

State Accountability Index Final Report to the State Board of Education

February 16, 2010

Pete Bylsma, EdD, MPA
BYLSMAPJ@COMCAST.NET

TABLE OF CONTENTS

Background.....	1
Accountability Index	2
Preliminary School Results.....	4
Preliminary District Results.....	8
Handling Special Cases	12
Alternative Schools	12
Counting ELL Results	13
Improvement by High Performers	15
Integrating the Federal and State Accountability Systems	15
Determining AYP	16
In-Depth Analysis Before Making AYP Decisions	18
Consequences	20
Advantages of the State System.....	23
Recognition.....	25
Appendix A – Indicators and Outcomes	29
Overview	29
Using the Index	31
Calculation Methods	32
Indicator 1: Achievement of Non-Low Income Students	33
Indicator 2: Achievement of Low Income Students	35
Indicator 3: Achievement vs. Peers	37
Indicator 4: Improvement	41
Extended Graduation Rate Measure	42
Other Useful Measures to Report	43
Distribution of Index	44
Tiers	47
Descriptive Statistics.....	50
Recognition System	52
Appendix B – In-Depth Analysis	55
Appendix C – Current State Assistance Program	58
Appendix D – Relevant Legislative Mandates.....	61
Appendix E – State Board of Education Accountability	64
Framework and Resolution	
Appendix F – Conceptual Framework for Support for	67
Struggling Schools (Mass Insight)	
Appendix G – Current AYP Sanctions.....	70

Appendix H – Additional Accountability Data.....	74
Appendix I – Advisory Group Members	80

ABBREVIATIONS

AMAO	Annual Measurable Achievement Objective
AYP	Adequate Yearly Progress
ELL	English language learner
NCLB	No Child Left Behind
OSPI	Office of Superintendent of Public Instruction
SBE	State Board of Education
WAAS	Washington Alternate Assessment System
WALA	Washington Association for Learning Alternatives
WLPT	Washington Language Proficiency Test

BACKGROUND

Educational accountability systems require several components: (1) measures of effectiveness, (2) goals to guide improvement efforts, (3) reports that provide useful information to policymakers, educators, and parents, and (4) a set of consequences that recognize exemplary performance and support those needing more help. To this end, the Legislature requires the State Board of Education (SBE) to develop a statewide accountability system to help improve academic performance. Changes being made to the state's assessment and data systems and at the U.S. Education Department provide an opportunity to consider new ideas related to accountability measures, goals, reporting, and consequences.

As a first step, SBE is required to adopt objective, systematic criteria to identify schools and districts for recognition and for receiving additional state support. To meet this requirement, SBE has developed an Accountability Index that sorts schools and districts into different "tiers" based on multiple measures. This Index plays a key role in providing feedback about the status of education reform in schools and districts and supporting continuous improvement efforts.

Various stakeholders and technical advisors provided input and feedback about the Index and how to identify schools and districts in most need. This document summarizes the way the Index is calculated, how it is used to identify schools and districts for recognition and additional state support, and its relationship with federal accountability system. The appendixes provide more detailed information about the state accountability system.

The Index was developed using a set of guiding principles. To be effective, the Index should:

- Be transparent and easy 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);
- Use 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 produce accurate results;
- 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;
- Provide multiple ways to reward success; and
- Be flexible enough to accommodate future changes.

Several assumptions were made during the development of the index.

- Struggling 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.
- Struggling schools and districts will be eligible to receive additional resources to make dramatic improvement in student outcomes.
- Struggling schools and districts will be required to participate in a state-supported initiative if offers of additional support are not accepted and substantial improvement does not occur after two years (see Appendix F for information about these issues).

The Index is only one part of the overall accountability system. While it provides evidence and a rationale for descriptive labels applied to schools and districts, the Index does not set improvement goals, nor does it provide all the information needed to improve the education system. Hence, improvement goals still need to be developed and more information needs to be reported. In addition, the state needs to design a system to support schools and districts in most need. This document describes the Accountability Index, examples of how Index results can be reported, how Index data will be used for recognition purposes, and other data that can be reported to keep stakeholders informed. SBE is still in the process of establishing improvement goals.

ACCOUNTABILITY INDEX

The Index is based on how schools and districts perform on a set of outcomes and indicators. Specifically, the 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 of students who are not from low-income families;
- (2) achievement of students from low-income families;
- (3) achievement of all students when compared to “peers,” i.e., those with similar student characteristics (the percentage of students who have a disability, are learning English, come from low-income families, are mobile, and are designated as gifted); and
- (4) the improvement in the achievement of all students from the previous year.

The results of these 20 measures form the matrix shown in Table 1.

Table 1: Matrix of Accountability Measures

INDICATOR	OUTCOME				
	Reading	Writing	Math	Science	Ext. Grad. Rate
Achievement of non-low income students					
Achievement of low income students					
Achievement vs. peers					
Improvement from the previous year					

Each cell of the matrix is rated on a 7-point scale (from 1 to 7) fixed benchmarks. The 7-point scale gives sufficient “spread” in the results. Each of the four subjects is rated using the same set of benchmarks across the entire school/district (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** and ranges from 1.0 to 7.0. The higher the Index, the better the level of performance of the school/district. High schools and districts have 20 measures, while most elementary and middle/junior high schools have only 16 measures because they do not have graduates (some schools may not test some subjects based on their grade configuration).

Table 2 shows how each of the five outcomes are measured using the four indicators and the benchmarks that produce the ratings. Achieving a high rating is a challenge, especially in content areas where performance has been low (i.e., math and science). The Learning Index is used to measure the assessment outcome for two indicators: *achievement compared to peers* and *improvement*. This index (not to be confused with the Accountability Index) accounts for the percentage of students performing at the five different assessment levels, not just those meeting standard. The Learning Index ranges from 0 to 4, with 4.00 as the highest score (similar to a grade

point average). This index is explained in detail in Appendix A, along with other details about how the indicators, outcomes, and benchmarks were selected and measured.

Table 2: Benchmarks and Ratings for Outcomes and Indicators

	READING	WRITING	MATH	SCIENCE	EXT. GRAD. RATE ¹	
ACHIEVEMENT (NON-LOW INCOME)	<u>% MET STANDARD</u>		<u>RATING</u>		<u>RATE</u>	<u>RATING</u>
	90 - 100%		7		> 95	7
ACHIEVEMENT (LOW INCOME)	80 - 89.9%		6		90 - 95%	6
	70 - 79.9%		5		85 - 89.9%	5
	60 - 69.9%		4		80 - 84.9%	4
	50 - 59.9%		3		75 - 79.9%	3
	40 - 49.0%		2		70 - 74.9%	2
	< 40%		1		< 70%	1
ACHIEVEMENT VS. PEERS²	<u>DIFFERENCE IN LEARNING INDEX</u>		<u>RATING</u>		<u>DIFFERENCE IN RATE</u>	<u>RATING</u>
	> .20		7		> 12	7
	.151 to .20		6		6.1 to 12	6
	.051 to .15		5		3.1 to 6	5
	-.05 to .05		4		-3 to 3	4
	-.051 to -.15		3		-3.1 to -6	3
	-.151 to -.20		2		-6.1 to -12	2
	< -.20		1		< -12	1
IMPROVEMENT (change from the previous year)	<u>CHANGE IN LEARNING INDEX</u>		<u>RATING</u>		<u>CHANGE IN RATE</u>	<u>RATING</u>
	> .15		7		> 6	7
	.101 to .15		6		4.1 to 6	6
	.051 to .10		5		2.1 to 4	5
	-.05 to .05		4		-2 to 2	4
	-.051 to -.10		3		-2.1 to -4	3
	-.101 to -.15		2		-4.1 to -6	2
	< -.15		1		< -6	1

Note: Assessment-related results are the combined results of both the regular state tests and the alternate assessments (WAAS) from all grades.

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

²This indicator adjusts the outcomes using statistical methods (multiple regression) to control for five student characteristics beyond a school's control: the percentage of low-income, ELL, special education, gifted, and mobile students. (Mobile students are those who are not continuously enrolled from October 1 through the entire testing period.) Scores are the difference between the actual and predicted levels of the Learning Index. Scores above 0 are "beating the odds" and negative scores are below the predicted level. Separate analyses are conducted for schools for each type of school (elementary, middle, high, multiple grade levels). District calculations also control for current expenditures (adjusted for student need).

The Index applies to districts using the same rules, indicators, and outcomes that are used for school accountability. The results are 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 79 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.

Tier assignments are determined based on the Index score. Schools and districts will fall into five tiers based on their accountability index score. The 5-tier system provides sufficient differentiation

among schools and districts to guide decisions about recognition and those needing further support. Table 3 shows the ranges for the 5-tiers and their descriptive names.

Preliminary School Results

Table 3 also shows the distribution of schools using the criteria shown in Table 2 and data from 2007. A total of 2,011 schools had a reportable Index score (the Index for schools with fewer than 10 students assessed across all tested grades is calculated but not reported). More detailed results are reported in Appendix A.

Table 3: Tier Ranges and 2007 School Results (N=2,011)

Tier	Index Range	Number of Schools	Percent of Schools	Number of Students	Percent of Students	Average Enrollment
Exemplary	5.50 – 7.00	81	4.0%	28,650	2.9%	354
Very Good	5.00 – 5.49	131	6.5%	64,500	6.4%	492
Good	4.00 – 4.99	591	29.4%	314,700	31.3%	532
Fair	2.50 – 3.99	980	48.7%	523,000	52.0%	534
Struggling	1.00 – 2.49	228	11.3%	74,000	7.4%	325

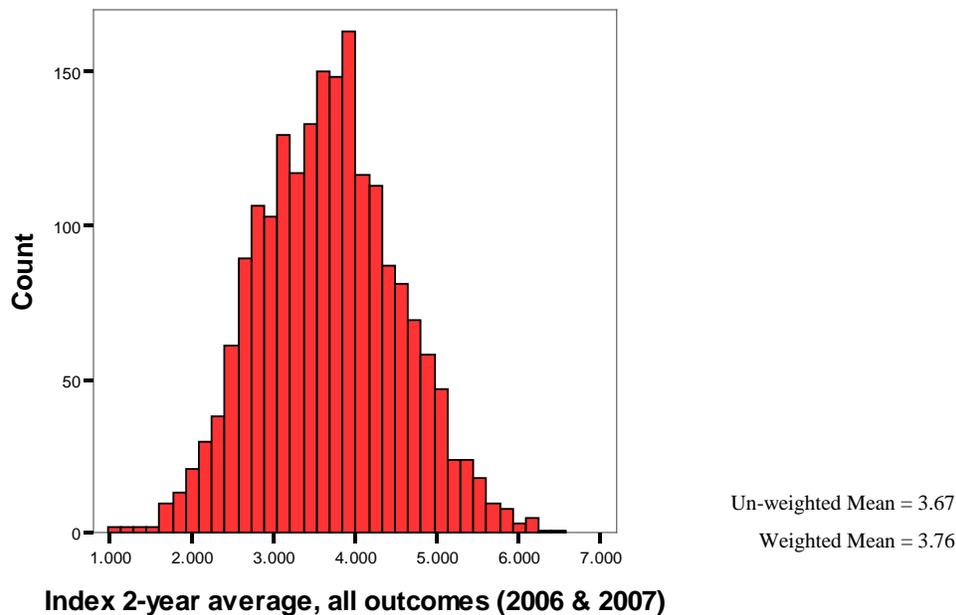
Note: Schools averaged 3.71, with 4.00 being the mid-point on a 7-point scale.

In 2007, nearly half the schools were in the Fair tier. The 228 schools in the Struggling tier enrolled 74,000 students, or roughly 1 in every 14 students statewide. Of the schools in this tier, 98 (43%) were alternative schools or served other special populations, with enrollment of 12,400 students. The 130 “regular” schools in the Struggling tier enrolled 61,600 students. Of the schools in the Struggling tier in 2007, 143 (63%) did not make AYP, and 62 were in “school improvement.”

There were 1,984 schools that had enough students to generate the Accountability Index in both 2006 and 2007. *Over a 2-year period*, only 2% (42 schools) had an average Index of 5.5 or greater (Exemplary tier), while 7.5% (149 schools) had an average Index below 2.50 (Struggling tier). Of the 149 schools in the lowest tier, 89 were regular schools that enrolled approximately 39,000 students (roughly 4% of statewide enrollment).

Figure 1 shows the distribution of the 2-year Index average for the schools that had reportable data in both years. There was little difference in the distribution of schools based on their grades served (i.e., elementary, middle, high). Alternative schools were more evenly distributed across the tiers, although they were more likely to fall into the Struggling tier than any other tier. (See Appendix A for more details on the types of schools that fall into the various tiers.)

Figure 1: Distribution of Schools by Index Score (Average of 2006 and 2007)



The Index and matrix results can be made available in a “Report Card” for use by policymakers, educators, and the public. Tables 4 and 5 give examples of how the individual ratings generate the Index for two actual schools using results available from 2007. The results can 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 a quick snapshot of where a school is strong and weak and its overall rating. 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 test 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 non-low income 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 non-low income student’s rate. Its Index of 3.45 puts it in the middle of the Fair 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 test scores are well above the state average in several grades but below the state average in one grade. It had sharp declines in its Learning Index from very high levels the previous year, resulting in low improvement ratings in 3 subjects (although most of the students still met standard in reading and writing). Its Learning Index was very high compared to schools serving similar students. Low-income students had lower rating than non-low income students in three subjects, which is typical in most schools. The graduation rate does not apply. Its Index of 4.44 is above the middle of the scale and is in the middle of the Good tier.

Table 4: “Actual” High School Ratings and Results, 2007

Indicator	Reading	Writing	Math	Science	Ext. Grad. Rate	Average
Non-low inc. ach.	6	6	3	1	5	4.20
Low-inc. ach.	4	4	1	1	7	3.40
Ach. vs. peers	2	2	2	2	6	2.80
Improvement	1	4	1	4	7	3.40
Average	3.25	4.00	1.75	2.00	6.25	3.45
Non-low inc. ach.	80.0%	85.0%	51.3%	32.9%	89.0%	
Low-inc. ach.	63.5%	63.9%	22.7%	16.7%	100.0%	
Ach. vs. peers	-.173	-.178	-.157	-.182	+9.4	
Improvement	-.330	-.038	-.238	+.021	+8.8	

ACCOUNTABILITY INDEX ←

Table 5: “Actual” Elementary School Ratings and Results, 2007

Indicator	Reading	Writing	Math	Science	Ext. Grad. Rate	Average
Non-low inc. ach.	7	7	5	1		5.00
Low-inc. ach.	5	4	4	1		3.50
Ach. vs. peers	7	7	7	5		6.50
Improvement	1	4	3	3		2.75
Average	5.00	5.50	4.75	2.50		4.44
Non-low inc. ach.	92.8%	92.0%	76.8%	38.1%		
Low-inc. ach.	77.2%	61.4%	61.8%	20.8%		
Ach. vs. peers	+.450	+.502	+.493	+.119		
Improvement	-.228	-.041	-.084	-.087		

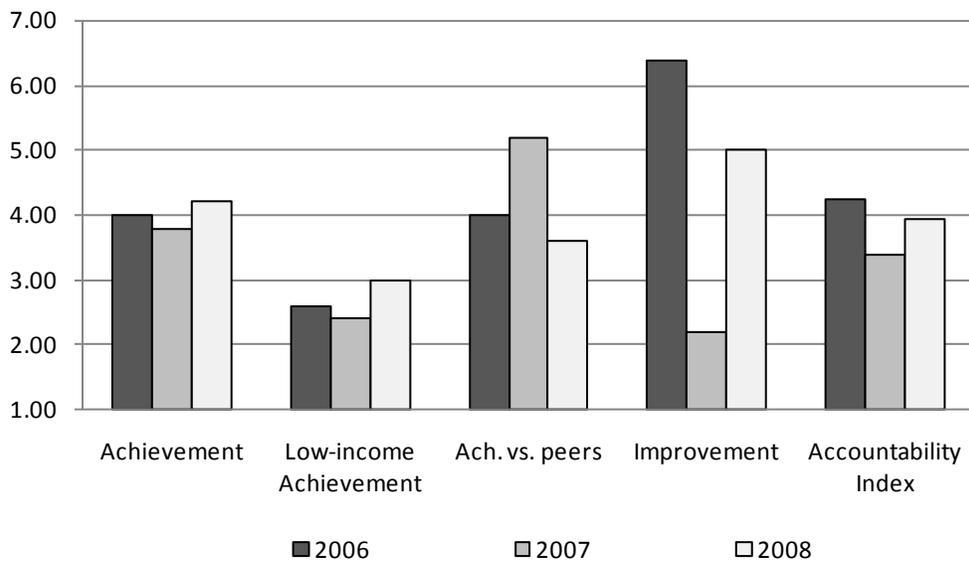
ACCOUNTABILITY INDEX ←

Other types of tables and charts can illustrate the results. Table 6 and Figure 2 show examples of how all the results for another high school can be shown across multiple years to show trends over time. Adding stars or other symbols can help present the results graphically.

Table 6: Showing Accountability Results Over Time (High School)

<i>Indicator/Outcome</i>	<i>2006</i>		<i>2007</i>		<i>2008</i>	
Non-low inc. ach.	4.00		3.80		4.20	
Reading	*****	6	*****	5	*****	6
Writing	*****	6	*****	6	*****	6
Math	***	3	***	3	***	3
Science	*	1	*	1	**	2
Ext. grad. rate	****	4	****	4	****	4
Low-income ach.	2.60		2.40		3.00	
Reading	*****	5	****	4	****	4
Writing	****	4	*****	5	*****	5
Math	*	1	*	1	*	1
Science	*	1	*	1	*	1
Ext. grad. rate	**	2	*	1	****	4
Ach. vs. peers	4.00		5.20		3.60	
Reading	****	4	*****	5	***	3
Writing	*****	5	*****	5	****	4
Math	****	4	*****	6	*****	5
Science	*****	5	*****	5	**	2
Ext. grad. rate	**	2	*****	5	****	4
Improvement	6.40		2.20		5.00	
Reading	*****	7	*	1	****	4
Writing	*****	7	****	4	*****	6
Math	*****	6	*	1	*****	5
Science	*****	5	*	1	****	4
Ext. grad. rate	*****	7	****	4	*****	6
INDEX	<i>Good</i>	4.25	<i>Fair</i>	3.40	<i>Fair</i>	3.95

Figure 2: Average Ratings, 2006-2008 (High School)



Preliminary District Results

The Index applies to districts using the same rules, indicators, and outcomes that are used for school accountability. The results are 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 79 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.

Figure 3 shows the distribution of the *number of schools* by tier for an *actual* district in 2007, while Figure 4 shows the *percentage of students* enrolled at those schools. This district had an Index of 4.30, which is in the Good tier. Figure 5 shows the Index results for all the schools in a different district using 2008 data. One elementary school had an index in the Exemplary tier.

