produces a rating. If the low-income variable was the only one used in

the analysis, School A would have a rating of 4 because its index is more than .20 points above its predicted level, while school B would have a rating of 0 because its index falls more than .20 points below the predicted level.

Figure 9: Scatterplot of Math Results in Elementary Schools by Percent Low Income



The advisors discussed other possible independent variables that could be included in the analysis. These include the percentage of students who are enrolled in a gifted program, the percentage of minority students, and school size (enrollment).

- A gifted variable was not included because of a lack of reliable data, although the system should somehow take into account when a school has concentrations of these students. These schools will likely have very high achievement ratings.
- A race/ethnicity variable was not included because it is highly correlated with the other variables. Statistical analyses that included this variable found it added very little to the explanatory power of the model. Moreover, using this variable would reduce our ability to identify schools where students of color are treated differently. Finally, many of these students are also from low-income families, which is a separate indicator.
- A school size variable was not included because research findings to date reveal mixed results about how school enrollment levels affect student outcomes. School size is also a factor that can be controlled somewhat at the district level through the use of specialized programs and boundary lines. Other methods can be used to help schools compare themselves to those with similar sizes once the accountability results are made known.

The **Learning Index** is the dependent variable used for this indicator and for the Improvement indicator described below. This index, which was developed by the Commission on Student Learning and refined by the A+ Commission,⁹ takes into account the

⁹ These Commissions are no longer in existence.

percentage of students performing at the different WASL levels. Specifically, the WASL and WAAS tests have five levels of performance:

- Level 0 – No score given¹⁰
- Level 1 – Well below standard
- Level 2 – Partially meets standard
- Level 3 – Meets standard
- Level 4 – Exceeds standard

This index is calculated like a grade point average with 4.0 as the highest score and reflects the level of student performance across the entire range of proficiency, not just those meeting standard. It gives greater weight to higher levels of proficiency on the state assessments and provides an incentive to support the learning of all students, including those well below standard (Level 1) and those that already meet the standard (Level 3) so they can move up to the next level. There is a “ceiling effect” when using this measure, but preliminary results show that even high-performing schools were achieving large gains because of the movement of students from Level 3 to Level 4. Once a school has all of its students in Level 4, there would not be any possibility to improvement any more, but all ratings together would still result in a school being in highest tier.

The following example shows how the Learning Index is calculated. The same method is used to calculate the index for all WASL tests (reading, mathematics, writing, science) in all the tested grades:

- Level 0: 5% of all students assessed
- Level 1: 15% of all students assessed
- Level 2: 20% of all students assessed
- Level 3: 40% of all students assessed
- Level 4: 20% of all students assessed

$$\begin{aligned} \text{Learning Index} &= (0*0.05) + (1*0.15) + (2*0.20) + (3*0.40) + (4*0.20) \\ &= 0 + .15 + .40 + 1.20 + .80 = 2.55 \end{aligned}$$

IMPROVEMENT INDICATOR

The Improvement indicator relies on changes in the Learning Index for the four assessed subjects and the graduation rate from one year to the next. Specifically:

- Improvement on *assessments* is rated on the levels of annual change in the Learning Index:

> .12	4
.051 to .12	3
-.05 to .05	2
-.051 to -.12.....	1
< -.12	0

¹⁰ The “No Score” designation includes unexcused absences, refusals to take the test, no test booklets but enrolled, incomplete tests, invalidations, and out-of-grade level tests.

- Improvement on *graduation rate* is rated on the level of percentage point change in the extended graduation rate from the previous year (see below for more information on how the graduation rate is calculated):

> 6	4
3.01 to 6.00	3
-3.00 to 3.00	2
-6.00 to -3.01	1
< -6	0

A one-year change is used rather than using averages of previous years or a change from a year further in the past because it is the simplest calculation, it reflects the most recent set of results, and it does not distort the most recent results (using a two-year average helps a school if scores go down and penalizes the school if scores go up). New schools would only need two years of data to generate an improvement score. Since results are created each year, changes over time are seen when examining the results across multiple years.

The advisors discussed other possible improvement measures, including a 10% reduction in those not meeting standard (the AYP “safe harbor” measure), a 25% reduction in those not meeting standard over a 3-year period (the goal used for grade 4 reading several years ago), a percentage point gain from the previous year (or over several years), and a change in the scale score. While each of these have merit, the advisors determined that the annual change in the Learning Index provided the best measure of improvement because it focused on more than just those meeting standard and uses available data. The other measures can be used when analyzing “struggling” schools and districts for possible designation in the Priority tier.