Figure 3: Distribution of Schools by Grade Level and Tier in District “A” (2007)

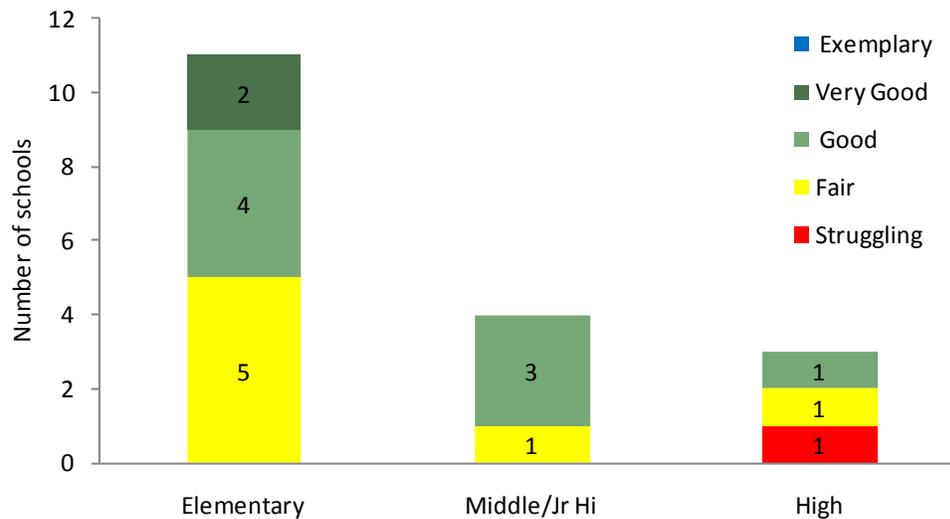
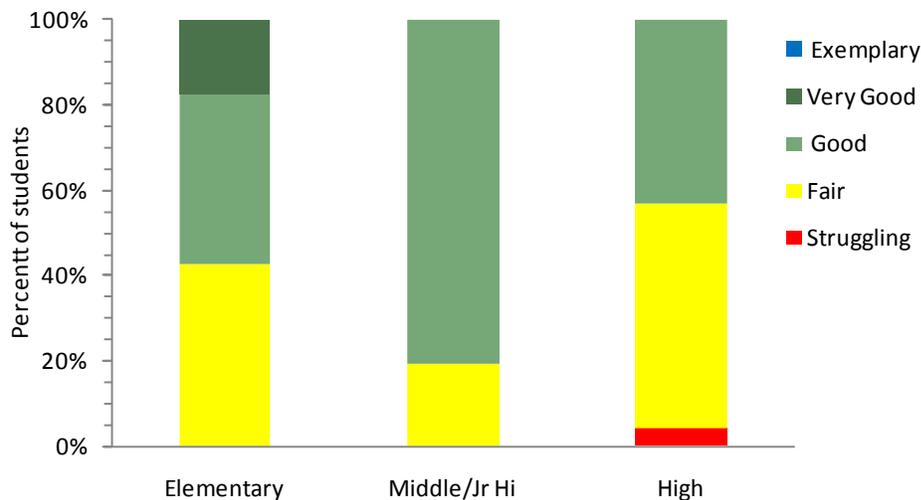


Figure 4: Distribution of Students by Grade Level and School Tiers in District “A”



Note: One alternative high school, with an Index in the Struggling tier, has relatively few students.

Figure 5: Index for Schools in District “B” (2008)

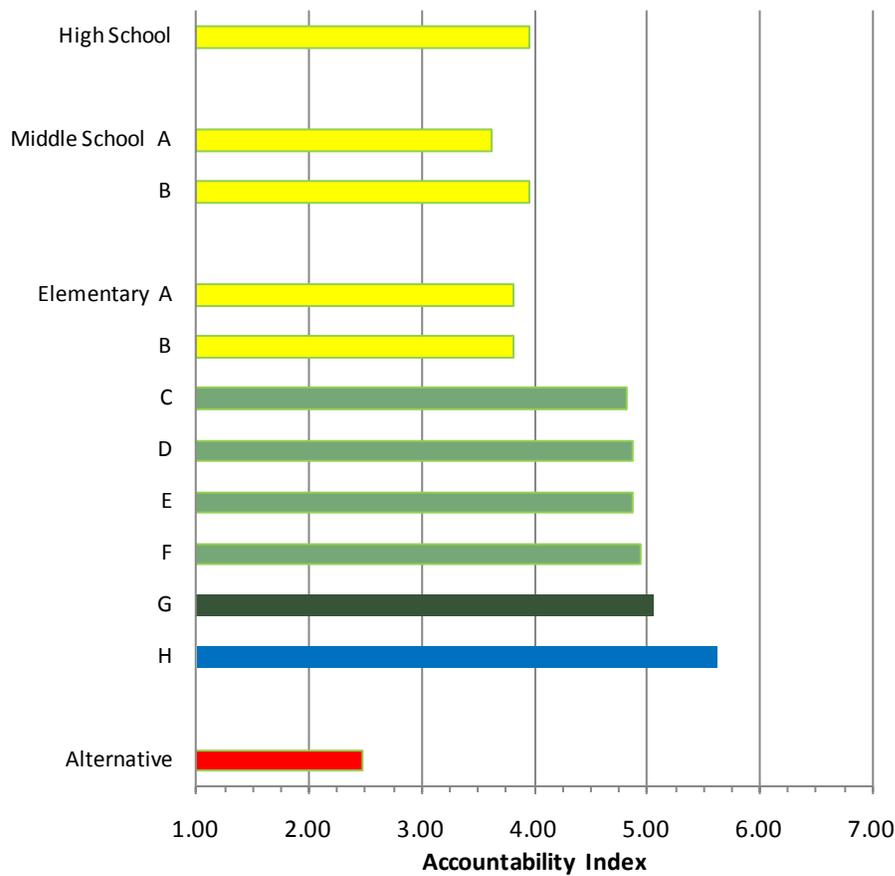


Table 7 shows the statewide district results using the same criteria and rating system used for schools. Districts are more tightly clustered in the distribution than schools, with fewer districts in the top and bottom tiers. More than 60% fell in the Fair tier. The 17 districts in the Struggling tier in 2007 tended to be small (fewer than 750 students). Seven of these 17 districts made AYP in 2007 and only three were in district improvement. This is due to the low AYP targets in 2007, a large margin of error for small districts, and fewer student groups meeting the minimum reporting number. Although larger districts were not usually in the Struggling tier, some had many of their schools in the Struggling tier. For example, 13 districts that were not in the Struggling tier had at least two regular schools with a 2-year Index average below 2.50, and three districts had at least five regular schools with a 2-year Index average below 2.50.

Table 7: Tier Ranges and 2007 Results for Districts (N=291)

Tier	Index Range	# of Districts	% of Districts	# of Students ¹	Average Enrollment ¹
Exemplary	5.50 – 7.00	1	.3%	360	360
Very Good	5.00 – 5.49	9	3.1%	31,500	3,500
Good	4.00 – 4.99	87	29.9%	278,500	3,200
Fair	2.50 – 3.99	177	60.8%	692,500	3,910
Struggling	1.00 – 2.49	17	5.8%	17,500	1,030
Priority	1.00 – 3.99	TBD	TBD	TBD	TBD

Note: Districts averaged 3.69, with 4.00 being the mid-point on a 7-point scale.

¹Approximate numbers

Figure 6 shows the distribution of all the *district* Index results in 2007. Figure 7 shows the 2-year average Index for each *county*.

Figure 6: Distribution of Index Score by District, 2007

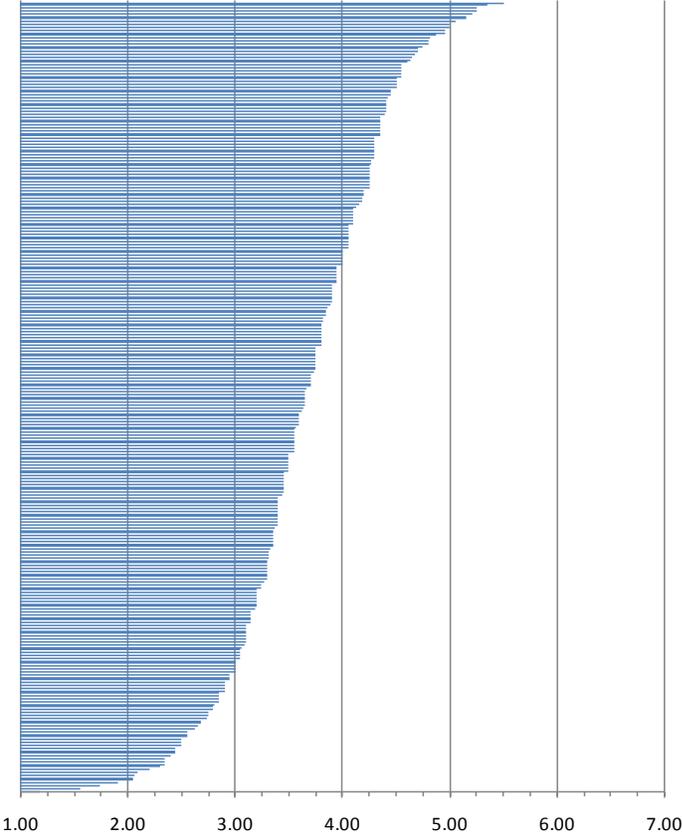
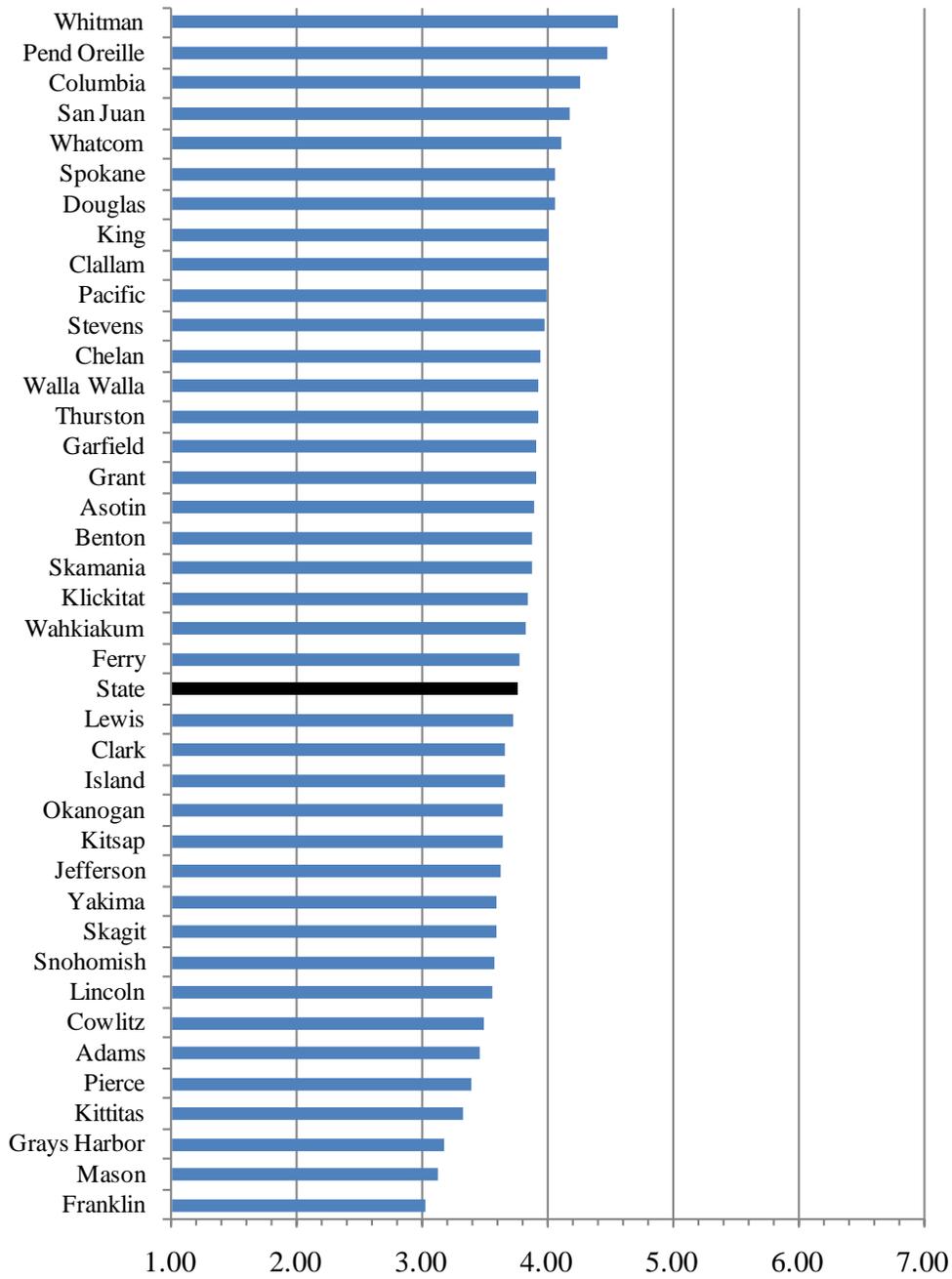


Figure 7: Distribution of Index Results by County



Average Index, 2006 & 2007

Note: The 2-year average is weighted by district enrollment, so larger districts count more than smaller districts. This gives an accurate representation of the county results.

HANDLING SPECIAL CASES

The accountability system needs to be flexible enough to accommodate some special situations. This includes holding alternative schools accountable using additional data, excluding some ELL results from the Index calculations, and not counting the improvement cells when achievement is at very high levels.

Alternative Schools

Many types of alternative schools exist in the state. More than half the “schools” with this designation serve at-risk students in grades 9-12. Some believe these schools have taken on more challenging students, which allows more traditional schools to generate better outcomes with their remaining students. On the other hand, some schools offer special programs for students who are not at-risk and who must meet rigorous academic requirements for admission. Some are considered “alternative” because they do not use a normal school approach. A growing number of schools serve students through distance or digital learning and offer instruction electronically, usually via the Internet. Parent Partnership Programs are a type of “school” where parents are the primary instructor, the district provides instructional materials, and a certificated teacher acts as a consultant. Alternative schools exist at the elementary and middle school levels as well, and some programs are offered through independent contractors. Some schools with this designation target special student populations (e.g., special education, gifted, ELL), and some are held on college campuses or at night.

The wide variation in the focus, structure, and clientele of alternative programs across the state poses unique accountability challenges. Their results are included in district results, but school-level outcomes may be very high or low because of the type of students served. As a result, no “peer” indicator is computed for these schools. Most of these schools are relatively small and their total 2007 enrollment was less than 4% of enrollment statewide, but many of them serve student populations facing significant challenges.

Given the specialized nature of many alternative schools, some believe these schools should be held accountable through more than just an Index score. Other states have designed accountability systems specifically for schools serving at-risk student populations. For example, California has set up a substitute method to hold these schools accountable.¹ Alternative schools in California can volunteer to be held accountable using 3 of 15 other outcomes over an extended period of time (e.g., 3 years). A school might choose to be evaluated by growth in the number of credits earned, improvement in the attendance rate, and gains on pre-post tests given during the year. However, California’s system is very complex.

Progress is being made in Washington that will help hold alternative schools accountable in a more appropriate way.

- The Washington Association for Learning Alternatives (WALA) has compiled research on best practices among alternative schools. Information on these best practices, as well as research on effective schools, will be used as a framework to hold alternative schools accountable. Specifically, these types of schools will be examined in eight areas using a set of rubrics that must still be developed: (1) student learning in the real world, (2) school atmosphere and support,

¹See <http://www.cde.ca.gov/ta/ac/am> for information on California’s Alternative School Accountability Model.

(3) staff quality and support for each other and students, (4) shared leadership and district support, (5) community partnerships, (6) student re-engagement, (7) family engagement, and (8) future (post-secondary) focus.

- The state now collects more information that will help educators and stakeholders understand outcomes in alternative (and other) schools. OSPI has created a database of students who achieve the Certificate of Academic Achievement (CAA) via state-approved alternatives to passing the state test. OSPI also has begun collecting dropout recovery data and will soon collect data on student credits and courses that will allow for analyses of credits earned.

In light of the above, accountability for alternative schools will begin using the approach used for all schools—each will receive an Index score using the normal process. However, those not making AYP two years in a row and those already in “school improvement” will undergo a deeper analysis to determine if they are using research-based best practices and showing progress. Rubrics must be developed that will provide a more appropriate framework for evaluating alternative schools. Once these rubrics are in place, OSPI will use them to determine if an alternative school not making AYP is following best practices and showing progress. OSPI will then make AYP-related decisions and recommendations about areas where improvement needs to occur in the future. If a school does not make AYP again the next year, the areas that needed improvement will be the main focus on the deeper analysis.

Counting ELL Results

Results for ELL students are currently included in AYP based on federal requirements. Their results are not included during their first year of enrollment in a U.S. public school, but their results are included beginning in the second year.² OSPI previously asked for permission to exclude ELL’s test results from AYP decisions until the student has been enrolled in a U.S. public school for three years or until the student achieves an advanced level of English proficiency on the Washington Language Proficiency Test (WLPT), whichever comes first.³ This request is based on research that shows it takes many years for ELLs to acquire “academic” proficiency in English, the state assessments are given in English, no translated versions are administered, and the students must be able to read and write English in order to understand and respond to the test items. Moreover, testing these students in English violates widely-adopted testing standards and ethics because of threats to validity and mistreatment of human subjects. However, the U.S. Education Department has denied OSPI’s request to change the way ELL students are included in accountability calculations. An alternate proposal to use a “sliding scale” was also rejected by the Department.⁴

²Per federal requirements, ELL students are not required to take the reading test in their first year of enrollment in a U.S. public school, but they are required to take the math test, regardless of how long they have been attending a U.S. public school, even though all the math test items are word problems given only in English.

³All ELLs must have a *placement* test score to determine *initial* eligibility in the state program. All ELLs must take an *annual* language proficiency test (WLPT-II) to determine *continued* eligibility. The composite score from the annual test, which reflects proficiency in reading, writing, speaking, and listening, generate a “Level” on the WLPT. The three-year exemption period reflects the views of most stakeholders and is the average time required for ELL students to exit the program.

⁴In this proposal by OSPI, ELL results were to be included in the students’ second year of enrollment when determining AYP, but proficiency would depend on the students’ English proficiency level as measured on the WLPT. For example, a grade 3 student scoring in Level 2 on the WLPT (intermediate English) would be considered proficient in reading with a test scale score of 359, while a student in Level 3 (advanced English) would be considered proficient with a scale score of 388.

Nevertheless, to improve the validity of the accountability system, the Index will exclude results for ELL students who are in their first *three* years of enrolling in a U.S. public school for any test that requires reading and writing in only English.⁵ The 3-year calendar year period begins when the student first enrolls in a school where English is the primary language of instruction.⁶ Although research has shown it takes longer than three years to acquire proficiency in English in an *academic* setting, this “extended exclusion” policy reflects the amount of time most ELLs take to meet state standards,⁷ and provides motivation to help ELLs acquire English skills.

This policy still includes the results of most ELLs. About 70% of all ELL students enter school in kindergarten, and they will have attended school for three years before taking the state assessment for the first time in grade 3. Although relatively few ELLs will have their results excluded, this policy increases the fairness and validity of the accountability results. Sensitivity analysis using data from a large district with many ELLs found that this policy created relatively few changes in the Accountability Index.

The state will take other steps to provide more accountability for helping ELL students.

- All ELLs must still take all the required state assessments after their first year of enrollment, and detailed results are reported on the state Report Card just like for other student groups.
- OSPI will start reporting WLPT results on its Report Card in a way that allows educators, parents, and other stakeholders to monitor the progress ELLs make in terms of learning English.⁸ The results will include the percentage of students achieving at each WLPT level in each subject, data on the length of time ELLs have been enrolled in the program, and the Annual Measurable Achievement Objectives (AMAOs) required by the U.S. Education Department as part of Title III. (Districts are required to publish their AMAO results, so having OSPI publish the results reduces districts’ reporting burden.) Making the results public provides more transparency, and simply making the results public often has a positive impact on student outcomes.
- OSPI will determine the percentage of ELL students who are on track to meet state content standards based on their level of English proficiency. This information will help educators in districts and schools know what students need extra help in specific content areas.

⁵The math and science tests were available in Spanish and Russian for the first time in 2009 but responses must still be made in English. Data show that students with “advanced” level of English but who are not yet proficient do not know enough English to meet standard on the content assessments. When data become available on ELLs’ initial English level, the exclusion period will vary based on their incoming English ability. The recommended exclusion period is 2 years for students with advanced English (Level 3), 3 years for students with intermediate English (Level 2), and 4 years for students with limited/beginning English (Level 1).

⁶The date of entry into a U.S. school is captured in the home language survey related to the ELL program. However, it does not have information about how long a student has been attending a U.S. public school. Although some ELLs are highly mobile and do not attend school the entire year, for the sake of simplicity the 3-year period includes time when students are not enrolled.

⁷A recent analysis of ELL students found that more than half passed the state test by the end of their third year in the program. Moreover, in 2003 OSPI conducted a survey of stakeholders (e.g., principals, ELL staff, parents) to determine their views about the amount of time to delay counting test results. Most said three years was the right level of delay (some said more years, others said fewer years).

⁸OSPI currently provides an annual report with ELL data, but it does not post the results on the Report Card in the same way that other state test results are posted. Further, WLPT data files are not available for downloading like the other sets of data.

Improvement by High Performers

Schools and districts that perform at very high levels are not able to improve much from the previous year. To avoid “penalizing” these schools and districts for a lack of improvement, ratings for the improvement indicator will be excluded from the Index calculations when performance is very high. Without this policy, schools and districts with nearly all of their students achieving Level 4 (exceeds standard) on the state tests and graduating nearly all their students would not be able to achieve a rating above 4 (little or no improvement). Specifically, the improvement indicators will be excluded when computing the Accountability Index when a Learning Index reaches 3.85 out of 4.00 and remains at or above that level for two consecutive years. (A school or district needs to improve by more than .15 to receive a rating of 7, which is impossible when their Learning Index is at or greater than 3.85.) The first year the Index falls above 3.85, the school/district will get a rating based on their improvement. If the Learning Index stays at or above 3.85, the maximum rating is not possible, so the indicator will not be calculated. Of the schools and districts with reportable data (at least 10 students assessed), two schools reached this level on the Learning Index in 2006 (one in both reading and writing and the other in writing), no district reached this level in 2006, and no school or district reached this level in any subject in 2007. Hence, it would not have applied to any school or district in either year.

The same policy applies to the extended graduation rate outcome. A school or district will not have its extended graduation rate results counted when computing the Accountability Index when the rate reaches or exceeds 94% and remains at that level for two consecutive years. (The graduation rate must improve by more than 6 percentage points to earn a rating of 7.) Of the schools with graduation data, 11% had a rate that was at least 94% in two consecutive years.

INTEGRATING THE FEDERAL AND STATE ACCOUNTABILITY SYSTEMS

Federal law requires states to have a single accountability system. Many states combine their own state accountability system with the federal NCLB system, and some states use a separate set of data and use it for state accountability purposes. Federal law also requires schools and districts to be assessed annually to determine if they make AYP. This involves accountability for nine groups of students in reading, math, and one more indicator.⁹ Accountability for performance at the student subgroup level is widely viewed as a positive feature of the law.

Stakeholders across Washington and in many other states believe the current federal system is flawed in many respects. Many believe it is overly complex, lacks transparency, does not provide an accurate or valid representation of school and district quality, and imposes sanctions that do not make sense. As a result, stakeholders provided advice on how the Accountability Index can be used to make AYP decisions. They also suggested changes in the consequences schools and districts face when they do not make AYP and when they make AYP while in “improvement” status. These suggestions, if approved, will increase the rigor and comprehensiveness of the accountability system as a whole, correct the flaws in the current system, and make it more valid, transparent and easier to understand.

⁹The nine groups are “all” students, five race/ethnic groups, two program groups (students with disabilities and English language learners), and students from low income families. In Washington, the unexcused absence rate is the additional indicator at the elementary and middle school levels, and the extended graduation rate is the additional indicator for high schools.

The Accountability Index incorporates three policies to increase transparency and make the results easier to understand. First, there must be at least 10 students assessed or in a graduation cohort for a cell of the matrix to be rated. This is consistent with the current state reporting policy. Second, all students are counted, including those who enter school after October 1. This is also consistent with the current state reporting policy. The current AYP policy uses the “N” of 30 students and does not include the test results of students who enroll after October 1. Using a larger N and excluding results for students who are enrolled for only part of the year means (1) the results are not consistent with the “real” results reported by the state, and (2) fewer student groups are held accountable.

A third policy to increase transparency and simplicity relates to the “margin of error.” Under the current AYP policy, the state uses a margin of error at the 99% level to avoid “false positives” to ensure a school/district is not identified as not making AYP based on the possibility of measurement error. This policy is justified when a single missed target results in the entire school not making AYP. However, this further obscures the true results by artificially increasing the percentage of students who meet standard (the fewer the students counted, the larger the extra percentage of students who are considered meeting standard). The measures used in the Accountability Index do not use any margin of error—test results are not inflated in any way and AYP results do not depend on the size of the student population.

Determining AYP

Washington will request using the Accountability Index and a similar matrix with subgroup data to create a different set of rules to determine AYP. In addition, a new set of consequences are proposed for those “in improvement.” Implementing this system requires the approval of the U.S. Education Department and on funding for struggling schools and districts that do not receive federal Title I funds.

For valid AYP determinations to be made, schools and districts must have at least 10 students in at least four of the 20 cells of the accountability matrix, i.e., at least four cells must be rated. This prevents determinations based on limited data.¹⁰ When at least four cells are rated, the following rules will be used to hold both schools and districts accountable.

- The **All students** group is held accountable using the Accountability Index. *Schools and districts do not make AYP if they have a 2-year average Accountability Index below 3.00 AND an Index that declines two years in a row.* Since the Index includes both writing and science, AYP decisions related to the entire student population are more comprehensive and include many more measures of effectiveness. The 3.00 level is still in the Fair tier but is justified if the Index has a declining trend.
- **Subgroups** are held accountable using a separate matrix that uses the same concepts as the Accountability Index. Two more subgroups (Pacific Islanders and multi-racial) are added to provide more complete coverage. However, only the outcomes used for federal accountability—reading, math, and the extended graduation rate—are used, and the two income-related indicators are combined. A “row average” is calculated for each subgroup. *Schools and districts do not make AYP if any row average declines two years in a row.*

¹⁰Schools with less than four rated cells must submit an improvement plan to OSPI for review. This is the current practice when a school does not meet the minimum number of students. Schools that do not have any assessed grades (e.g., K-2, 11-12) must also submit its school improvement plan for review.

Table 8 gives an example of a hypothetical high school with at least 10 students in each subgroup (in reality, no school has at least 10 students in every group). Up to nine ratings are based on the performance of each group in three outcomes (reading, math, extended graduation) and three indicators (achievement of all students, achievement vs. peers, and improvement).¹¹ Colors highlight ratings that are better or worse than the previous year. In this example, six groups had a row average in 2009 that was less than the row average in 2008. If the row average in 2010 declines again for any of these groups, the school would not make AYP in that group.

Table 8: 2009 Results, Hypothetical High School

Subgroup	READING			MATH			EXT. GRAD. RATE			Average rating	Change from previous year
	Met Std. (All stud.)	Peers	Improve.	Met Std. (All stud.)	Peers	Improve.	Met Std. (All stud.)	Peers	Improve.		
American Indian	4	4	4	1	5	4	1	4	4	3.44	0.33
Asian	6	4	5	4	5	4	5	4	6	4.78	0.56
Pacific Islander	4	4	4	1	5	4	1	4	4	3.44	0.22
Black	3	3	3	1	3	5	1	3	2	2.67	-1.00
Hispanic	3	4	4	1	4	4	1	4	4	3.22	-0.11
White	5	4	4	3	3	4	3	4	4	3.78	-0.22
Multi-racial	4	4	4	2	4	4	2	4	4	3.56	-0.22
Special education	2	4	4	1	4	4	1	2	1	2.56	-1.22
ELL	1	5	5	1	4	4	1	3	3	3.00	-0.11
Low income	4	5	5	1	5	4	1	4	4	3.67	0.22
Average	3.6	4.1	4.2	1.6	4.2	4.1	1.7	3.6	3.6	3.41	-0.16

NOTE: Ratings in red are less than the previous year, ratings in green are more than the previous year.

Using the modified matrix in this way has a number of benefits. It keeps the simplicity of the Accountability Index intact, uses the same metrics as the Index to provide greater simplicity,¹² provides more detailed information about subgroups, focuses on improvement from each group's baseline, relies on multiple cells when computing row average to reduce fluctuations in averages from year to year, and treats every group with equal value regardless of the size of the group.

- **The ELL group has an alternate method for making AYP** based on the results of the WLPT and the content assessments. As noted above, OSPI has developed a method to calculate the percentage of ELLs who are on track to meet the content standard when they become proficient in English.¹³ Hence, if the ELL subgroup does not make AYP via the row average, *schools and districts can make AYP if the percentage of ELLs in WLPT Levels 2 and 3 who are on track to meet standard in the content areas (i.e., reading and math) when they become proficient in English does not decline two years in a row.* This alternative method is a fairer way to hold this group accountable because it emphasizes improvement in both English proficiency and academic performance and considers English language ability when examining students' performance in the academic subjects. It also recognizes that ELL students move in and out of the group, which means there is a different group of students each year. (This group will always

¹¹The current AYP system uses of unexcused absence rates at the elementary and middle school levels. Data for these rates are not included because they are not part of the index system, and nearly all schools meet the required goals. However, they will continue to be used when making AYP determinations.

¹²The modified matrix relies on the same rules as the Accountability Index. For example, the results for all grades are combined, there must be at least 10 students to report results, the percent meeting standard is used for the achievement indicators, the Learning Index is used for the peers and improvement indicators, and the same rating system is used.

¹³ELLs are counted in WLPT Level 1 for only one year to provide an incentive to help new ELLs as much as possible.

have students who do not meet standard as new students enter the system, so this group will never have all its students be proficient in the academic subjects.)

- **Schools and districts must show continuous improvement from their own baseline over each 4-year period.** Specifically, they must reduce the combined percentage of all tested students not meeting standard across all tested subjects by at least 33% every four years, with 2010 as the baseline year. This means each school and district has its own target in 2014 and every four years thereafter (e.g., 2014, 2018).¹⁴ This is a similar concept currently used to access “safe harbor” under NCLB. Under safe harbor, a proficiency cell that does not meet an annual target is still considered to have met the target if it has 10% fewer students not meeting standard from the previous year. The 33% level is approximately the same level of improvement required to meet the safe harbor criterion over a 4-year period. This approach holds all schools and districts accountable for improvement in an equitable manner, regardless of their baseline. Those that do not meet their improvement goal by 2014 and 2018 will not make AYP that year. This level of improvement is more realistic than requiring all students to meet standard by 2014, regardless of their starting point.

Hence, the system is “compensatory” in nature—having one low rating in a matrix does not automatically result in a school/district not making AYP. The index blends performance across multiple ratings, and low ratings are compensated by higher ratings, a concept similar to how a GPA works. This is different from the “conjunctive” model now used to determine AYP. In a conjunctive model, a single missed target results in a school/district not making AYP. This is analogous to labeling a student as a failure when a single low grade occurs. The increasing level of proficiency currently required to make AYP will make it even less likely a school/district will meet the target. So in this analogy, a student would have to get higher and higher grades in all subjects to avoid being labeled a failure.

Finally, all the current data reporting requirements of NCLB will continue to be met (i.e., making public the disaggregated data for the assessments, participation, and “other indicators” for the various student subgroups). To enhance the accountability system, a more extensive set of data will be made public and may be used when conducting deeper analyses of schools and districts that do not make AYP (see Appendix H).

In-Depth Analysis Before Making AYP Decisions

Each fall OSPI will compute the Accountability Index and subgroup results for the modified matrix and then apply the rules for making AYP. All schools and districts in all tiers will be given an AYP status, not just those receiving Title I funds. The first time a school or district does not make AYP, it is in a “warning” year. To determine the AYP status of schools and districts that do not make AYP two years in a row, OSPI will first conduct a deeper analysis using additional data. After this analysis, OSPI will determine if the school/district should move into an “improvement” step and be required to take certain actions (see Consequences section below). A deeper analysis is needed to make AYP decisions because “test based accountability indices should trigger further

¹⁴ For example, a K-6 school with only 25% of its students meeting standard in 2010 had 75% not meeting standard. Since 33% of 75% is 25%, this school needs to have 50% of its students meeting standard by 2014 (25% plus 25%).

investigation rather than portend immediate sanctions.”¹⁵ Thus, schools and districts not making AYP two years in a row do not automatically enter “improvement” status.

OSPI staff will analyze a comprehensive set of quantitative and qualitative data and may need to contact the district and ESD to get more information. After completing its analyses, OSPI will confirm that a school/district will move into improvement status if it has not made AYP two years in a row as long as there is no compelling evidence not to do so. Districts can appeal to have OSPI review any AYP-related decision using the process described below.

The data to be reviewed fall in four general categories. The list below provides examples in each (see Appendix B for a complete list).

- **Contextual Data**
 - Type of school
 - Changes in student population
 - Programs served by the school
- **Assessment Results (State content assessments/WAAS/WLPT)**
 - Trends over multiple years for each subject area
 - Subgroup trends (e.g., race/ethnicity, ELL, special education)
 - Results for students who have been enrolled for at least two years
- **Teaching and Learning Issues**
 - Teacher education and experience levels
 - Student/teacher ratio
 - Recent changes in leadership (key central office staff and principals) and teachers
 - Alignment of curriculum and materials across grades and with state standards
- **Other Data**
 - Graduation and dropout rates for subgroups
 - Funding from local levies/bonds and outside sources
 - Problems with data that generate the index (e.g., reporting errors related to graduates)
 - Participation rates for all subgroups
 - Other indicator data (unexcused absence and graduation rates) for all subgroups
 - Perception survey results

For a school/district that has not made AYP *four or more* years in a row, OSPI will review the data and may determine that it would benefit from a significant amount of additional support. At this point, the school/district may be moved into “Voluntary Action” for at least two years and are offered substantially more help. However, if the extra assistance is not accepted and significant improvement does not occur during the two-year period, the school/district will move into “Required Action” status and a binding corrective action plan will be established between the district and the state, if authorized by the Legislature.¹⁶

¹⁵Recovering the Promise of Standards-Based Education (2008), Education Policy Briefing Sheet, National Academy of Education, www.naeducation.org.

¹⁶ESHB 2261, passed by the 2009 Legislature, contains language on this issue.

Consequences

NCLB currently requires schools and districts to undergo increasing levels of “sanctions” if they do not make AYP over an extended period of time. NCLB also requires schools and districts that are in an improvement step to make AYP two years in a row in order to exit improvement status.¹⁷ Washington state law currently prohibits the use of some consequences authorized by NCLB. For example, the state has no authority to require schools/districts to accept state help, and it has no authority to require corrective action or restructuring. However, the state legislature will likely approve giving the state this additional authority during its 2010 session.

Many stakeholders believe the federally-mandated sanctions have flaws that need to be corrected. For example, students must be allowed to transfer to another school before their current school is required to provide additional services to help these students. Schools must also allow all students to transfer, including those performing well. Even when a school in “improvement” status performs well and makes AYP, all the sanctions must remain in effect. Finally, the consequences do not apply to non-Title I schools that do not make AYP, even though in Washington these schools outnumber Title I schools, enroll more students, and are more likely to have low Index results (see Table 9). In fact, almost half the students in the state attended non-Title I schools that did not make AYP in 2008, and a large number of these schools are in “improvement” status but evade the teeth of the accountability system because they are not required to face any of the NCLB sanctions.

Table 9: Index Results for Schools Not Making AYP in 2008

Tier	Title I Schools		Non-Title I Schools	
	Number	Students enrolled	Number	Students enrolled
Exemplary	5	2,445	8	4,973
Very Good	27	13,211	31	30,637
Good	165	82,518	217	184,184
Fair	326	157,312	333	227,112
Struggling	56	19,184	83	24,388
Total	579	274,670	672	471,294

As a result, Washington proposes using a different set of consequences that reflect common sense changes to the current NCLB rules. If approved, these consequences will apply to all schools and districts, not just those receiving federal Title I funds.¹⁸

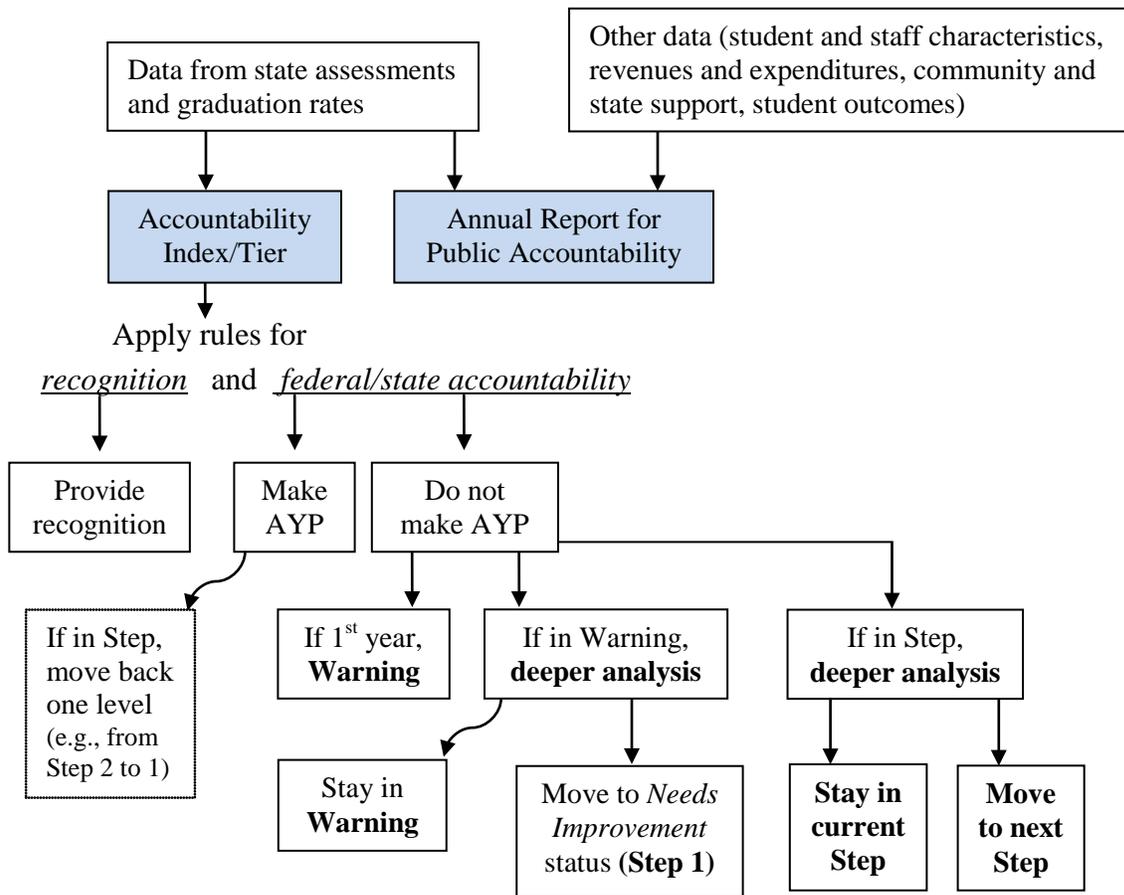
¹⁷The sanctions include providing students with the option to transfer to another school at district expense (Step 1), providing supplemental educational services, such as tutoring, to low income and low achieving students (Step 2), undergoing “corrective action” that requires either hiring external advisors to help improve the school, extending the school day or year, or making curriculum and instruction changes (Step 3), and preparing and then implementing a “restructuring” plan that would change the form of governance (Steps 4 and 5). Each year a school “in improvement” does not make AYP in the same content area, it moves to the next step, and each step along the way includes the sanctions of the previous step (e.g., Step 2 requires the sanctions of both Steps 1 and 2). District improvement involves only two steps, and moving into district improvement and to the next step is more complicated than for schools. A school or district in improvement does not “move back” a step or forego sanctions if it makes AYP. If it makes AYP two years in a row, it exits improvement status and sanctions no longer apply.