ACHIEVEMENT OF STUDENTS FROM LOW-INCOME FAMILIES

Much research has shown that student achievement is highly correlated with a family’s socioeconomic status (SES). Specifically, academic achievement among students who live in a low-income family is usually far below students from families that are not considered low income. This indicator focuses on the performance of low-income students. It uses the same five outcomes as the Achievement indicator: the four subjects tested by the WASL/WAAS statewide (reading, writing, math, and science) and the extended graduation rate. However, the outcome measures are the percentages of assessed students who are from low-income families who meet standard on the assessments and who graduate by the age of 21. The same rating scales are used as the achievement indicator.

A low-income student is one who is eligible to receive a federally-subsidized meal (e.g., free or reduced-price lunch). The percentage of students in high schools who are eligible is often higher than what is reported, but this measure is still the best available proxy for SES. This indicator is highly correlated with the percentage of ELL students and students of color, two groups of students that often have lower levels of student achievement. The indicator is also positively correlated with students with disabilities and mobility.¹¹ This does not imply that a

¹¹ The statewide correlations between the percentage of students considered low-income and the percentage of students of color and ELL students in a school are .70 and .68 respectively. More than 86% of the ELL students are from low-income families. The correlations with mobility and special education are .49 and .27 respectively.

student’s socioeconomic status captures all the unique needs of students of color, students with disabilities, those learning English, or those who are mobile. These students face additional challenges in Washington schools that affect their learning.¹² Finally, the results for this indicator will not be different from the Achievement indicator if most or all of the students in a school come from low-income families.

EXTENDED GRADUATION RATE MEASURE

The Washington State definition of the on-time graduation rate is the percentage of students who graduate from public high school with a regular diploma (not including a GED or any other diploma not fully aligned with the state’s academic content standards) in the standard number of years. The period of time required for students with disabilities to graduate is specified in each individualized education program (IEP). Students with disabilities who earn a diploma by completing the requirements of an IEP in the required period of time are counted as on-time graduates. The period of time required for EL and migrant students to graduate is determined on an individual basis when they enter the district and may be longer than the standard number of years. The period of time required to graduate for a migrant student who is not LEP and does not have an IEP can be one year beyond the standard number of years. LEP and migrant students who earn a diploma in the required period of time are counted as on-time graduates.

The on-time graduation rate is calculated as follows:¹³

$$\text{On-Time Graduation Rate} = 100 * (1 - \text{grade 9 dropout rate}) * (1 - \text{grade 10 dropout rate}) * (1 - \text{grade 11 dropout rate}) * (1 - \text{grade 12 dropout rate} - \text{grade 12 continuing rate})$$

with *Dropout Rate* = $\frac{\text{number of students with a dropout, unknown, GED completer code}}{\text{total number of students served (less transfers out and juvenile detention)}}$

To encourage schools to serve students who remain in school beyond 4 years, a separate graduation rate is calculated that includes students who graduate in more than 4 years. This “**extended rate**” is used for AYP purposes and the rate used in the accountability index. The formula for calculating this rate is as follows:

$$\text{Extended Graduation Rate} = \frac{\text{number of on-time and late graduates}}{\# \text{ of on-time graduates} / \text{on-time graduation rate}}$$

Dropouts are not counted as transfers. Since graduation data are not reported until after the beginning of the school year, the rates from the previous year are used.

¹² The Center for the Improvement of Student Learning (CISL) has convened an advisory committee to develop a strategic plan to address the achievement gap for African American students, as outlined in HB 2722.

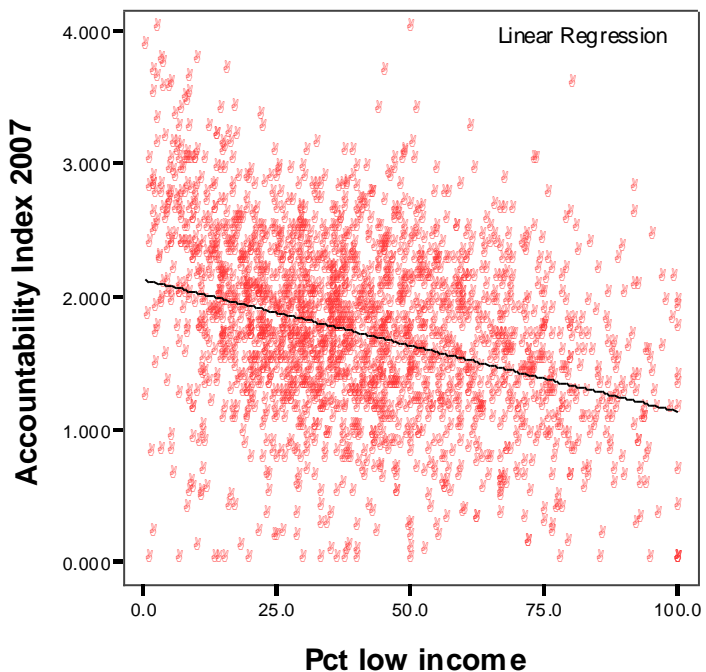
¹³ See <http://www.k12.wa.us/DataAdmin/pubdocs/GradDropout/03-04/Graduationanddropoutstatistics2003-04Final.pdf>, chapter 1, for more information about these formulas.

The calculation method may change in the future when the state has enough data to track students over the entire time period. The cut scores for determining the ratings may need to change if another method produces substantially different results.

DISTRIBUTION OF INDEX

Given the high correlation between family income and student performance, analyses were conducted to see how the school index related to schools' percentage of low-income students. Figure 10 shows these results for the 2,046 schools used in the analysis, while Figure 11 shows the results for the 296 districts. These figures show a much weaker relationship between the two variables than what would be seen if the dependent variable was achievement. Many schools and districts that have relatively few low-income students still have rather low index scores, while many that have high concentrations of low-income students have rather high index scores. The trend line is still sloping downward, but the correlations and r-squares are relatively weak (-.33 and .11 for schools, -.22 and .05 for districts). These are much weaker than the relationship between student achievement and socioeconomic status. This is because achievement represents only half the index and is moderated by two of the other variables (improvement, peers) that have low correlations with socioeconomic status (all the school correlations with the improvement and peers variables were less than $\pm .08$). It is harder for a school or district that has a high percentage of students who are low-income to achieve a very high index because the "all" students results are very similar to the low-income students results.

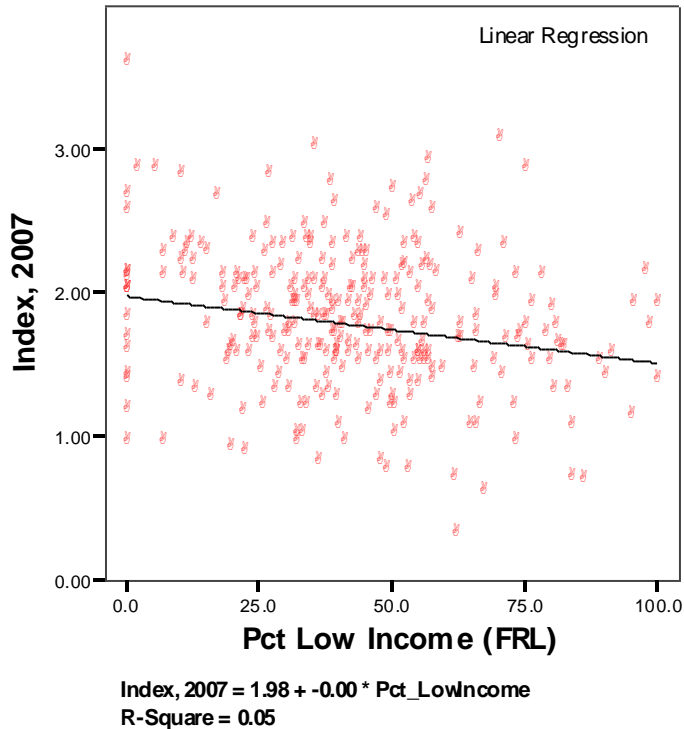
Figure 10: Scatterplot of Index for All Rated Schools, by Percent Low Income



$$\text{Accountability Index 2007 (rating average using student data)} = 2.14 + -0.01 * \text{PctLowInc}$$

R-Square = 0.11

Figure 11: Scatterplot of Index for Districts, by Percent Low Income



RECOGNITION SYSTEM

Many of the guiding principles apply to the recognition system. The system should:

- Be transparent and simple to understand;
- Rely on multiple measures;
- Encourage the improvement of student learning and cooperation among educators;
- Focus at both the school and district levels;
- Rely mainly on criterion-referenced measures; and
- Provide multiple ways to demonstrate success and earn recognition.

With these principles in mind, the same matrix that is used to generate the index is also used to identify schools and districts for recognition. Cut points were developed for all 30 cells of the matrix after looking at distributions of the ratings for all schools. (The impact of the cut points on *districts* was not calculated for this analysis. Districts have fewer high ratings, as noted in Figures 1 and 5, so they would receive recognition less often than schools). To ensure recognition does not occur based on one good year alone, two years are averaged, and the average must meet minimum criteria.

Different cut points are used for different parts of the matrix because it is harder to achieve high ratings for some cells.

- For the “inner” 20 cells of the matrix, at least a 3.0 average is needed to receive recognition. To meet this level, a school/district needs to receive at least two straight ratings of 3, which are the second highest ratings (or it could have a rating of 2 and 4 in a

2-year period). Cells that average 3.5 or better (receive ratings of 3 & 4 or a 4 & 4) would receive recognition with “honors.”