¹⁸This assumes the state can provide funding for assistance to non-Title I schools and does not lose any Title I funding due to federal “supplant” rules.

- Schools/districts not making AYP for the same reason (e.g., same subgroup) in consecutive years move into “improvement” unless there is a compelling reason not to based on the results of a deeper review (see below).
- If the reason for not making AYP is due to the performance of a different group than a group responsible for not making AYP in the previous year, the school/district does not move to the next step of the process.
- School choice and supplemental educational services must be made available to the students in the subgroup(s) whose results are responsible for the school not making AYP. (Currently it applies to all students in the school, even if they are in a group that has performed well.) All service providers must be “highly qualified” and the district is allowed to be a provider.
- Those in improvement status that make AYP move back a step (e.g., from Step 2 to Step 1). Those making AYP two years in a row exit improvement status. This allows a gradual withdrawal of state support over time. (Under the current AYP rules, schools and districts in “improvement” must make AYP in two consecutive years to exit this status entirely, and no credit is given for making AYP in one year.)

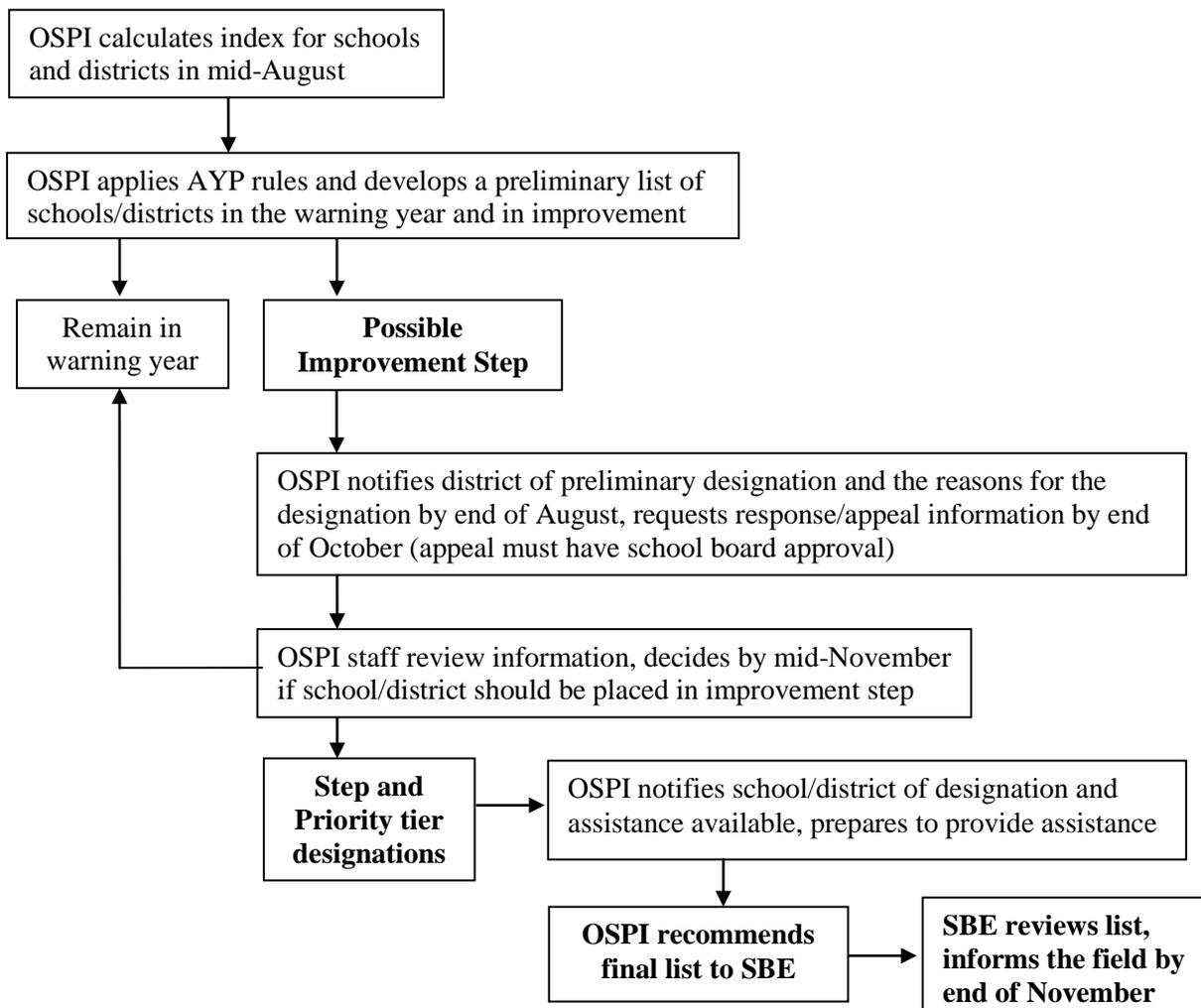
Figures 8 provide an overview of the process for making AYP decisions.

Figure 8: Process for Making AYP and Step Determinations



In terms of timing, OSPI will compute the accountability index and apply the rules for making AYP by mid-August of each fall. The more intensive analysis of the data will then occur, and by the end of August, OSPI will make a preliminary decision about placing a school/district in an improvement step, notify the district of the possible designation, and give the reasons why this designation is possible. During September and October, the district/school is given a chance to appeal the designation by providing more information (e.g., what explains the low index results, other favorable results, feeder school information, district assessment results, personnel changes, type of interventions made to date) and any plans being made for the future. Any appeal to OSPI needs local school board approval. After receiving the August high school assessment results, OSPI will review the information again to see if the new data provides any changes to the results. By mid-November, OSPI will recommend to SBE the schools and districts that should be placed in an improvement step and the rationale for the designation. SBE will review the list, receive comments, finalize the list, and inform the schools and districts by the end of November about how they can respond to the designation. Figure 9 provides a flow chart of this process.

Figure 9: Timing for Identifying Schools and Districts in an Improvement Step



Schools/districts placed in a Step will be offered additional state support that will be tailored to meet their specific needs. Appendix C provides information about the current OSPI system of support. Appendix F outlines some initial concepts for assistance as described by Mass Insight.

The deadline for requesting changes in how AYP is calculated in 2010 has passed, so the proposed system of rules will not be used to make AYP determinations this year unless granted a waiver by the U.S. Education Department. However, the Accountability Index will still be calculated and made public so the results of the matrix can be used for educational purposes and by OSPI in its assistance decisions. Index results will also be calculated retroactively and used for recognition purposes. Making the results public provides a more valid picture of school/district performance than the current AYP results, and it introduces the concepts to the various stakeholders prior to its full implementation.

Advantages of the State System

Using the Accountability Index and the modified matrix for subgroups when making AYP decisions has a number of advantages compared to the current system.

- The proposed system is *more valid and less volatile* because it is based on the performance of all students in more subjects across all grades in a school/district rather than using results for individual grades where students change from one year to the next. Moreover, it counts all students the same number of times.¹⁹ Using five labels provides more accurate judgments about performance than the current “Pass/Fail” system.
- The proposed system is *more inclusive and comprehensive* because it uses a smaller minimum number for reporting (10 students across the entire school/district), includes the results of all students (not just those who are continuously enrolled), includes both writing and science.
- The proposed system is *more transparent and easier to understand* because it does not include a margin of error, the benchmarks are the same over time and for the different subjects at both the school and district levels, there are fewer rules and subgroup “cells” to examine, and both schools and districts have the same minimum number for reporting results.
- The proposed system *addresses unintended side effects* created by the current system. Including writing and science helps prevent a narrowing of the curriculum; using the Learning Index to measure performance across the range of assessment results reduces the focus on students who perform close to the proficiency cut point at the expense of students who are farther above and below that level; and using rigorous benchmarks in a compensatory manner encourages the state to maintain high performance standards and does not provide an incentive to lower these standards so all students can be counted as proficient by 2014, a target viewed as unrealistic if standards are kept high. (Two recent studies found that Washington state has some of the nation’s toughest AYP requirements, resulting in a relatively high percentage of schools not making AYP and “needing improvement.”²⁰)
- The proposed system has *more appropriate consequences* and provides *stronger intervention* measures when improvement does not occur.

¹⁹In the current AYP system, a single student can be counted in five groups, while others may be counted only twice.

²⁰See “*The Accountability Illusion*,” Thomas B. Fordham Institute (February 2009) and “*Schools Struggling to Meet Key Goal on Accountability*,” Education Week (January 7, 2009).

Table 10 summarizes the differences between the current federal AYP system and the proposed system that, if approved, will be used for both federal and state accountability.

Table 10: Comparison Between Federal and Proposed State Accountability System

Issue	Current Federal System	Proposed State System
Number of Cells	37 cells for elementary and middle schools, 45 cells for high schools, 119 cells for a district	20 cells for districts and high schools, 16 cells for elementary and middle schools, participation and unexcused absence rates are tracked separately
Unit of Measurement	Three groups of grades (3-5, 6-8, 10)	Combine all grades together to look at the whole school and district regardless of their grade configuration
Outcomes Measured	Reading, math, extended graduation rate, unexcused absence rate, participation rate	Reading, math, science, writing, extended graduation rate, unexcused absence rate, participation rate
Indicators Measured	Achievement compared to a “uniform bar” for nine categories of students: all, five racial groups, ELLs, special education, low income	Achievement of low income and non low income students (shows the achievement gap), achievement vs. peers, improvement from previous year, two more subgroups added to make 11 categories of students, separate matrix used for subgroups, improvement measured from group baseline
Assessment Measure	Proficiency (meets standard)	Proficiency (meets standard) for two indicators and full range of possible results (Learning Index) for two indicators
Goals	All are proficient by 2014	Requires continuous improvement from own baseline, goals set every four years beginning in 2010
Students Included	Only those continuously enrolled from October through testing period in May	All students regardless of period of enrollment
Minimum Number of Students Required	At least 30 students in a grade span; schools and districts may have a different N	At least 10 students in a school and district; schools and districts have the same N
Model	Conjunctive: Not meeting the goal in any cell generates a negative label	Compensatory: Average of all cells determines tier and descriptive label
Margin of Error	Proficiency level adjusted using 99% confidence level to avoid Type I errors (false positives) due to use of a conjunctive model	No margin of error used due to the use of a compensatory model
ELL Students	Reading and math results included after one year of U.S. enrollment; WLPT results not required to be made public	Test results in four subjects included after three years of U.S. enrollment (results are still reported); additional accountability via reporting and feedback using the WLPT data

Consequences	Five steps for schools, two steps for districts; same consequences for all schools, regardless of magnitude of need; apply only to Title I schools and districts; weak intervention authority	Used for recognition; sanctions and support depend on deeper analysis; applies to all schools and districts, even those not receiving Title I funds; strong intervention authority
Exiting Improvement Status	Must make AYP two years in a row	Can gradually move “backward” out of improvement status by making AYP

RECOGNITION

Providing recognition is another “consequence” of the accountability system. Several guiding principles apply to the new recognition system. It should be transparent and simple to understand, provide multiple ways to reward success, rely mainly on criterion-referenced measures, and be based on reaching challenging yet attainable goals. The recognition system approved by SBE is consistent with these principles and is based on a “theory of change” that people are motivated more by achieving success than by avoiding negative labels, blame, and guilt.

Providing recognition using data from the Accountability Index results is considered “Phase I” in the implementation of the new state accountability system. Theoretically all schools can be recognized because it is a criterion-based system. OSPI is developing its own criteria for providing recognition using Index results. The two agencies are working together to announce the winners in Spring 2010. The federal government also provides funding for four types of awards, primarily for schools receiving Title I funds. Appendix A provides more information on the current forms of federal and state recognition.

SBE will give recognition for “Outstanding Overall Performance” in seven areas using the following criteria.

- (1) For the **Index**, the 2-year average must be at least 5.50, at least 10 cells of the matrix are rated each year, and there are fewer than 10% students designated as gifted each year.
- (2-5) For **language arts** (reading and writing combined), **math**, **science**, and the **extended graduation rate**, the overall (column) 2-year average is at least 6.00, at least 2 of the 4 cells in the column are rated each year, and there are fewer than 10% students designated as gifted each year.²¹
- (6) For the **achievement gap**, there are at least 10 students in at least 2 of the 5 outcomes (columns) in both of the income-related cells (non-low income and low income), there is no rating of 1 in any income-related cell or peer cell, there is no more than a 1-point difference in the rating between the two income-related cells,²² and there are fewer than 10% students designated as gifted each year. Each of the above criteria must be met two years in a row.

Each of the above six recognition areas requires fewer than 10% of the students to be gifted each year. Statewide 3% of all students received this designation in 2008, so schools with 10% or more gifted students have unusually high concentrations of the most capable students. The exclusion

²¹For language arts, both reading and writing must have a 2-year average of at least 6.00 and at least 2 of the cells rated in each column each year.

²²For example, if the reading non-low income cell is rated 5, the reading low-income cell could be rated no lower than 4 and no higher than 6.

criterion prevents a school from receiving recognition because it will likely have much higher than normal ratings. To ensure these types of schools are eligible to receive recognition, a seventh recognition area was established.

(7) For **schools with gifted programs** (those with at least 10% gifted students), the top 5% of schools in grade band—elementary, middle, high, and multi-level—will receive recognition based on the 2-year average peer ratings in all four subjects.²³

The highlighted sections of the Accountability Index shown in Table 11 relate to the areas where the SBE recognition will be given.

Table 11: State Board Recognition Categories

Indicator	Outcome					Average
	Reading	Writing	Math	Science	Ext. Grad.	
<i>Non-low inc. achievement</i>	Compare ↑	Compare ↑	Compare ↑	Compare ↑	Compare ↑	
<i>Low inc. ach.</i>	↓	↓	↓	↓	↓	
<i>Ach. vs. peers</i>						Top 5% gifted
<i>Improvement</i>						
Average	6.00*		6.00*	6.00*	6.00*	5.50*

* Minimum 2-year average rating required to earn recognition.

Table 12 shows the number of schools that would have been recognized by SBE in 2008. In all, 99 schools—nearly 5% of all schools statewide—would have been recognized if these criteria were in effect at that time. Elementary schools represent the largest percentage of schools statewide, so they would have been recognized most often. Eight of the 99 were alternative schools. A total of 50,500 students were enrollment in the 99 schools in 2008 (5% of all students), with an average size of 510 students per school. Figure 10 shows the percentage of schools that met the criteria in each area. Very few schools would have received recognition in math, science, the index, or for having closed the achievement gap in all areas.²⁴

²³Results for the peer indicators control for the types of students attending the school (the percent gifted, low income, ELL, special education, and mobile). This ensures schools with the highest concentrations of gifted students do not automatically receive this form of recognition.

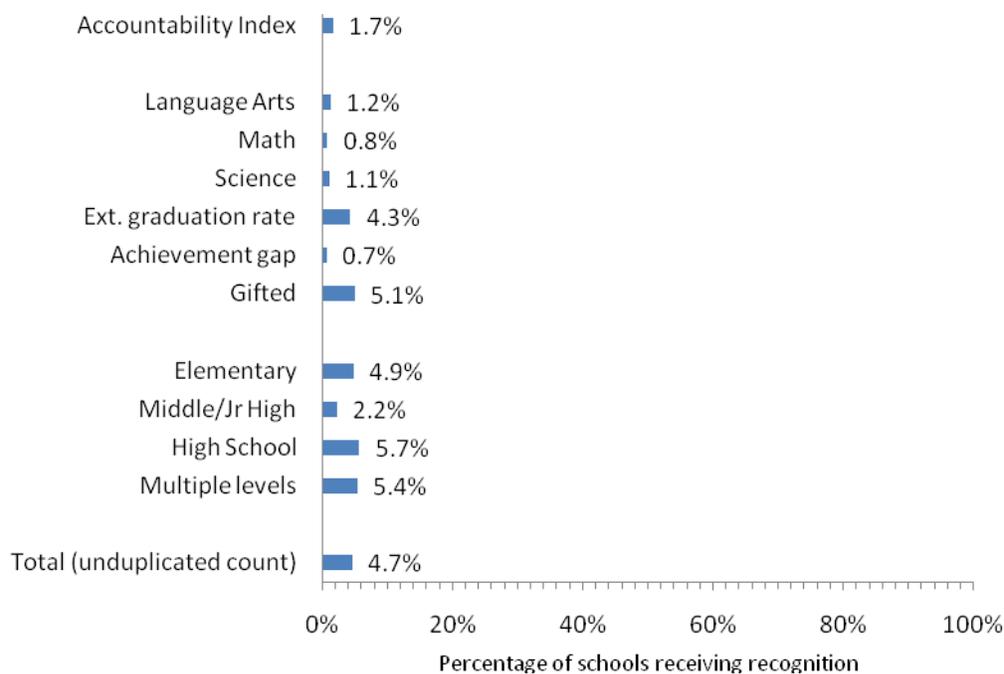
²⁴The uneven results occur because recognition is given based on a set of criteria rather than on a percentage basis (a norm-referenced approach) and because of differences in the relative difficulty of the assessments.

Table 12: Distribution of Schools Meeting State Board Recognition Criteria, by Grade Type (2008)

	Elementary	Middle/ Jr. High	High	Multiple Levels	Total
Total Recognized*	52 (1)	8 (3)	23 (0)	16 (4)	99 (8)
Number of All Schools	1,059	359	400	298	2,116
Percentage of All Schools Recognized	4.9%	2.2%	5.7%	5.4%	4.7%

* Number of alternative schools are shown in parentheses

Figure 10: Percentage of Schools Meeting “Outstanding Overall Performance” Criteria (2008)



* * * * *

The system needs to be flexible to accommodate changes. Revisions to the assessment system and cut scores, NCLB requirements, graduation requirements, graduation rate formulas, and content standards may have an impact on some measures and could require changes to the system. As data systems improve statewide and more information becomes available, more data will be available to aid our understanding about other student outcomes and factors that could contribute to them (e.g., post secondary outcomes, local levy data, student-teacher ratios). Other more sophisticated analyses could be used as well (e.g., growth models).

A number of issues must still be resolved before the accountability system can be implemented effectively. Various OSPI and SBE activities need to be integrated and aligned with one another to avoid duplication and confusion. These relate to how the index is used for AYP-related decisions (e.g., what determines “making AYP,” how and when assistance and other consequences occur), how and when recognition occurs, and how index results are represented and made available to the

public. Finally, since the Accountability Index provides only a snapshot of school and district performance, additional information needs to be made available to educators, policymakers, and the public to provide a more complete picture of the school and district context, resources, and student outcomes. This includes data for “reciprocal accountability” to recognize the role the community and state play in supporting education.

The appendixes provide more additional information related to the current and proposed accountability system.