- For the 10 “averaged” cells on the outside of the matrix, at least a 2.75 is needed. This lower average is justified because it is much harder to achieve an average of 3.0 in the multiple categories. Relatively few schools and districts would be recognized even at this lower level—on average only 14% of schools reached this level in each of the 10 cells, and even fewer districts reached this level (districts do not have as many high ratings). If a 3.0 were required instead of a 2.75, only about 9% of schools, on average, would receive recognition in these cells.
 - ✓ To meet an average of 2.75 in the *five outcome categories* (assessments and graduation rate), a school/district needs to have a total of 11 points in the four indicator ratings ($11/4=2.75$). This would usually require a majority of ratings of at least a 3 in two consecutive years.
 - ✓ To meet this level in the *four indicator categories* (achievement, improvement, achievement vs. peers, low-income achievement), a school/district needs to have a total of 14 points in the five outcome ratings ($14/5=2.80$). This would usually require 4 out of 5 ratings of at least a 3 in two consecutive years.
 - ✓ Like the “inner” cells of the matrix, any “averaged” cell with a 2-year average of 3.5 or better would receive recognition with “honors.”

The number of schools and districts that receive recognition depends on the criteria described in Table 2. If the Board wanted to increase or decrease the amount of recognition provided, it could either change the criteria in Table 2 or change the cut points for recognition. Changes in the criteria in Table 2 would also affect the index scores for districts and schools. The Board could also request that a more formal “standard-setting” process take place to confirm or adjust the criteria used in Table 2.

The Board could establish additional criteria in order for a school/district to receive recognition. For example, the Board could require that recognition be given only if the achievement gap (e.g., between genders or between various groups of students) was decreasing. It could also require a closer analysis of the data before a school/district receives recognition with honors to ensure data problems (in their favor) or other factors are not responsible for very high ratings. This would prevent inappropriate designations that could undermine the accountability system.

A number of issues still need to be resolved related to the recognition. This includes what benefits accrue when a school or district meets the recognition criteria. The consequence could be as simple as highlighting the results on a Web site and issuing a press release about the winners. It could also generate financial rewards in certain cases. Another issue is what happens when a school and district are one in the same. The Board would need to make sure that any recognition is not duplicative (e.g., issuing a banner or financial reward for both the school and the district). Further, the Board could create other types of recognition, such as special recognition for a few outstanding schools/districts and some that could be competitive in nature (e.g., require nominations or applications). Finally, the proposed recognition should be integrated with existing awards being given by OSPI. It currently gives recognition through federal and state programs.

- **Federal Awards**

- ✓ *Blue Ribbon Schools* are nominated by OSPI and selected by the U.S. Department of Education based on high academic performance. In order to be selected, nominated schools must provide detailed information about their school, they can be any type of school (including private schools), and they must make AYP in the year of the nomination and the following year.
- ✓ For the *Academic Achievement Award* program, Title I Part A schools that met AYP for three consecutive years in math and/or reading can apply for recognition of improving student achievement in one or both content areas. Up to nine schools can receive an award of \$10,000. The application provides details about successful math and/or reading strategies, and these strategies are showcased at state conferences and on OSPI's website in order to assist other schools.
- ✓ For the *Distinguished Schools Award*, four Title I Part A schools are selected, two in the national category and two in the state category. Schools must apply for this award, which focuses on either exceptional student performance for two or more years or significant progress in closing the achievement gap. National award winners receive \$10,000 while state award winners receive \$5,000.

- **State Awards**

- ✓ OSPI began recognizing *Schools of Distinction* in 2007 based on improvement over an extended period of time and achievement that exceeds the state average. Only the top 5% of schools received this award.
- ✓ OSPI has been giving *Improvement Awards* since 2004 to schools and district that make at least a 10% reduction in the percentage of students not meeting standard in reading, writing, and math in grades 4, 7, and 10. Wall plaques with metal plates for updates are provided to those receiving this award. In 2007, there were 1,255 schools that received a total of 2,190 awards in the three grades and subjects; 241 districts received a total of 804 awards in the three grades and subjects. OSPI does not provide any recognition or results based on how schools or districts compare to their peers.

Table 9 provides the data used in Figure 9. It shows the number and percentage of schools that would have received recognition if the proposed system were in place in 2007. Out of the 2,046 schools, the largest number would have received recognition in just one or two of the 30 areas, and 330 schools (16%) would not have received any type of recognition. At the other extreme, 291 schools (14%) would have received recognition in 10 or more areas; 2 schools would have received recognition in 22 of the 30 cells of the matrix.

Table 10 provides the data used in Figure 10, the number and percentage of all schools that met the recognition criteria in each of the 30 areas in 2007. Reading achievement had the largest number of schools meeting the criteria; achievement in math, science, and among low-income students had far fewer schools meeting the criteria. Only 4% had an overall average of 2.75 on the accountability index over the 2-year period. Although schools would have received recognition in a total of 9,082 areas, this represents less than 15% of the maximum number of areas (30 cells x 2,046 schools). Roughly 40% of the recognitions would have been considered "with honor" based on schools averaging 3.5 or better. These "honor" recognitions represent less than 6% of the maximum number of possible areas.

Table 9: Number of Schools with Recognition, by Number of Recognitions (2007)

Number of recognitions at a school	Number of schools	Pct of all schools	Cumulative percent
0	330	16.1%	16.1%
1	338	16.5%	32.6%
2	260	12.7%	45.4%
3	185	9.0%	54.4%
4	169	8.3%	62.7%
5	143	7.0%	69.6%
6	104	5.1%	74.7%
7	85	4.2%	78.9%
8	77	3.8%	82.6%
9	64	3.1%	85.8%
10	59	2.9%	88.7%
11	55	2.7%	91.3%
12	33	1.6%	93.0%
13	41	2.0%	95.0%
14	18	0.9%	95.8%
15	20	1.0%	96.8%
16	14	0.7%	97.5%
17	18	0.9%	98.4%
18	12	0.6%	99.0%
19	10	0.5%	99.5%
20	6	0.3%	99.8%
21	3	0.1%	99.9%
22	2	0.1%	100.0%

Table 10: Distribution of Schools with Recognition, by Type of Recognition (2007)

Type of Recognition	# of “Schools of Distinction”	# of “Schools of Distinction” with Honor	Total # of schools recognized	Pct of all schools**
Reading achievement	727	330	1,057	51.7%
Writing achievement	309	255	564	27.6%
Math achievement	204	60	264	12.9%
Science achievement	37	9	46	2.2%
Ext. grad rate achievement	75	83	158	36.0%
Subtotal, Achievement¹	1,352	737	2,089	
Reading improvement	135	100	235	11.5%
Writing improvement	322	446	768	37.5%
Math improvement	230	209	439	21.5%
Science improvement	286	265	551	26.9%
Ext grad rate improvement	54	50	104	23.7%
Subtotal, Improvement¹	1,027	1,070	2,097	
Reading among peers	210	210	420	20.5%
Writing among peers	221	254	475	23.2%
Math among peers	176	312	488	23.9%
Science among peers	191	313	504	24.6%
Ext graduation rate among peers	46	46	92	21.0%
Subtotal, Peers¹	844	1,135	1,979	
Low-income reading achievement	259	105	364	17.8%
Low-income writing achievement	128	78	206	10.1%
Low-income math achievement	26	17	43	2.1%
Low-income science achievement	5	4	9	0.4%
Low-income ext grad rate	38	61	99	22.6%
Subtotal, Low Income¹	456	265	721	
Achievement overall	179	41	220	10.8%
Improvement overall	297	29	326	15.9%
Achievement vs peers overall	311	125	436	21.3%
Low-income achievement overall	30	7	37	1.8%
Reading overall	306	30	336	16.4%
Writing overall	374	48	422	20.6%
Math overall	103	8	111	5.4%
Science overall	33	6	39	1.9%
Grad rate overall	153	40	193	44.0%
Accountability Index	75	1	76	3.7%
Total¹	5,540	3,542	9,082	

** N=2046 for academic measures; N=439 for extended graduation rate measures

¹ Duplicated count

APPENDIX B

IDENTIFYING PRIORITY SCHOOLS AND DISTRICTS

The advisors (see Appendix D) generated a comprehensive list of quantitative and qualitative data that could be used to determine which schools in the “struggling” tier should be identified as needing more significant support from the state over a longer period of time (the Priority tier). Schools in the Priority tier would have the greatest need based on consistent underperformance on multiple measures (grades, subjects, indicators) over multiple years. The advisors assumed that being in this tier would generate the opportunity for substantially more support. The following factors were initially identified (the advisors did not discuss data for identifying Priority *districts*).

Contextual Data

- Type of school (alternative school, institution)
- Changes in student demographic profile (e.g., rapid increase in low-income or ELL students)
- What programs are included in the school (e.g., concentrations of ELL, special education, gifted)
- Program changes (e.g., establishing new ELL or special education programs)
- Student mobility
- Number of languages spoken by students
- Feeder schools
- Boundary changes (closures, consolidations)
- Construction or renovation projects

Analysis of Assessment Results (annual and trends over time)

- Achievement trends over multiple years for each subject area
- Size of the gap between WASL scores in different subjects
- Size of the achievement gap
- Percent students meeting 3 of 3 and 4 of 4 standards
- Trends for subgroups (gender, race/ethnicity, low-income) and programs (ELL, special education)
- Level of growth over time
- Changes in scale scores
- How performance compares to similar schools
- Results of students who have been in the school for longer periods of time (track cohorts of students to see how percent meeting standard changes over time, review results for just “continuously enrolled” students, the percentage of students meeting standard the next year in the next grade compared to the previous year, e.g., the percent in grade 4 in one year compared to the percent in grade 5 the next year)
- Results from retakes (high school) and collection of evidence
- WLPT results for students from different language backgrounds, percentage of students exiting ELL program