- Appendix A provides more details about how the index is calculated and used.
- Appendix B provides information about the data that can be used to conduct in-depth analyses of schools and districts that do not make AYP.
- Appendix C gives an overview of the current state assistance system that is funded primarily by the federal government.
- Appendix D provides information on the State Board’s legislative mandate to create an accountability system.
- Appendix E provides language related to the Board’s positions on accountability.
- Appendix F describes the initial concepts of the Priority Tier presented by Mass Insight.
- Appendix G provides an overview of the current AYP rules and sanctions.
- Appendix H provides a list of additional data that can be used to supplement the Accountability Index data.
- Appendix I lists the names of those who provided advice and feedback during the development of this proposal.

President Obama has cited his concerns about the current AYP system, such as its failure to acknowledge when significant improvement has taken place. The President believes we can improve and evolve in all aspects of education, including how AYP decisions are made. Stakeholders in Washington state join him in his desire to change NCLB and the current AYP system in order to hold our schools and districts accountable in a more rigorous, more valid, and more transparent way. Washington has taken the initiative to lay out a new accountability model using a set of outcomes and indicators. Use of these data for making AYP decisions addresses fundamental weaknesses in the existing system and encourages the state to maintain rigorous content and performance standards. Stakeholders in Washington believe this new system and the use of the new Accountability Index paves a way forward to increased clarity and accuracy in assessing our education system, thereby offering educators and stakeholders a transparent means to ensure each and every student receives an excellent and equitable education.

APPENDIX A

INDICATORS AND OUTCOMES

This appendix provides detailed information about how the indicators and outcomes were selected, how the Accountability Index is calculated and used, and statistical results that were prepared to estimate the effect of the proposed recommendations.

OVERVIEW

One of the guiding principles for the accountability system is the use of multiple measures. After much discussion, the advisors (see Appendix I) recommended using five outcomes and four indicators, resulting in a 5x4 matrix with 20 measures. Other indicators and outcomes were discussed besides the state tests 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.

Each cell of the matrix is rated, and the Index is 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 test results that generate the ratings. By using the simple average, schools without data for some indicators are still included in the system and a separate system is not needed for different types of schools.

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 benchmarks over time and the use of measures (e.g., standard deviations) that vary by subject. This means that benchmarks are fixed for all subjects and recognition is given when schools/districts meet certain criteria (there is not a limit to how many 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 knows in advance what is required to achieve a rating and receive recognition, regardless of how others perform. Using fixed criteria also encourages cooperation among educators because they are not competing with one another for recognition. Finally, it allows all schools and districts to reach the Exemplary tier.

The advisors discussed other types of analyses that could provide more accurate results (e.g., hierarchical linear modeling, value-added growth models, creating norms through z-scores). 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.

The Index simplifies a complex set of measures that reflect different constructs and rigor. Each assessment in each grade differs from the others on the basis of cut scores to reach standard and the material being tested. As a result, the distribution of student performance differs on the various assessments. There is also measurement error in the data being used to generate the ratings. For example, the percentage of low income students (especially at the high school level) is known to have some reporting error. Moreover, some schools have all 20 cells rated while others have

relatively few, and the graduation rate applies only to high schools and districts. The Index becomes harder to change as the number of rated cells increase. Further, the “peers” result is a norm-referenced measure while the others are criterion-referenced. Finally, the ratings may reflect little difference in performance (e.g., 49.9% meeting standard is essentially the same as 50% but receives a different rating). Thus, the Index is analogous to a grade point average (GPA) that is computed for a student who takes different courses with different teachers, each with their own set of expectations, grading scales, and standards. The Index, like the GPA, is a simplified estimate of student performance (although unlike the Index, a GPA is cumulative over time) and is better understood when the underlying details are examined more closely.

Comparison with Current AYP Methods

All stakeholder groups believed that AYP results do not provide a valid measure to evaluate schools and districts for recognition and additional support. While the rules the state adopted to calculate AYP conformed to the federal requirements, nearly all stakeholders believe the federal system is too complex, has too many adjustments, and is neither transparent nor fair in its accountability determinations. As a result, AYP results are increasingly being ignored. Here are the main problems that have been identified:

- AYP has different goals for reading and math at three different grade levels, and the goals change over time. The goals are the same for every student group, regardless of their starting point. As a result, some groups have great difficulty reaching the goals, even when substantial improvement occurs.
- The AYP results do not reflect the results reported by the state. 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.
- Determining if a school/district meets the annual assessment targets involves complicated calculations of a “proficiency index.” 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. As a result, computing school results can be very complex and therefore lack transparency.
- AYP has unrealistic goals and is almost entirely punitive in nature. All students must be deemed “proficient” by 2014. Schools must meet up to 37 goals each year, and districts must meet as many as 119 goals.²⁵ Not meeting just one goal leads to negative consequences and labeling, and not making a goal in two consecutive years puts a school in an “improvement step” that requires an increasing number of sanctions. These sanctions are the same regardless of how many goals are missed and by how much.
- All students a school that “needs improvement” must be allowed to transfer to another school, including those in groups that meet the goals, with transportation costs paid by the district. Allowing students who are performing well to transfer can reduce the school’s academic performance even further. Title I funds must be set aside to provide transportation to students that request it, which limits the use of these funds to help those who are eligible to receive Title I services.

²⁵Federal regulations issued in October 2008 added eight more goals for high schools and districts due to the requirement to use disaggregated graduation rates when determining AYP.

- AYP focuses only on reading and math at the expense of other subjects, which has resulted in a narrowing of the curriculum.
- AYP looks only at students who meet standard, which 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 AYP system for *districts* has different rules than the system for schools. As a result, district AYP results sometimes are 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 currently the major determinant in its AYP results—in 2008 all but one district in improvement status enrolled more than 1,000 students.

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) has a total of 16 students with assessment results and will 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. The district accountability system is essentially the same as the system for schools, which makes it relatively easy to understand and compute. Finally, using Index data to determine AYP does not penalize the state for maintaining high and challenging performance standards (defined as “internationally competitive” by the Legislature). There is no incentive to lower the cut scores or definition of proficiency in order to ensure all students become “proficient” by 2014.

Two measures used in the current AYP are not used when calculating the Index. The unexcused absence rate—the “other indicator” for elementary and middle schools—is not used because it is not measured the same way across the state. Very few schools have not met AYP due to this measure. Participation rates, which are required to be at least 95% to make AYP, are not used because virtually all schools and districts have exceeded this rate by a wide margin. These two measures will still be calculated and reported to meet federal accountability requirements, and schools and districts that do not meet the current goals will not make AYP in these areas.

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 five “tiers” based on their index number (4.00 is the middle of the 1-7 scale). The method used to determine the cut scores and names for the various tiers is discussed later in this appendix.

- Those with the highest index numbers, from 5.50 to 7.00, are in the “Exemplary” tier.
- Those with an index of 5.00 to 5.49 are in the “Very Good” tier.
- Those with an index of 4.00 to 4.99 are in the “Good” tier.
- Those with an index of 2.50 to 3.99 are in the “Fair” tier.
- Those with an index of 1.00 to 2.49 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 do not receive ratings for every cell of the matrix, which can affect their index number. Moreover, schools that administer assessments with lower scores overall (e.g., math and science) tend to have a lower index score than those that do not. For example, schools serving grades 5, 8, and 10 give the science test, 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. Hence, the Index is only comparable across schools serving the same grades. In addition, the Index does not reflect how close a school is to the benchmarks—small differences in results can generate different ratings (e.g., 89%=6 and 91%=7). Moreover, schools serving very few students may have more volatile ratings from year to year. Finally, the lack of vertical alignment of the assessments in the same content area presents another complicating factor when making comparisons across schools serving different grade levels.

Given the different types of schools being rating and the complexity of the variables used to generate the rating, school results are best reported in groups of 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 serving multiple grade levels* (e.g., K-8, K-12, 6-12)
- *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 they have fewer than 10 assessed students).

CALCULATION METHODS

To calculate the assessment-related measures, student-level data are used and aggregated to the school and districts levels. This provides more accurate results than using aggregated school and district results available to the public. Using student-level data also allows for the aggregation of results from the grade level that would be suppressed because fewer than 10 students were assessed. Results are only suppressed when there are fewer than 10 students assessed in the *combined grades*, and they are not computed when there are less than five students in a cell to avoid having results for a small number of students affect the index results.²⁶ Students who take alternate assessments (WAAS) are included in the calculations, and there is no restriction on the percentage of these students who can meet standard.²⁷ Students with disabilities that meet standard

²⁶Very 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 it is in the Struggling tier on a consistent basis, the school could be placed in Voluntary Action. Some schools have grade configurations that do not include some tested subjects (e.g., a K-4 school does not give any science test), but there may still be a small number of students that are assigned a grade where the test is given (e.g., a 5th grader attending a K-4 school). When this occurs, the results for the few students are not included in the school index calculation to prevent them from generating another set of indicators for that subject. The results for these students are still included in the district results.

²⁷Although there are no limits to the number of students that can take the WAAS, NCLB limits the number of students that can count as proficient via an alternative performance standard at the district level to 1% of all assessed students in the district. In Washington, this applies only to the WAAS Portfolio. In contrast, the state accountability system encourages appropriate testing of students with disabilities and does not penalize a district if it is able to have a higher percentage of students meeting standard.

on the state test at Level 2, as stipulated in their IEP, are counted as if they met standard at Level 3. Students who are not continuously enrolled are included in the calculations. Students who previously passed the assessments are also included in the calculations (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 are exempted from taking the assessments (i.e., those with excused absences and medical exemptions, first-year ELL students, home-based and private school students) are not included in the calculations. ELL students whose results are exempted from Index calculations but who meet standard are included in the Index calculations. This provides an incentive to help ELL students meet standard as soon as possible. If students who met standard in a previous year do not have their level available in a student-level database, they are considered to have performed at Level 3. For grade 10 results, the highest score achieved through the August testing period is used (this includes results for students who met standard in grade 9 and any retakes that result in higher scores).²⁸ Results during grades 11 and 12 will be considered when making accountability decisions for schools and districts not making AYP several years in a row. This recognizes the extra effort that districts make to help students who are in danger of not graduating unless they pass the required assessments. Finally, students who are required to take a test but receive no score (e.g., refused, invalidated results, unexcused absence) are considered performing at Level 0.

When computing the Index, all the ratings are counted equally (they are not weighted). Achievement is considered most important, so it represents three of the four indicators. Averages are computed only for cells of the matrix that have data (e.g., an elementary school has no graduation data, so the averages for the indicators use only assessment results). District results are based on OSPI's aggregation rules, so the district results do not include results from correctional institutions, independent 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 any district. Finally, Index results are based on ratings generated for a single year rather than averages over multiple years. This provides greater simplicity and avoids 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).²⁹ However, the annual results will be viewed across multiple years before drawing conclusions about schools or districts.

INDICATOR 1: ACHIEVEMENT OF NON-LOW INCOME STUDENTS

This indicator examines outcomes for students who are not identified as living in low-income families (i.e., not eligible for a free or reduced-price meal). The five outcomes are the four subjects tested statewide (reading, writing, math, science) and the extended graduation rate (see the explanation below on how this rate is calculated). Using results for non-low income students separate from those for low-income families (used in the second indicator described below) means no student is double counted. Under the current AYP rules, some students are counted as many as five times, while others are counted only twice. Counting students once reflects the belief that all students have equal value and no group of students is more important than any other group.

²⁸Any passing results from grade 9 are not counted until grade 10.

²⁹In 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 percent meeting standard includes both the results of the regular tests and the WAAS, which is given to students with disabilities. Subgroups results (for the various race/ethnicity groups, ELL, students with disabilities, gender) are reported as required by NCLB and are used when examining the schools and districts that do not make AYP two years in a row or are in an improvement step (see Appendix B). Results for low-income students are used in aggregate in a separate indicator described below.

The benchmarks and ratings for this indicator in 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:
 - 90 - 100%7
 - 80 - 89.9%6
 - 70 - 79.9%5
 - 60 - 69.9%4
 - 50 - 59.9%3
 - 40 - 49.9%2
 - < 40%1

- Achievement on the *extended 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%7
 - 90 - 95%6
 - 85 - 89.9%5
 - 80 - 84.9%4
 - 75 - 79.9%3
 - 70 - 74.9%2
 - < 70%1

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.³⁰

A school/district must have at least 10 students for results to be counted in the matrix. The minimum number used by OSPI is 10, but it applies at the grade level. Using an N of 10 for a *school* means that very small schools will now be included in the accountability system because

³⁰The 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 the test is not being taken seriously by some students, resulting 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 is to treat it like the other subjects. Keeping the same rating system as the other subjects also keeps the system consistent and simple, and it 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 conveys the wrong message about the importance of students learning science concepts and content.

they are more likely have at least 10 students assessed in 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 are fewer 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 at least 30 continuously enrolled students must be assessed when making decisions about sanctions.³¹ 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.

INDICATOR 2: ACHIEVEMENT OF LOW INCOME STUDENTS

This indicator focuses on the performance of low-income students, i.e., those who are eligible to receive a federally-subsidized meal (free or reduced-price lunch). This indicator uses the same five outcomes as the non-low income achievement, and the same benchmarks and rating scales are used as well. The percentage of low-income students in high schools is often higher than what is reported, but this measure is still the best available proxy for socioeconomic status.

Having a separate indicator for low-income students highlights how well these more vulnerable students are performing. Much research has shown that student achievement is highly correlated with a family's socioeconomic status. 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. The federal Title I program focuses largely on helping these students. 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.³² This does not imply that a 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.

The results of both the non-low income and low-income indicators are provided on OSPI's Report Card. Unlike the current AYP results, the performances of both groups are not adjusted in any way (i.e., no margin of error or exclusions of non-continuously enrolled students). However, the Report Card does not show how income affects performance for the different race ethnic groups. Further analysis shows that the difference in performance between the non-low income and low income students is very pronounced. Figures A1 and A2 show the average reading and math scale scores in 2007. In both subjects in all grades, there is a large difference between the scores of the two income groups. In reading, every non-low income group had an average scale score that meet the standard (400), regardless of the grade or race/ethnic group. The differences between the two income groups are more pronounced in math. The average math scale score gradually declines as

³¹In the past, the N was larger (40) for the ELL and special education groups and smaller (10) for the "all students" group. The N is now required to be 30 across a grade band (3-5, 6-8) and in high school. At the district level, the N can be higher when there are more than 3,000 students enrolled.

³²The 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. The federal Title III program targets ELL students, and the federal IDEA targets students with disabilities.

the grades increase. While there are still differences among the race/ethnic groups, multiple regression analyses consistently find that income is a much more powerful predictor of performance than race. Initial versions of the Accountability Index used race as an indicator, but it did not produce results that were as accurate as when income status was used.

Figure A1: Average Reading Scale Scores by Income Level and Race, All Grades (2007)

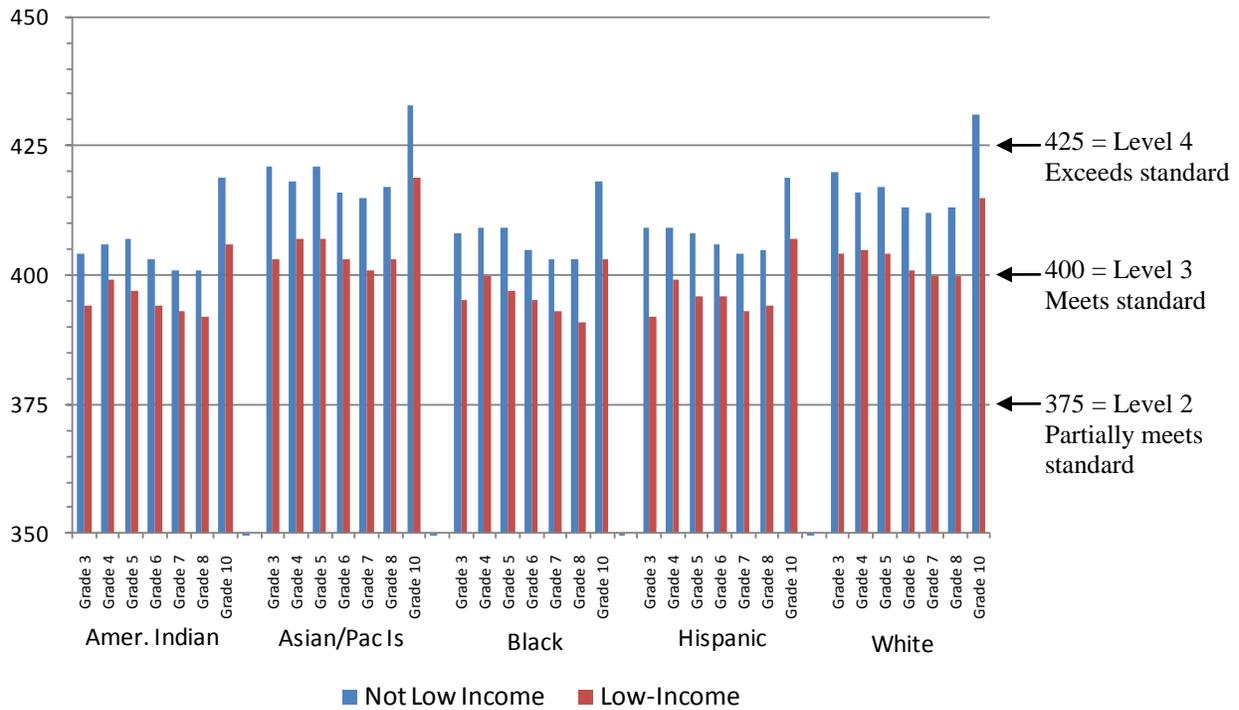
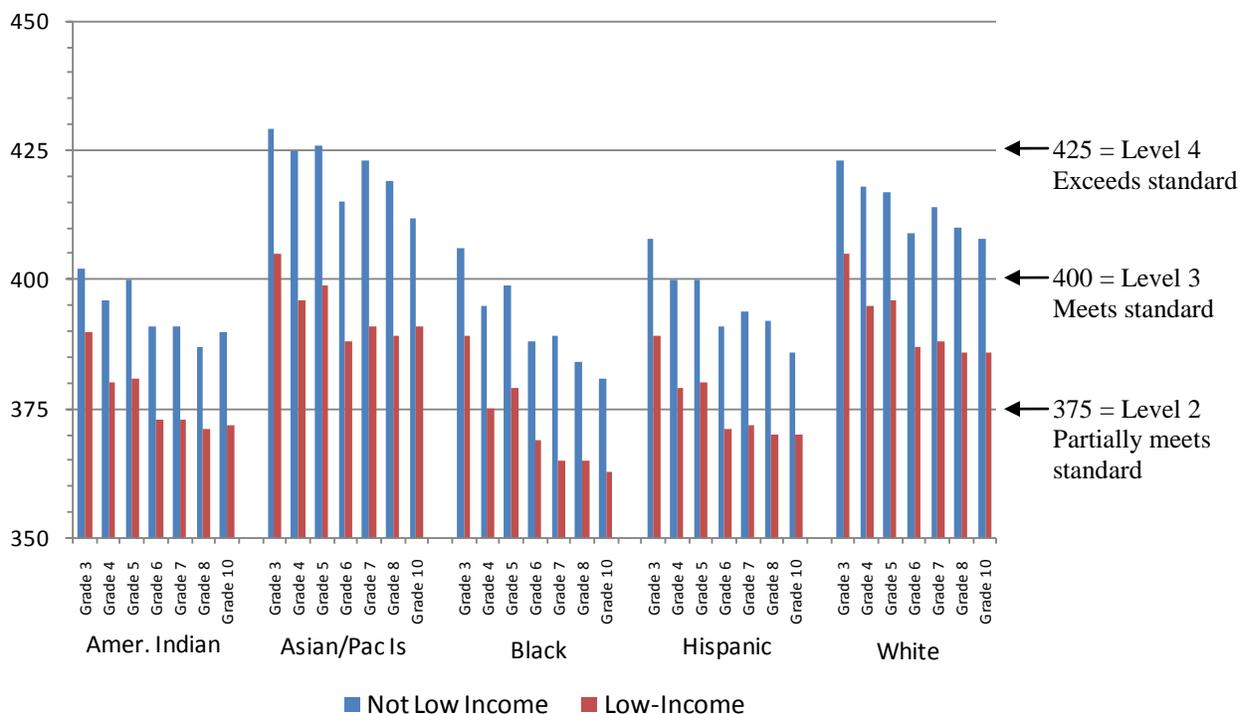


Figure A2: Average Math Scale Scores by Income Level and Race, All Grades (2007)



INDICATOR 3: ACHIEVEMENT VS. PEERS

This indicator uses the Learning Index (described below) and controls for student characteristics beyond a school's control. The score is the difference between a school's adjusted level and the average (predicted) level among schools/districts with similar characteristics (i.e., "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 similar student characteristics; those with scores below 0 are performing below those with similar student characteristics.