AYP Results

- Results generated with minimum Ns, confidence intervals, and continuously enrolled students (helps prevent false positives)

- How far the “all” group is from the annual goal
- Proficiency, participation, and other indicator results for all subgroups
- Number and percentage of cells not making AYP
- Which subgroups and subjects did not make AYP (ELL, special education, and participation rates count less, the all and race/ethnic groups count more)

Other Quantitative Data (some may only be available at the district or school levels)

- *Graduation data:* On-time and extended graduation rates for all students and subgroups, difference in rates, percentage of students still enrolled after four years
- *Dropout data:* Annual and cohort dropout rates for all students and subgroups, difference in rates
- *Discipline data:* Number of suspensions and expulsions, source of referrals, types of infractions, types of students being disciplined the most
- *Perception results:* Surveys of staff, parents, students about school conditions and how the results differ from one another
- *Classroom conditions:* Class sizes, student/teacher ratios by grade and subject
- *Staff characteristics:* Percentage of staff with certificates, teacher education/experience levels
- *Staff turnover:* Teacher and leadership changes at school and district levels
- *District assessments:* Results from any other assessments (e.g., MAP, grade 2 reading, portfolios)
- *Volunteers:* Number of parents volunteers, how they are used
- *Retention:* Number and percentage of students retained in grade, number and type of subjects not passed, level of credit deficiency
- *Finances:* Amount generated by local levies/bonds, fund balances, amount and sources of outside funding, stability in funding over time
- *District characteristics:* Number and percentage of schools in Tier 3, percentage of district students enrolled in Tier 3 schools
- *Data anomalies:* Incorrect data reported that could affect analyses, missing data, reason for missing data, number of ratings generating the average index

Qualitative Data

- *District role:* Resource amounts and types allocated to school, type of staff and programs provided, funding levels, type and intensity of interventions made to date, appropriateness of district policies, data analysis capacity, role of the district in school improvement efforts
- *Initiatives:* Number being attempted, focus and validity of initiatives, level of integration/cohesion among activities
- *Data use:* Quality of data system, capacity to use data, how information is used
- *Self-assessments:* Quality and use/implementation of school improvement plans
- *Staff relations:* Level of collaboration among staff and administrators within the school, union relations
- *Results from external reviews:* Results from accreditation and OSPI’s Comprehensive Program Review (CPR), input from ESDs

Given the comprehensive nature of this list and the limited capacity to analyze all these data for every school in the “struggling” tier, the list was re-examined to determine which were the most important factors to review. Those factors appear in the body of this document.

Schools serving special populations require separate analyses. For example, schools serving high concentrations of more challenging student populations (e.g., alternative schools, institutions, those primarily serving ELL students and those with disabilities) often have low index results that would put them in the “struggling” tier. These schools have great need and should not be automatically excluded from being a Priority school. A closer look into the quality of programs serving these students is needed to see if more support should be provided. These kinds of schools may require an alternative accountability system (states like Texas have set up such a system). Some institutions should be excluded (e.g., jails & detention centers) but other included (e.g., long-term psychiatric facilities).

Other types of schools may need special analyses as well. For example, results for *very small schools* ($N < 10$) are available but cannot be revealed to protect confidential information about students. However, the results could still be examined for trends over time. The number of *virtual schools* is increasing, often serving home-based students who are not required to take state assessments and may not be authorized to grant diplomas, which could mean there are few or no outcomes to measure. While some of these schools will generate results, they often serve many students outside the district, which means the school’s results are not included in the district results.

Certain preconditions need to exist for schools and district for them to use the additional resources effectively. For example, schools in the lowest tier need to be ready to benefit from the extra support. Without their buy-in, the chances for a successful reform are minimal. Size and location may need to be considered. If the number of schools in the “struggling” tier is high and exceeds the level of resources available to support them, the state may want to require a minimum number of students per school before providing assistance to ensure cost-effectiveness of the assistance. Similarly, those identified for the Priority tier may have a wide geographic distribution. A single small school in a remote location may have the same level of need as a cluster of larger schools in a more accessible location. The state will need to determine how best to allocate its limited resources to ensure the cost effectiveness of its support. Finally, the state may want to consider providing support by geographic location to ensure equity in the distribution of the assistance.

APPENDIX C

CURRENT STATE ASSISTANCE PROGRAM

SCHOOL IMPROVEMENT ASSISTANCE

The mission of the Office of Superintendent of Public Instruction's School Improvement Assistance (SIA) program is to help build capacity for districts and schools to improve student achievement through the use of the continuous school improvement model. This comprehensive model of support is unique in the United States. While many states have accountability systems that focus on rewards, punishments and takeovers, the SIA program provides comprehensive support for schools. Independent studies of the program have noted that the schools that received assistance for three years showed greater achievement gains than their respective comparison groups and the state as a whole. Nearly 60% of schools that have participated in SIA have exited federal improvement status and have made Adequate Yearly Progress (AYP) in the last two years of the program. The studies found further evidence that achievement gaps have been reduced in SIA schools.

Program Components

- **School Improvement Facilitator (SIF):** The facilitator works with OSPI, the school district, school, and a **School Improvement Leadership Team (SILT)** to develop a plan to address identified needs and to prepare and implement a jointly developed performance agreement between the school, school district and OSPI. The school improvement facilitators are experienced educators who have been successful in improving student performance and work approximately 1.5 days a week with each school for the three years of school improvement plan development and implementation. The school improvement leadership team includes representatives from the district and school staff, parents, and community members. Additional members may include educational service district (ESD) staff, OSPI staff and students.
- **Comprehensive Needs Assessment/School Performance Review:** The needs assessment/school performance review is completed jointly by the school improvement leadership team, school district, OSPI, and a team of peer educators and experts. The school's strengths and challenges are identified and recommendations for improvement are developed. The school's curriculum, leadership, instructional practices and resources, assessment results, allocation of resources, parental involvement, support from the central office, and staff, parent, and student perceptions are examined. Student performance data, indicators from the "Nine Characteristics of High Performing Schools" and the results of a review of the school's reading and math instructional practices and program, are used to identify areas to consider for improvement. The assessment/audit includes the administration of survey instruments and an on-site visit.
- **School Improvement Process, Tools, and Support:** Schools are given the necessary processes, tools and expertise for the school improvement leadership team to develop a comprehensive *School Improvement Plan*. Funds are provided to contract with individuals to assist with components of the plan, and the school improvement facilitator are responsible for organizing and facilitating meetings in coordination with school and district staff.

- **Funds for Staff Planning and Collaboration:** Funds for planning time related to the development of the school improvement plan are provided. These funds may be used to provide stipends for school improvement leadership team members. A minimum of three days must be devoted to planning time for all staff during the development of the school improvement plan. The funds can be used to pay staff stipends or to pay substitute teachers.
- **Performance Agreement:** Once the school improvement plan is completed, a two-year performance agreement is jointly developed by the school, school district and OSPI. The agreement identifies specific actions and resources the school district, the school and OSPI will commit to implement the school improvement plan. The agreement also includes a timeline for meeting implementation benchmarks and student improvement goals.
- **Implementation and Sustainability:** Tools and resources for the implementation of the performance agreement are provided during years two and three. The resources and expertise are determined on a case-by-case basis for each school, but could include such support as the provision of expertise in working with diverse student populations (e.g. special education, English language learners), funding and expertise to implement research-based practices and programs, and funding for time for staff collaboration. Schools and school districts are expected to ensure that existing funds are used effectively and to dedicate school district resources as identified in the jointly developed Performance Agreement.
- **Training Workshops:** Funds are provided to send a team of representatives to workshops during the school year to effectively plan for school improvement.
- **Professional Development:** Professional development opportunities for the school's principal and other school instructional leaders are provided in partnership with OSPI and the Association Washington School Principals (AWSP). Workshops are available during the school year.

The Process

Year 1: School Improvement Planning and Performance Agreement

- Conduct needs assessment through school performance review (formerly educational audit)
- Support staff training
- Develop school improvement plan/ performance agreement
- Develop student performance goals and evaluation criteria

Year 2: Implementation

- Tools and resources to implement the school improvement plan and performance agreements
- Evaluate student progress based on goals in the agreement

Year 3: Sustainability

- Tools and resources to build capacity and develop sustainability
- Evaluate student progress based on goals in the agreement

DISTRICT IMPROVEMENT ASSISTANCE

For 2008-2009, districts fall in four district improvement groupings: (1) New in Step 1; (2) Continuing in Step 1; (3) New in Step 2; and (4) Continuing in Step 2. The technical assistance provided to districts in improvement status varies to meet the needs of districts either as they are developing their improvement plans or in various stages of implementation of their plans. The following areas are the most common types of support.