Separate analyses are run for the four different types of schools—elementary, middle, high, and comprehensive (e.g., K-12)—because of the variation of the variables at each grade level. Schools serving non-regular schools (e.g., alternative schools, ELL and special education centers, private schools on contract, institutions) are not included in the regressions (they are self-identified as non-regular schools in the OSPI database). 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. The Index for non-regular schools is based on an average of their remaining ratings. Schools without a federal meal program are not included in the regressions because there is no information about their percentage of low-income students.

For *schools*, five student characteristics are the independent variables in the multiple regression: the percentage of (1) low-income students (percent eligible for free or reduced-price lunch³³), (2) English language learners, (3) students with disabilities, (4) mobile students (not continuously enrolled), and (5) students designated as being gifted. The dependent variables are a school's Learning Index for each of the four assessments and the extended graduation rate. The regressions are weighted by the number of students assessed in the subject (and the number of students in grades 9-12 for the extended graduation rate) to prevent a small "outlier" school from distorting the regression (predicted) line. The regressions showed that all five variables helped improve the quality of the predicted levels. The regression uses a "stepwise" method with its five variables.

For *districts*, three student characteristics used in the school analysis are 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 is not used because analyses including this variable do not provide meaningful results. The percentage of students designated as gifted is not included because there is little variation at the district level. The same five dependent variables from the school-level analyses are used in the district analyses (the Learning Index for the four subjects and the extended graduation rate).

Financial information is 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 than others due to differences in the maximum levy amounts and the relative wealth of the community. The financial variable used is the total amount of operating expenditures per weighted pupil. This variable controls for the level of funds spent in the district and does not include spending for capital projects. The "weighted pupil" count "inflates" the enrollment figure

³³The percentage of students in high schools who are eligible is sometimes higher than what is reported, but this proxy for socioeconomic status is still the best available.

because certain students require more resources to educate and receive extra funding in the state formula. The extra weight for ELL and low-income students is .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 is .93, which is consistent with both the national research and the level of funding provided by the state. This weighting system effectively “subtracts” the extra amount of funding that districts receive from their total based on the level of students in their district who generate additional funding, which makes the financial amounts comparable.

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 not continuously enrolled in a school 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 current 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 who are new to the district and are still enrolled during the entire year. The current measure—the percentage of non-continuously enrolled students—can be used until a better measure is identified.

The benchmarks and ratings for this indicator in the four assessed subjects and the extended graduation rate are as follows:

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

> .207
.151 to .206
.051 to .155
-.05 to .054
-.051 to -.153
-.151 to -.202
< -.201

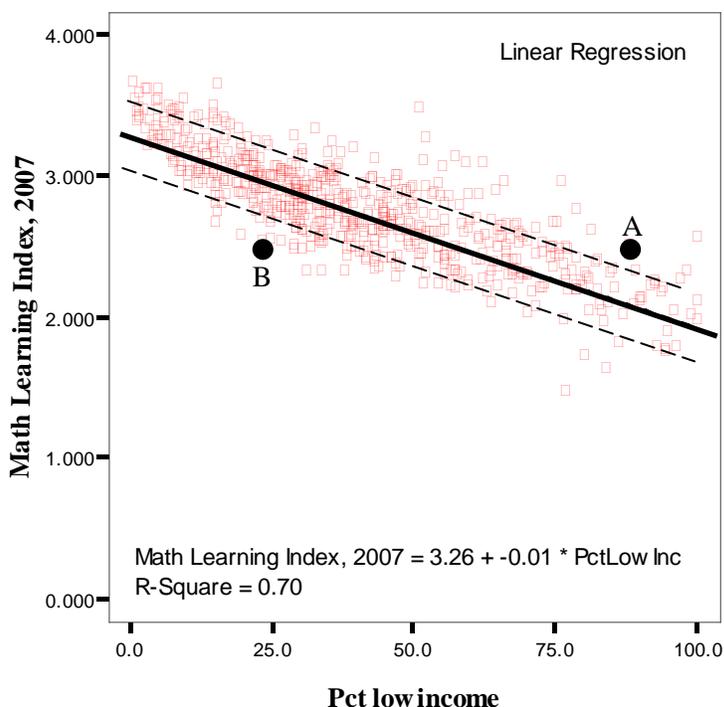
- Achievement vs. Peers on the *extended graduation rate* is rated based on the percentage point difference between the actual and predicted extended graduation rate:

> 127
6.1 to 126
3.1 to 65
-3 to 34
-3.1 to -63
-6.1 to -122
< -121

The scatterplot in Figure A3 illustrates the concept for how this “peer” indicator is calculated, although it shows only one of the independent variables (percent low income students) in relation to one outcome (elementary math results). Each dot represents an elementary school. The dark line is the average (predicted) level for a given Learning Index and low-income percentage. The vertical

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 school A falls well above the line while school 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 were the only one used in the analysis, school A would have a rating of 7 because its index is more than .20 points above its predicted level, while school B would have a rating of 1 because its index falls more than .20 points below the predicted level. (Note that excluding alternative schools from the analysis removes many schools that would appear well below the bottom dashed line. Including them would move the predicted line downward, resulting in more schools being above the predicted line.)

Figure A3: 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 minority students and school size (enrollment).

- A race/ethnicity variable was not included because it is highly correlated with the other variables. Statistical analyses (stepwise regression using student data at each grade level) using this variable found it had very little explanatory power in the model and the other variables had stronger coefficients. Using this variable would also reduce our ability to identify schools where students of color are treated differently. Finally, students of color are more likely to come from low-income families, and the performance of low income students is a separate indicator already in the matrix.
- While school size is used as a weight in the multiple regression, a school size variable was not included as an independent variable because research findings to date reveal mixed results about how school enrollment levels affect student outcomes. Moreover, statistical analyses

found that this variable added little to the explanatory power of the model. School size is also a factor that can be controlled somewhat at the district level through the use of specialized programs and boundary lines, so it is not an “external variable” like the others. Once the accountability results are made known, other methods can be used to help schools compare themselves to those with similar sizes.

The **Learning Index** is the dependent variable used in this indicator and for the Improvement indicator described below. This index was developed by the Commission on Student Learning and refined by the A+ Commission.³⁴ It takes into account the percentage of students performing at the five different performance levels based on their scale score:

- Level 0 – No score given³⁵
- 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 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, so their improvement rating will not be calculated (see Indicator 4 below).

The following example shows how the Learning Index is calculated. The same method is used to calculate the index for all the assessments in 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

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

The U.S. Education Department allows the use of this type of index when determining AYP. However, it does not allow the use of levels above proficiency (e.g., Level 4).

Another method to look at peer groups is through the use of a new **Challenge Index**. This index recognizes that some schools and districts need more resources to educate some types of students. Although this index should not be used when calculating the Accountability Index, data related to the relative challenge can still be reported elsewhere to give educators, the community, and other

³⁴These Commissions are no longer in existence.

³⁵The “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.

stakeholders a general sense of the overall challenge schools and districts face compared to others in the state.

Rather than create complex statistical models to determine the level of challenge faced, a simpler and more transparent approach is to approximate the relative challenge each school and district faces by adding the percentage of higher-need students together. Specifically, a proposed Challenge Index adds the percentage of assessed students who are (1) low-income, (2) ELLs, (3) in special education, and (4) mobile, and then subtracts the percentage of gifted students. For example, in 2008 the state had 37.9% low income students, 3.4% ELLs, 12.1% special education, 7.3% mobile, and 4.1% gifted assessed in the various grades. This equals a Challenge Index of .566 as computed below.

$$\text{State Challenge Index (2008)} = .379 + .034 + .121 + .073 - .041 = .566$$

The Challenge Index results can then be divided into 10 groups (deciles), with each group representing about 10% of the students assessed statewide, as shown in Table A4. Those with the least challenge (a Challenge Index less than .20) have the lowest decile of 1, while those with the greatest challenge (a Challenge Index more than 1.00) have the highest decile of 10. The higher the index, the greater the need for additional resources. This information can be helpful when determining which schools and districts may be in most need of additional support.

Table A4: Challenge Index Data¹

Decile	Decile Range	Pct. of assessed students statewide
1	< .200	9.8
2	.200 – .300	9.5
3	.301 – .390	10.6
4	.391 – .460	10.0
5	.461 – .520	9.2
6	.521 – .600	10.5
7	.601 – .690	9.7
8	.691 – .810	10.4
9	.811 – 1.00	10.1
10	> 1.00	10.1

¹State mean = .566; state median = .524

INDICATOR 4: IMPROVEMENT

The Improvement indicator relies on changes from one year to the next in the Learning Index for the four assessed subjects and for the extended graduation rate. The benchmarks and ratings for this indicator are as follows:

- Improvement on *assessments* is rated on the annual change in the Learning Index:
 - > .157
 - .101 to .156
 - .051 to .105
 - .05 to .054
 - .051 to -.10.....3
 - .101 to -.15.....2
 - < -.151

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

> 6	7
4.1 to 6	6
2.1 to 4	5
-2 to 2	4
-2.1 to -4	3
-4.1 to -6	2
< -6	1

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). Moreover, new schools 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 previously used for grade 4 reading), 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 it uses available data. The other measures can be used when analyzing the schools and districts that do not make AYP in consecutive years.

Schools and districts performing at very high levels are not able to improve much, so this indicator is excluded from the index calculations when a rating of 7 cannot be reached. Specifically, a school or district will have its improvement indicator suppressed when computing the accountability index when a Learning Index reaches 3.85 out of 4.00 and remains at or above that level for two consecutive years. (A school or district needs to improve by more than .15 to receive a rating of 7, which is impossible when their Learning Index is at or greater than 3.85.) The first year the index falls above 3.85, the school/district would get a rating based on their improvement. If the index stays at or above 3.85, the maximum rating is not possible, so the indicator would not be calculated. The same policy applies to the extended graduation rate outcome. A school or district will have the improvement results suppressed when the rate reaches or exceeds 94% and remains at that level for two consecutive years. (The graduation rate must improve by more than 6 percentage points to earn a rating of 7.)

EXTENDED GRADUATION RATE MEASURE

Washington defines the on-time graduation rate as the percentage of students who graduate from public high school with a regular diploma (not including a GED or any diploma not fully aligned with the state’s academic content standards) in the standard number of years. This is usually 4 years, although ELLs and students with disabilities are allowed a longer period of time to graduate “on time.” The 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 ELL 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.

Due to data limitations, the **on-time** graduation rate is currently estimated and calculated in the following manner using data from one year:³⁶

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

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

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 is 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.

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.

The extended graduation rate is used as the “other indicator” for AYP. At the elementary and middle school levels, the “other indicator” is the unexcused absence rate. This rate will continue to be calculated as required by NCLB and used in AYP determinations, but it is not used in the Accountability Index because it is measured differently among schools. No other outcome was identified for use in the Index for elementary and middle schools, so these schools have fewer cells in the matrix.

OTHER USEFUL MEASURES TO REPORT

Other outcome measures were considered to be part of the accountability matrix in order to have the Index reflect more than just assessment results. The advisors initially discussed three options because they are already collected: unexcused absence data (currently used as the “other indicator” for AYP), safety reports, and perception data (required for use in the school improvement process). However, none of these are collected in a consistent manner.

³⁶See <http://www.k12.wa.us/DataAdmin/pubdocs/GradDropout/03-04/Graduationanddropoutstatistics2003-04Final.pdf>, chapter 1, for more information about these formulas.

Other possible variables were discussed that relate to high schools and post-secondary outcomes. Among those considered were the percentage of students (1) taking the minimum number of courses that make them eligible to attend a 4-year public Washington college, (2) enrolled in Advanced Placement (AP) or International Baccalaureate (IB) courses, (3) participating in a dual enrollment program, such as Running Start, (4) taking a college entrance exam, such as the SAT, (5) enrolling in college the year after graduation, and (6) required to take remedial courses while in college. However, stakeholders suggested not including any of these variables in the matrix because of data problems associated with each. None is collected by the state, although legislation proposes having the state collect transcript data in a standardized manner. Adding more outcomes to the matrix would also make the system more complicated. As a result, the matrix is, by design, limited in scope.

Nevertheless, other types of data will be made available to provide a more comprehensive picture of educational outcomes. Contextual factors play a significant role in student performance, and data on these factors at the school, district, and state levels for multiple years will be provided to provide more transparency about the operating environment and acknowledge the role that factors outside the education system have on student performance. In some cases, these data are used to calculate the results in the accountability matrix (e.g., percentage of low-income, ELL, special education, gifted, and mobile students). The disaggregated results required by NCLB and the outcomes discussed above will be made available as well. Having additional student outcome data in one location provides a comprehensive look at all student groups at the same time. In most cases, the data are already available on separate pages of the OSPI Report Card, so combining them in one place is relatively easy. In a few cases, the data have not been collected but could be done easily. OSPI is planning to collect standardized transcript data at the secondary level, which allows for reporting the percentage of students that have taken the minimum number of course in the various content areas that make them eligible to attend a public 4-year state university.³⁷

Appendix H provides a list of suggested data to be provided. SBE and OSPI should work together and post data on the variables at the school, district, and state level for at least three years at a time. The variables relate to student and staff characteristics, district revenues and expenditures, community and state support, and student outcomes.

DISTRIBUTION OF INDEX

Some stakeholders have voiced concern that the index might be unfair and tend to place schools and districts with more challenging student populations in the Struggling tier. Given the high correlation between family income and student performance, analyses were conducted to see how the Accountability Index related to socioeconomic status.

Figure A4 shows these results for the 2,011 schools used in the analysis, while Figure A5 shows the results for the 291 districts.³⁸ These figures show a very weak relationship between the two

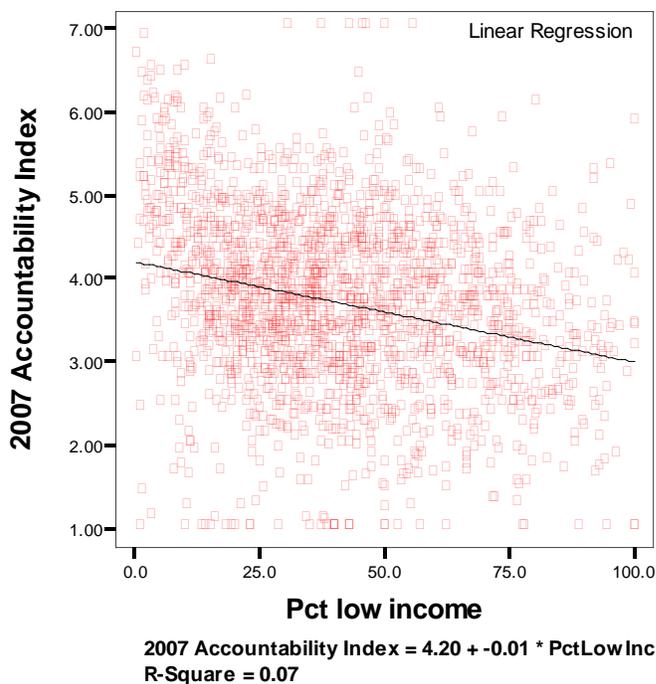
³⁷The Higher Education Coordinating Board sets these requirements, which are 4 English courses, 3 math courses, 2 lab sciences, 3 social studies courses, 2 foreign language courses, and 1 course in the arts. A BERC Group analysis of transcripts from a representative sample of graduates statewide in the Class of 2008 found that 49% had met the college-eligible course requirements.

³⁸Schools and districts that have fewer than 10 students assessed (this includes schools with no tested grades, such as K-2 or 11-12) are held accountable through a review of their improvement plans (the current policy).

variables than what would be seen if the dependent variable was achievement for all students. 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 high index scores. The trend line is still sloping downward, but the correlations and r-squares are rather weak (-.26 and .07 for schools, -.18 and .03 for districts). These are weak relationships because achievement for low-income students represents only 25% of the index and is moderated by the improvement and peers variables that have very low correlations with socioeconomic status. (The strongest correlation for schools was for writing improvement at -.091.) Figure A6 shows the distribution of district results, which are more tightly grouped than schools.³⁹

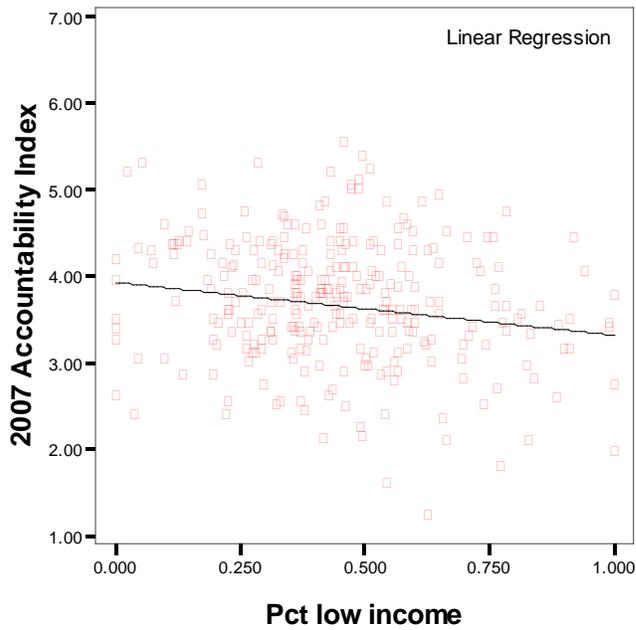
Other analyses were conducted to determine if the Achievement vs. Peers indicator was biased against higher poverty schools or if there was a non-linear relationship in the results. The analyses examined the residuals for low, medium, and high poverty schools by grade level for each subject. The mean residuals for all groups were very close to 0 (as expected) and the standard deviations were very similar in size among the different socioeconomic groups at the different grade levels. This shows that the results of the distribution of the residuals are not biased toward any group and are linear in nature.

Figure A4: Scatterplot of 2007 Index for All Rated Schools, by Percent Low Income (2007)



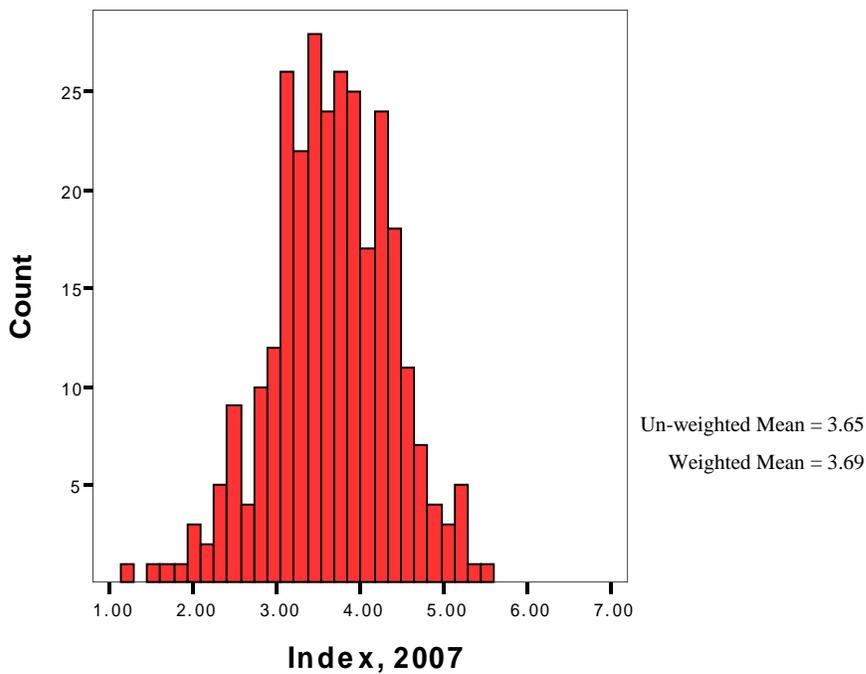
³⁹District results do not include correctional institutions, tribal schools, contract schools, and schools serving more than 50% of students outside the district boundary. The aggregation rules used in these calculations are the same as those used by OSPI when calculating district results. Results are not published when there are fewer than 10 students assessed in all the tested grades.

Figure A5: Scatterplot of Index for Districts, by Percent Low Income (2007)



$2007 \text{ Accountability Index} = 3.93 + -0.61 * \text{PctLowIncome07}$
R-Square = 0.03

Figure A6: Distribution of Districts by Index Score (2007)



TIERS

Tier “cut scores” (e.g., 2.50 to reach Fair) were created through an iterative process with multiple stakeholders using both empirical data and professional judgment. The distribution of all the Index results, such as those shown earlier (Figures 1 and A6), were reviewed for several years. Initial cut scores were created and advisors reviewed the Index results and tier assignments for schools in their district to determine if they reflected what they knew about the quality of the schools. The Indexes of schools known to be high performers were reviewed to see where they fell in the distribution. Revisions were made to the cut scores as a result of these reviews. When a change was made from a 5-point to a 7-point scale, cut scores were adjusted to create more tiers. As other policies were proposed (e.g., changing how ELL results are counted), the distribution changed and the tier cut scores were reviewed to see if they still reflected a valid description of the schools. Changes in the benchmarks also affect the ratings, which affect the Index. Hence, any change in one needed to be reviewed in relation to the other. In addition, stakeholders discussed how differences in assessment rigor affected a school’s scores. For example, middle school assessments are more difficult than those in the other grade bands, which makes it more difficult for these schools to achieve a higher tier (they also do not have graduation rates in their matrix, which tend to have higher ratings than the assessments.) Further, staff in more than 100 districts, which together enrolled more than half the students statewide, reviewed preliminary results for their schools to ensure the ratings and Index produced results that were consistent with their understanding of current conditions. Adjustments were made to the system based on feedback from district staff in order to increase the system’s validity. District results, which have a narrower range, were also reviewed to see how districts fell in the tiers. Stakeholders decided to keep the same set of cut scores for all types of schools and for districts to enhance simplicity and transparency, even though it meant having fewer middle schools and districts in the highest tiers. (Research has found overall district quality tends to be lower than the quality in individual schools.⁴⁰) Ultimately, the cut scores reflect policy decisions about what level of performance should generate the right tier label.

The names of the tiers were debated and changed over time. Options considered include colors (e.g., blue for the high tier, red for the lowest tier), metals (e.g., platinum, gold, silver, bronze), numbers, and other value-based descriptors (e.g., excellent, adequate, failing). The use of letters (e.g., A, B, C, etc.) was rejected because they are used for student grades and can cause confusion about school/district quality.

Table A5 shows the distribution of the different types of schools among the different tiers, based on their average index in 2006 and 2007. Alternative schools are over-represented in the Struggling tier. High schools are more likely to have a higher index because of five factors: (1) dropouts reduce the number of low performers, (2) students who struggle may go to an alternative school, (3) graduation rate outcomes help improve the index numbers, (4) the test is easier to pass compared to other grades, and (5) high school students have multiple chances to take the test, including in grade 9 and in August of grade 10.

⁴⁰See Characteristics of Improved School Districts: Themes from Research (2004), Shannon & Bylsma, OSPI, <http://www.k12.wa.us/research/pubdocs/DistrictImprovementReport.pdf>.

Table A5: Distribution of 2-Year Average Results by Tier and School Type

Tier	Elementary	Middle/ Jr. High	High	Multiple Levels	Alternative
Exemplary	29	1	1	2	9
Very Good	51	7	26	2	4
Good	280	100	137	6	33
Fair	657	221	147	15	93
Struggling	68	17	2	3	48
Total	1,085	346	313	28	187
% of total					
Exemplary	2.7%	0.3%	0.3%	7.1%	4.8%
Very Good	4.7%	2.0%	8.3%	7.1%	2.1%
Good	25.8%	28.9%	43.8%	21.4%	17.6%
Fair	60.6%	63.9%	47.0%	53.6%	49.7%
Struggling	6.3%	4.9%	0.6%	10.7%	25.7%

Table A6 shows a more detailed distribution of 2008 index results among the 1,994 schools that had at least four rated cells. This information is useful to see how a change in the benchmarks affects the tier distribution. For example, if the minimum benchmark for the Fair tier was raised from 2.50 to 3.00, another 209 schools would be in the Struggling tier.

Table A6: Distribution of Schools by Index, 2008

Index Range	Current Tier	# of schools	% of all schools
6.50-7.00	Exemplary	4	0.2%
6.00-6.49		22	1.1%
5.50-5.99		79	4.0%
5.00-5.49	Very Good	162	8.1%
4.50-4.99	Good	275	13.8%
4.00-4.49		380	19.1%
3.50-3.99	Fair	369	18.5%
3.00-3.49		303	15.2%
2.50-2.99		209	10.5%
2.00-2.49	Struggling	116	5.8%
1.50-1.99		43	2.2%
1.00-1.49		32	1.6%

Table A7 shows the number of schools in each tier according to their AYP results in 2008 and according to their two-year average index (2007 and 2008). In 2008, many more schools did not make AYP and were identified as being “in school improvement” because of a large increase in the goals. The table shows a positive relationship between AYP results and index results: the lower the tier, the more likely a school is in school improvement. However, a few high-performing schools (those in the Exemplary and Very Good tiers) did not make AYP, and some were “in improvement.” Far fewer schools are in the Struggling tier than in school improvement: over 600 schools were “in improvement” but the number with a 2-year average in the Struggling tier was 1/5 this amount. Table A8 shows results for Title I schools, which are the only schools subject to federal sanctions. Table A9 shows the number of schools in each tier that did not make AYP in 2008 in the “all students” cell in reading and math.

Table A7: Distribution of Schools by Tier and 2008 AYP Results

2008 Tier	Total Schools	Made AYP	Did not make AYP	# in School Improve.	Step					Pct. made AYP	Pct. did not make AYP	Pct. in School Improve.
					1	2	3	4	5			
Exemplary	116	104	12	1	1	0	0	0	0	89.7%	10.3%	0.9%
Very Good	164	108	56	27	17	2	7	1	0	65.9%	34.1%	16.5%
Good	659	289	370	182	121	13	20	5	23	43.9%	56.1%	27.6%
Fair	895	219	676	332	200	27	55	17	33	24.5%	75.5%	37.1%
Struggling	221	72	149	74	72	8	10	7	7	32.6%	67.4%	33.5%
Total	2,055	792	1,263	616	411	50	92	30	63	38.5%	61.5%	30.0%

2-Yr Ave Tier	Total Schools	Made AYP	Did not make AYP	# in School Improve.	Step					Pct. made AYP	Pct. did not make AYP	Pct. in School Improve.
					1	2	3	4	5			
Exemplary	47	45	2	0	0	0	0	0	0	95.7%	4.3%	0.0%
Very Good	106	82	24	6	5	0	1	0	0	77.4%	22.6%	5.7%
Good	681	329	352	142	97	9	21	5	10	48.3%	51.7%	20.9%
Fair	1,017	247	770	399	239	32	61	19	48	24.3%	75.7%	39.2%
Struggling	121	31	90	64	38	7	8	6	5	25.6%	74.4%	52.9%
Total	1,972	734	1,238	611	379	48	91	30	63	37.2%	62.8%	31.0%

Note: There are fewer schools with a 2-year average because some schools did not exist or have data in both years.

Table A8: Distribution of Title I Schools by Tier and 2008 AYP Results

2-Yr Ave Tier	Total Schools	Made AYP	Did not make AYP	# in School Improve.	Step					Pct. made AYP	Pct. did not make AYP	Pct. in School Improve.
					1	2	3	4	5			
Exemplary	9	9	0	0	0	0	0	0	0	100%	0%	0%
Very Good	36	24	12	2	2	0	0	0	0	66.7%	33.3%	5.6%
Good	299	164	135	50	33	2	6	4	5	45.2%	54.8%	16.7%
Fair	520	133	387	182	115	11	24	7	25	25.6%	74.4%	35.0%
Struggling	53	10	43	28	17	3	5	1	2	18.9%	81.1%	52.8%
Total	917	340	577	262	167	16	35	12	32	37.1%	62.9%	28.6%

Table A9: Distribution of Schools Not Making AYP in “All Students” Cell in 2008

2-Yr Ave Tier	Total schools in tier	Did not make AYP in 2008 in “All” Reading	Did not make AYP in 2008 in “All” Math
Exemplary	47	0	0
Very Good	106	0	1
Good	681	20	83
Fair	1,017	183	434
Struggling	121*	41	62
Total	1,972	244	580

* Of the 121 schools in the struggling tier, 51 had too few students in both the reading and math “all students” cells in 2008 to receive an AYP determination (i.e., they had fewer than 10 continuously enrolled students). Most of these were alternative schools. The majority of the remaining 70 schools that met the minimum N requirement and had a 2-year average below 2.50 (Struggling tier) did not meet the 2008 reading or math goal in the “all students” cells (59% did not make AYP in reading and 89% did not make AYP in math).

DESCRIPTIVE STATISTICS

Table A10 shows the average (mean) rating and standard deviation for each cell in the matrix among schools, along with the number of schools receiving a rating in the cell. Table A11 shows the correlations of the ratings with the 2008 index (all were positive and all but one were statistically significant).

Table A10: Descriptive Statistics for All Schools, 2008

Measure	N	Mean*	Std. Deviation
Non-Low Income Reading Achievement Rating	1,919	5.31	1.33
Non-Low Income Writing Achievement Rating	1,748	4.95	1.57
Non-Low Income Math Achievement Rating	1,923	3.81	1.55
Non-Low Income Science Achievement Rating	1,717	2.86	1.58
Non-Low Income Ext. Graduation Rate Rating (Class of '07)	516	3.96	2.41
Low Income Reading Achievement Rating	1,838	3.81	1.30
Low Income Writing Achievement Rating	1,618	3.53	1.66
Low Income Math Achievement Rating	1,838	2.04	1.17
Low Income Science Achievement Rating	1,602	1.29	.71
Low Income Ext. Graduation Rate Rating (Class of '07)	459	2.96	2.15
Reading Peer Rating	1,802	4.03	1.73
Writing Peer Rating	1,746	3.97	2.05
Math Peer Rating	1,803	4.00	2.00
Science Peer Rating	1,715	4.00	2.19
Ext. Graduation Rate Peer rating (Class of '07)	336	4.34	1.81
Reading Improvement Rating	1,933	4.55	1.84
Writing Improvement Rating	1,853	4.74	2.22
Math Improvement Rating	1,935	4.01	1.88
Science Improvement Rating	1,825	4.90	2.17
Ext. Graduation Rate Improvement Rating (Class of '06 to '07)	459	4.11	2.42
Accountability Index	2,058	3.83	1.09

*Unweighted

Table A11: Correlations of 2008 Index with All Ratings for All Schools and Regular Schools by Grade Band*

Variable Rated	All schools	Regular high schools	Regular middle schools	Regular elem. schools	Regular multi-level schools
Non-Low Income Reading Achievement <i>N</i> 1,919	.692	.615 267	.714 345	.686 1,021	.685 118
Non-Low Income Writing Achievement <i>N</i> 1,748	.615	.490 267	.655 329	.699 913	.522 97
Non-Low Income Math Achievement <i>N</i> 1,923	.686	.632 268	.671 345	.718 1,021	.741 118
Non-Low Income Science Achievement <i>N</i> 1,717	.659	.627 268	.657 330	.703 885	.644 94
Non-low income ext. grad. rate (Class of '07) <i>N</i> 516	.625	.474 268			.529 62
Low Income Reading Achievement <i>N</i> 1,838	.622	.495 262	.676 347	.631 1,009	.672 113
Low Income Writing Achievement <i>N</i> 1,618	.484	.465 262	.642 328	.597 856	.361 96
Low Income Math Achievement <i>N</i> 1,838	.585	.409 262	.492 347	.680 1,009	.540 114
Low Income Science Achievement <i>N</i> 1,602	.391	.284 263	.339 328	.462 831	.408 98
Low income ext. grad rate (Class of '07) <i>N</i> 459	.486	.402 268			.460 56
Reading Peer <i>N</i> 1,802	.727	.692 277	.729 348	.747 1,043	.707 127
Writing Peer <i>N</i> 1,746	.643	.626 276	.688 338	.660 1,009	.475 116
Math Peer <i>N</i> 1,803	.716	.624 278	.720 348	.760 1,043	.632 127
Science Peer <i>N</i> 1,715	.630	.631 278	.612 339	.653 974	.537 117
Ext. grad rate compared to peers (Class of '07) <i>N</i> 336	.427	.453 271			.367 ^x 65
Reading Improvement <i>N</i> 1,933	.436	.416 264	.409 345	.414 1,026	.417 118
Writing Improvement <i>N</i> 1,853	.332	.417 264	.318 334	.311 989	.212 ^y 110
Math Improvement <i>N</i> 1,935	.416	.497 264	.450 345	.363 1,026	.466 118
Science Improvement <i>N</i> 1,825	.444	.448 264	.341 335	.464 959	.407 112
Change in ext. grad rate (Class of '06 to '07) <i>N</i> 459	.325	.329 260			.202 ^z 59

* Pearson correlation (2-tailed); all are statistically significant at the .000 level unless otherwise noted.

^x Significant at .01

^y Significant at .03

^z Not statistically different from 0 (no correlation with the index)

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 used to generate the index is also used to identify schools and districts for recognition. The distribution of school results in each cell of the matrix was reviewed to determine possible cut points. (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 A6, so they would receive recognition far less often than schools.)

SBE approved providing recognition for “Outstanding Overall Performance.” The rationale for this recognition system is as follows:

- The required 2-year averages are challenging but reachable targets. If a goal is too high, few will think they can reach it and the reward of recognition loses its motivational power.
- The same criteria are used for each subject for schools and districts for simplicity.
- The recognition system is based on a “theory of change” that people are motivated more by success than by blame or guilt and need clear, challenging, and attainable goals.
- The goals are criteria-based so schools/districts know what needs to be done to be recognized, and they do not have to worry about the performance of others. This goals are clear and encourages collaboration and cooperation among educators.
- Requiring minimum ratings and a 2-year average ensures recognition is given only for sustained exemplary performance and not based on one good year.
- A lower average is justified for the Index because it is harder to have a high average in multiple categories. The 5.50 average is the beginning of the Exemplary tier, so it would include all schools/districts with an average in the Exemplary tier.
- Recognizing relatively few schools in high priority areas demonstrates a commitment to these areas and provides more incentive to improve where the greatest improvement needs to occur.
- A more limited system ensures that any recognition that occurs is truly special. Having too many schools getting many awards reduces the significance of the recognition.
- The strongest predictor of the achievement gap is the difference between the two socioeconomic groups (non-low income and low income). The gap is measured in terms of the cells in the matrix rather than other gaps outside the matrix (e.g., the differences between race/ethnic groups).
- Outstanding sustained performance in schools with a “regular” student composition deserves recognition. Restricting the percentage of gifted students that are assessed provides a more accurate picture of school performance. High concentrations of gifted students generally inflate the results, making it easier for schools with special programs to receive recognition.

One other type of recognition—“Noteworthy Performance” based on lower performance levels in every cell of the matrix—was discussed but not implemented. Specifically, stakeholders discussed implementing a system in which the state could also provide recognition to schools and districts for each of the 20 cells of the matrix when the 2-year average for a cell is at least 5.50, and for the

index when the 2-year average is at least 5.00. To receive this type of recognition, schools and districts would also have needed to meet the following conditions:

- (a) No rating below 5 can occur in either year in the 20 cells of the accountability matrix.
- (b) Recognition for non-low income cells in reading and writing requires a minimum 2-year average of the low income group of 4.00.⁴¹

Table A12 shows the areas where recognition would be given and the minimum average.

Table A12: Required 2-Year Average for Noteworthy Performance

INDICATORS	OUTCOMES					Average
	Reading	Writing	Math	Science	Ext. Grad. Rate	
Achievement of non-low income	**	**	5.50			
Achievement of low income						
Achievement vs. peers						
Improvement						
Average						5.00

**Recognition in these cells requires the low-income cell to have a 2-year average of at least 4.00.

The rationale for this form of recognition was as follows:

- Giving recognition for all five outcomes and four indicators conveys the belief that all parts of the system are important. Recognizing fewer cells of the matrix could generate extra focus in some areas and not others.
- Requiring the low income reading and writing cells to have at least a 4.00 average ensures that cells that have high levels of performance do not get recognized if there is a significant achievement gap.
- There is no restriction on schools receiving recognition if they have 10% or more of their students designated as gifted. This allows all schools to be eligible for this type of recognition.
- Research has found that “small victories” support continuous improvement efforts. Education stakeholders viewed even minor forms of state recognition as a way to support improvement.

This option would have provided recognition to far more schools because it is based on performance in each of the 20 cells of the matrix as well as the index. More than 80% of the schools statewide (1,618 in total) met the criteria in some way, and some schools would have received recognition for performance in many of the cells of the matrix. Ultimately, neither SBE or OSPI decided to use this approach.

OSPI has given two types of awards in the past, both for improvement.

1. *Schools of Distinction* were recognized in school years 2006-07 and 2007-08 based on average improvement in the Learning Index in reading and math over an extended period of time (e.g., comparing 2008 to the average of 2002 and 2003) and required achievement to exceed the state average. Only the top 5% of schools receive this award based on their improvement. This is a “norm-referenced” system, so schools with high levels of improvement may not receive the

⁴¹This requirement is not used for math and science because so few schools/districts are meeting the 5.50 average, and there needs to be incentives to encourage overall performance in these two subjects.

award if they do not meet the state average or others improve by a greater amount. In 2008, a total of 101 schools (53 elementary, 21 middle, 20 high, and 7 alternative) received this award (two schools received recognition for performance at two grade levels). The average index for these schools in 2008 as 4.68, which is in the Good tier. Of these schools, 41% did not make AYP and 15 were in School Improvement. One alternative school receiving this recognition in 2008 had an index in the Struggling tier. Many of the schools receiving this recognition had a relatively high percentage of gifted students (as a group, they averaged nearly twice the state average), and their percentage of low income students was less than the state average.