- A. Providing a School System Resource Guide (SSIRG):** OSPI and WASA collaborated in developing a resource planning guide that supports districts as they analyze existing systems, structures, data, research findings, and more as they develop/revise their district improvement plan. A revision to the SSIRG is planned to be completed in 2008-09.
- B. Providing a Part-time, External District Improvement Facilitator:** District Improvement Facilitators are experienced educators who have been successful in improving student performance and receive continuous training through a partnership with WASA throughout the year. The selection of the facilitator is a collaborative effort between OSPI and each district. The facilitator works to help build the district's capacity to support high-quality, data-driven, research-based district improvement efforts.
- C. Providing or Arranging for Professional Development:** Additional resources for professional development to expand capacity of district and school personnel to sustain continuous improvement focused on improvement of instruction may be provided to meet the needs of districts.
- D. Provide for a District Educational On-Site Review:** Districts can request an educational on-site review to be completed by a team of peer educators and experts. The district's strengths and challenges are identified and recommendations for improvement are developed and provided to the district.
- E. Providing Identified Expertise:** Additional resources and expertise OSPI could provide is determined on a case-by-case basis for each district, but could include such support as expertise in working with diverse student populations (e.g., special education, English language learners), funding and expertise to implement research-based practices and programs, and funding for team collaboration time.
- F. Providing Limited Grant Money:** Districts may apply for two levels of grant support to assist in implementing one or more of the technical assistance opportunities listed A-E above.

OSPI recognizes the need to emphasize internal capacity building in districts and to revise its support systems and procedures over time.

APPENDIX D

ADVISORY GROUP MEMBERS

Dr. Pete Bylsma, an independent consultant and former state director of research and accountability at OSPI, was hired to help prepare the proposed index for Board review. He was assisted by a number of advisors. This diverse set of advisors reviewed the work that had been done to date, discussed numerous technical issues related to the proposed index, discussed the criteria for recognizing schools and districts, and identified quantitative and qualitative data that can be used to examine schools in the “struggling” tier to determine if they should be a Priority school needing much greater state assistance. Other stakeholders from OSPI were included in some of the discussions, and a State Board working group that focused on System Performance Accountability also provided feedback on the draft proposal.

Members of the advisory group were:

- Dr. Karen Banks, Shelton SD (District Improvement Facilitator)
- Ms. Maggie Bates, Hockinson SD (Assistant Superintendent)
- Ms. JoLynn Berge, OSPI (Federal Policy and Grant Administrator)
- Dr. Phil Dommies, North Thurston SD (Assessment Director)
- Dr. Linda Elman, Tukwila SD (Assessment/Research Director)
- Mr. Doug Goodlett, Vancouver SD (Special Services Director)
- Dr. Peter Hendrickson, Everett SD (Assessment Director)
- Dr. Feng-Yi Hung, Clover Park SD (Assessment/Evaluation Director)
- Dr. Nancy Katims, Edmonds SD (Assessment Director)
- Dr. Bill Keim, ESD 113 (Superintendent)
- Ms. Linda Munson, South Kitsap SD (Special Programs Director)
- Dr. Michael Power, Tacoma SD (Assistant Superintendent)
- Mr. Bob Silverman, Puyallup SD (Executive Director for Assessment)
- Ms. Nancy Skerritt, Tahoma SD (Assistant Superintendent)
- Dr. Lorna Spear, Spokane SD (Executive Director for Teaching and Learning)
- Dr. Alan Spicciati, Highline SD (Chief Accountability Officer)

APPENDIX E

RELEVANT LEGISLATIVE MANDATE

RCW 28A.305.130 Powers and duties — Purpose.

The purpose of the state board of education is to provide advocacy and strategic oversight of public education; implement a standards-based accountability system to improve student academic achievement; provide leadership in the creation of a system that personalizes education for each student and respects diverse cultures, abilities, and learning styles; and promote achievement of the goals of RCW 28A.150.210. In addition to any other powers and duties as provided by law, the state board of education shall:

(4) For purposes of statewide accountability:

(c) Adopt objective, systematic criteria to identify successful schools and school districts and recommend to the superintendent of public instruction schools and districts to be recognized for two types of accomplishments, student achievement and improvements in student achievement. Recognition for improvements in student achievement shall include consideration of one or more of the following accomplishments:

(i) An increase in the percent of students meeting standards. The level of achievement required for recognition may be based on the achievement goals established by the legislature and by the board under (a) of this subsection;

(ii) Positive progress on an improvement index that measures improvement in all levels of the assessment; and

(iii) Improvements despite challenges such as high levels of mobility, poverty, English as a second language learners, and large numbers of students in special populations as measured by either the percent of students meeting the standard, or the improvement index. When determining the baseline year or years for recognizing individual schools, the board may use the assessment results from the initial years the assessments were administered, if doing so with individual schools would be appropriate;

(d) Adopt objective, systematic criteria to identify schools and school districts in need of assistance and those in which significant numbers of students persistently fail to meet state standards. In its deliberations, the board shall consider the use of all statewide mandated criterion-referenced and norm-referenced standardized tests;

(e) Identify schools and school districts in which state intervention measures will be needed and a range of appropriate intervention strategies after the legislature has authorized a set of intervention strategies. After the legislature has authorized a set of intervention strategies, at the request of the board, the superintendent shall intervene in the school or school district and take corrective actions. This chapter does not provide additional authority for the board or the superintendent of public instruction to intervene in a school or school district;

(f) Identify performance incentive systems that have improved or have the potential to improve student achievement.