2. *Academic Improvement Awards* have been given since 2004 to both schools and districts that make at least a 10% reduction in the percentage of students not meeting standard from the previous year in reading, writing, and math in grades 4, 7, and 10. (This is the level required for a school to make “safe harbor” under AYP.) Wall plaques with metal plates for updates are provided. In 2007, there were 1,255 schools (60% of schools statewide) that received a total of 2,190 awards in the three grades and subjects (a similar number of schools received awards in 2008); 241 districts (81% statewide) received a total of 804 awards in the three grades and subjects. All these awards are given regardless of AYP status.

OSPI has decided to give recognition using data from the Accountability Index beginning in Spring 2010. The details of OSPI’s recognition system are still being finalized.

The U.S. Education Department provides limited recognition. Federal awards are only given to schools and are competitive in nature. Four types of awards are given and only to schools that make AYP. In 2008, 59 schools receive these awards (3% of all schools statewide). Changes in the federal administration could result in changes in the criteria for federal awards.

1. *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. In 2008, four schools were recognized (seven schools had been nominated).
2. 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, and four received the award in 2008. 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.
3. The *Academic Improvement Award* is given to Title I Part A schools that have made AYP the past three years and shown significant gains overall, preferably among subgroups of students. Of the 48 schools receiving recognition in 2008, most were elementary schools.
4. 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. In 2008, three schools received this award.

No recognition is given at the federal or state levels based on how schools or districts compare to others with similar student characteristics or for achievement by any student group.

APPENDIX B

IN-DEPTH ANALYSIS

Schools and districts that do not make AYP two years in a row do not automatically fall into “improvement” status. Instead, they will undergo an in-depth review by OSPI staff in August of each year. This deeper analysis will use available data for the schools and districts that could be placed in an improvement step and when determining those that need more intensive help. This analysis may involve contacting the district and/or ESD to get more information.

Besides data from the index matrix itself, a comprehensive set of quantitative and qualitative data can be reviewed. These data fall into four general areas. This appendix provides the list of data that can be reviewed, although not all data will need to be reviewed in every case.

1. Contextual Data

- Type of school (alternative school, institution)
- Changes in student demographic profile (e.g., rapid increase in low-income or ELL students)
- Changes in the community (e.g., employment, immigration)
- 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

2. Analysis of Assessment Results (annual and trends over time)

- Achievement trends over multiple years in each subject area
- Size of the gap between test 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
- Number of students who receive a “zero”
- 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 (CAA)
- WAAS results (including Portfolios)
- WLPT results for students from different language backgrounds, percentage of students exiting ELL program

- District assessments: Results from any other assessments (e.g., MAP, grade 2 reading, portfolios)

3. Teaching and Learning Issues

- Classroom conditions: Class sizes, student/teacher ratios by grade and subject
- Staff characteristics: Percentage of teachers and other staff with certificates, teacher education/experience levels
- Staff turnover: Teacher and leadership changes at school and district levels
- Resource allocation: Where staff and other resources are deployed in the district
- Alignment of curriculum and materials across grades and with state standards
- Volunteers: Number of parents volunteers, how they are used
- Initiatives: Number being attempted, focus and validity of initiatives, level of integration and cohesion among activities
- ELL program model(s) being used
- Professional development: Types and focus, involvement of professional learning communities
- Data use: Quality of data system, capacity to use data, types of data made available to educators, how information is used by different staff
- Supplemental education services: Experience of providers, amount of support provided, number and grades of students receiving help
- Extended learning opportunities: After-school programs, summer school classes offered, number of students receiving support
- Community involvement: Type and level of partnerships with local businesses, community groups, and philanthropies

4. Other 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
- 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 not making AYP, percentage of district students enrolled in schools not making AYP or in school improvement
- Data anomalies: Incorrect data reported that could affect analyses, missing data, reason for missing data, number of ratings generating the average index
- 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, role of the district in school improvement efforts
- 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 OSPT's Comprehensive Program Review (CPR), input from ESDs

Once schools and districts are placed in an improvement step, the state will need to consider how best to provide additional support. *Readiness to benefit* from the extra support is important in order for the recipients to use additional resources effectively. Without buy-in and capacity, the chances for successful reform are minimal. *Size* and *location* may need to be considered as well. If the number of schools and districts in a step is high and exceeds the level of resources available to support them, the state may need to require a minimum number of students per school before providing assistance to ensure cost-effectiveness of the assistance. Similarly, those in an improvement step 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 resources to ensure the cost effectiveness of its support. The state may also consider providing support by geographic location to ensure equity in the distribution of the assistance.

APPENDIX C

CURRENT STATE ASSISTANCE PROGRAM

The Office of Superintendent of Public Instruction has created several assistance programs to help schools and districts that have not made Adequate Yearly Progress. This appendix summarizes the two programs, one for schools and the other for districts, that were in place in 2009. These programs are continually being revised to reflect the latest research on improving the education system and to conform to federal funding requirements.

SCHOOL IMPROVEMENT ASSISTANCE

The mission of OSPI'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 programs 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

OSPI has recently placed more emphasis on a systems approach in its work with districts. A special "Summit District Improvement" initiative was launched in 2008 for five districts that applied on a competitive basis. It focuses on (1) Data collection of student achievement, perception surveys, classroom observation study, transcript analysis, and college eligibility and attendance

indicators; (2) a district needs analysis and action plan; and (3) school level data reports and school level alignment and implementation of the action plans.

Districts not participating in this initiative 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

RELEVANT LEGISLATIVE MANDATES

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.

Note: The Powers and Duties section was amended and sections 4c-f were eliminated and replaced with new language in ESHB 2261 as noted below.

Engrossed Substitute House Bill 2261 (Signed into law in May 2009)
PART V: SHARED ACCOUNTABILITY FOR SCHOOL AND DISTRICT IMPROVEMENT

NEW SECTION. Sec. 501.

(1)(a) The legislature intends to develop a system in which the state and school districts share accountability for achieving state educational standards and supporting continuous school improvement. The legislature recognizes that comprehensive education finance reform and the increased investment of public resources necessary to implement that reform must be accompanied by a new mechanism for clearly defining the relationships and expectations for the state, school districts, and schools. It is the legislature's intent that this be accomplished through the development of a proactive, collaborative accountability system that focuses on a school improvement system that engages and serves the local school board, parents, students, staff in the schools and districts, and the community. The improvement system shall be based on progressive levels of support, with a goal of continuous improvement in student achievement and alignment with the federal system of accountability.

(b) The legislature further recognizes that it is the state's responsibility to provide schools and districts with the tools and resources necessary to improve student achievement. These tools include the necessary accounting and data reporting systems, assessment systems to monitor student achievement, and a system of general support, targeted assistance, recognition, and, if necessary, state intervention.

(2) The legislature has already charged the state board of education to develop criteria to identify schools and districts that are successful, in need of assistance, and those where students persistently fail, as well as to identify a range of intervention strategies and a performance incentive system. The legislature finds that the state board of education should build on the work that the board has already begun in these areas. As development of these formulas, processes, and systems progresses, the legislature should monitor the progress.

Sec. 502. RCW 28A.305.130 and 2008 c 27 s 1 are each amended to read as follows:

The purpose of the state board of education is to provide advocacy and strategic oversight of public education; implement a standards-based accountability ~~system~~ framework that creates a unified system of increasing levels of support for schools in order 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: **Sections (c) through (f) are repealed**

NEW SECTION. Sec. 503.

A new section is added to chapter 28A.305 RCW to read as follows:

(1) The state board of education shall continue to refine the development of an accountability framework that creates a unified system of support for challenged schools, that aligns with basic education, increases the level of support based upon the magnitude of need, and uses data for decisions.

(2) The state board of education shall develop an accountability index to identify schools and districts for recognition and for additional state support. The index shall be based on criteria that are fair, consistent, and transparent. Performance shall be measured using multiple outcomes and indicators including, but not limited to, graduation rates and results from statewide assessments. The index shall be developed in such a way as to be easily understood by both

employees within the schools and districts, as well as parents and community members. It is the legislature's intent that the index provide feedback to schools and districts to self-assess their progress, and enable the identification of schools with exemplary student performance and those that need assistance to overcome challenges in order to achieve exemplary student performance. Once the accountability index has identified schools that need additional help, a more thorough analysis will be done to analyze specific conditions in the district including but not limited to the level of state resources a school or school district receives in support of the basic education system, achievement gaps for different groups of students, and community support.

(3) Based on the accountability index and in consultation with the superintendent of public instruction, the state board of education shall develop a proposal and timeline for implementation of a comprehensive system of voluntary support and assistance for schools and districts. The timeline must take into account and accommodate capacity limitations of the K-12 educational system. Changes that have a fiscal impact on school districts, as identified by a fiscal analysis prepared by the office of the superintendent of public instruction, shall take effect only if formally authorized by the legislature through the omnibus appropriations act or other enacted legislation.

(4)(a) The state board of education shall develop a proposal and implementation timeline for a more formalized comprehensive system improvement targeted to challenged schools and districts that have not demonstrated sufficient improvement through the voluntary system. The timeline must take into account and accommodate capacity limitations of the K-12 educational system. The proposal and timeline shall be submitted to the education committees of the legislature by December 1, 2009, and shall include recommended legislation and recommended resources to implement the system according to the timeline developed.

(b) The proposal shall outline a process for addressing performance challenges that will include the following features: (i) An academic performance audit using peer review teams of educators that considers school and community factors in addition to other factors in developing recommended specific corrective actions that should be undertaken to improve student learning; (ii) a requirement for the local school board plan to develop and be responsible for implementation of corrective action plan taking into account the audit findings, which plan must be approved by the state board of education at which time the plan becomes binding upon the school district to implement; and (iii) monitoring of local district progress by the office of the superintendent of public instruction. The proposal shall take effect only if formally authorized by the legislature through the omnibus appropriations act or other enacted legislation.

(5) In coordination with the superintendent of public instruction, the state board of education shall seek approval from the United States department of education for use of the accountability index and the state system of support, assistance, and intervention, to replace the federal accountability system under P.L. 107-110, the no child left behind act of 2001.

(6) The state board of education shall work with the education data center established within the office of financial management and the technical working group established in section 112 of this act to determine the feasibility of using the prototypical funding allocation model as not only a tool for allocating resources to schools and districts but also as a tool for schools and districts to report to the state legislature and the state board of education on how the state resources received are being used.

APPENDIX E

STATE BOARD OF EDUCATION ACCOUNTABILITY FRAMEWORK AND RESOLUTION

DRAFT ACCOUNTABILITY FRAMEWORK

Guiding Principles

The Board is creating an accountability framework that:

- Affirms that the call for stronger accountability must be accompanied by comprehensive funding reform
- Recognizes the sense of urgency to address the needs of all our students
- Works toward a unified system of federal and state accountability
- Recognizes the critical role of local school districts in addressing continuous improvement in student achievement
- Recognizes schools and districts that have demonstrated significant learning and/or improvement by their students by identifying and rewarding best practices and exemplary work
- Advocates for the state to provide proactive support for districts to make improvement in student achievement
- Creates a new collaborative mechanism to require certain school district actions if student achievement does not improve
- Recognizes the need for support from the local community, parents, staff in the schools and districts, regional and business partners, and state officials to improve our education system for all students
- Demonstrates “your money at work” in a new basic education funding system with its focus on student achievement.

Key Components of the Proposed System

1. An **Accountability Index** that (a) uses criteria that are fair, consistent, transparent and easily understood, (b) provides feedback to schools and districts to self-assess their progress, and (c) identifies schools with exemplary performance as well as those that are experiencing problems.

2. **Proactive, Targeted and Intensive Voluntary Programs** that build the capacity of districts to help their schools improve student achievement. Programs offered will be tailored to the magnitude of need. As part of this system of assistance the Board will ensure that all efforts are administered as part of one unified system of state assistance including the Innovation Zone – a new effort to help districts dramatically improve achievement levels by implementing exemplary leadership and instructional practices.

3. A **Timeline for Improvement** that defines what student achievement improvement would be expected by a district.

4. **Required Action:** After the allotted time period, if there is inadequate improvement, the state will require action be taken by the relevant district(s). The collaborative process the Board envisions includes various partners with specific responsibilities.

- a) SBE will task OSPI to conduct an academic performance audit using a peer review team.

- b) The local district, in collaboration with OSPI, will develop an improvement plan for implementation based on that review.
- c) SBE reviews and approves the local district plan which, once approved, becomes a binding performance contract between the State and District.
- d) Local districts will remain responsible for implementation.
- e) OSPI will monitor the implementation of the plan and provide SBE updates.
- f) SBE and the local school will report to the community on progress in improving student achievement.

SBE will continue to seek input from all interested parties. In January 2009, the Board plans to adopt a resolution to the legislature that states the guiding principles and key components for a new statewide accountability system that it believes needs to be a part of the revisions made to the basic education funding system. The Legislature will need to provide the Board and OSPI with the appropriate authority and resources to implement the new system. The Board will continue to refine the details of the accountability system by working with its education, parent, business and community partners.

FINAL SBE ACCOUNTABILITY RESOLUTION (APPROVED JANUARY 15, 2009)

WHEREAS, the State Board of Education believes that all students deserve an excellent and equitable education and that there is an urgent need to strengthen a system of continuous improvement in student achievement for all schools and districts; and

WHEREAS, the Legislature charged the State Board of Education to develop criteria to identify schools and districts that are successful, in need of assistance, and those where students persistently fail, as well as to identify a range of intervention strategies and performance incentive systems; and

WHEREAS, the State Board of Education affirms the call for stronger accountability must be reciprocal between the state and local school district and accompanied by comprehensive funding reform for basic education that demonstrates “taxpayer money at work” in improving student achievement; and

WHEREAS, the State Board of Education will work with its education partners to create a unified system of federal and state accountability to improve student achievement; and

WHEREAS, the State Board of Education recognizes the need for a proactive, collaborative accountability system with support from the local school board, parents, students, staff in the schools and districts, regional educational service districts, business partners, and state officials to improve student achievement; and

WHEREAS, the State Board of Education believes that schools and districts should be recognized for best practices and exemplary work in improving student achievement; and

WHEREAS, the State Board of Education recognizes the critical role of local school boards in addressing student achievement in developing a new state accountability system as well as the need to create a new collaborative mechanism to require certain school district actions if student achievement does not improve;

THEREFORE, BE IT RESOLVED that the State Board of Education will develop an accountability index to identify schools and districts based on student achievement using criteria that are fair, consistent, transparent, and easily understood for the purposes of providing feedback to schools and districts to self-assess their progress as well as to identify schools with exemplary performance and those with poor performance; and

BE IT FURTHER RESOLVED that the State Board of Education will work with its education partners to build the capacity of districts to help their schools improve student achievement. Programs will be tailored to the magnitude of need. As part of this system of assistance, the Board will ensure that all efforts are administered as part of one unified system of state assistance including the Innovation Zone – a new effort to help districts dramatically improve achievement levels; and

BE IT FURTHER RESOLVED that after a time set by the State Board of Education where there is no significant improvement based on an Accountability Index and other measures as defined by the Board, the district will be placed on Academic Watch and the State Board of Education will:

- Direct the Office of Superintendent of Public Instruction to conduct an academic performance audit using a peer review team
- Request the local school board, in collaboration with the Office of Superintendent of Public Instruction, to develop an Academic Watch Plan based on the review findings, which would include an annual progress report to the local community
- Review, approve, or send back for modification to the local board, the Academic Watch plan, which once approved becomes a binding performance contract between the state and district
- Ensure that the local school board will remain responsible for implementation
- Request the Office of Superintendent of Public Instruction to monitor implementation of the plan and provide updates to the State Board of Education, which may require additional actions be taken until performance improvement is realized
- Declare that a district is no longer on Academic Watch when the Office of Superintendent of Public Instruction reports to the Board that the district's school or schools are no longer in Priority status; and

BE IT FURTHER RESOLVED that the Board believes this accountability framework needs to be a part of the revisions made to the basic education funding system and that the Legislature will need to provide the State Board of Education, the Office of Superintendent of Public Instruction, and the local school boards, with the appropriate legal authority and resources to implement the new system; and

BE IT FURTHER RESOLVED that the State Board of Education will continue to refine the details of the accountability system by working with its education, parent, business and community partners over the next year.

APPENDIX F

CONCEPTUAL FRAMEWORK FOR SUPPORT FOR STRUGGLING SCHOOLS

Mass Insight, Executive Summary, December 2008

The State Board hired Mass Insight early in the process to prepare ideas for helping struggling schools. This appendix provides a summary of these ideas.

OVERVIEW OF THE INITIATIVE

The Need

- Like all states, Washington has a small number of schools where students persistently achieve at significantly lower levels than at peer schools.
- Also like all states, Washington has not been able to eliminate – or even to narrow, appreciably – the large achievement gap between “have” and “have-not” students and schools.
- Finally – like all states – Washington’s public schools are not yet broadly and successfully preparing most high school graduates with college and work-ready skills, after 15 years or more of standards-based reform.

The Context

- The Washington Legislature has charged the State Board of Education with developing a state system to identify Washington’s most successful and least successful public schools, and to recommend an approach to improve the latter.
- The Basic Education Funding Task Force is reviewing the state’s investments in public schools and the ways those funds are being spent, with an eye towards recommending a new funding formula capable of meeting 21st Century expectations for proficiency.
- 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.

Core Strategies

- **Prioritize success.** Establish bold exemplars of systematic, comprehensive turnaround by focusing resources and capacity, rather than attempt to serve every needy school at once and, in doing so, produce inadequate results.
- Generate change by **enabling local leaders and their partners**, rather than through state mandates and alternate governance.
- Enable local leaders to **earn** the opportunity to participate by insisting on **transformation** with this initiative, not incremental change.
- **Hold everyone accountable**, from the state through the districts to the schools and the students.

SPECIFIC RECOMMENDATIONS

The Proposed Plan

- Districts with Priority Schools as determined by the State Board of Education’s new Accountability Index will have the option to apply to the Innovation Zone as one option for voluntary intensive assistance in an overall continuum of support and intervention. That continuum will be designed to provide graduated levels of assistance to schools and districts depending on their performance record (as measured by the state’s new Accountability Index) and demonstrated need.
- Districts will be admitted to the Zone after being vetted by the State Board for readiness (i.e., strong signals of commitment to transformative change) and for a solid turnaround plan. Districts will be encouraged to apply on behalf of small clusters 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 not limited to a focus on individual schools.
- The Zone will offer \$50,000 in planning and preliminary implementation grants to districts and a significant dollar amount per school in implementation grants for periods of up to five years, with benchmark expectations at two years (leaving Priority status) and at four years (moving into the state’s “adequate” tier of school performance). Districts will be strongly encouraged to work with a lead partner in designing and implementing their Zone initiative.
- Districts with Zone initiatives will maintain good standing and continue to receive support so long as (a) their Priority Schools meet the benchmark expectations or (b) they can develop a revised plan that addresses analysis of the reasons for continued under-performance.
- Other Options for Intervention: Districts with Priority Schools that do not join the Zone’s first cohort, either because they elect not to apply or because their proposal was deemed inadequate by the State Board, will participate in one of two other options open to them: OSPI’s comprehensive district reform initiative (also called Summit Districts), or a school turnaround initiative designed and implemented with minimal state assistance (which we call the Consulting Assistance model in this proposal).
- Academic Watch: Across all three of these options, districts whose Priority Schools are not able to leave that status after two full implementation years will be placed on Academic Watch for further review and action. OSPI and newly formed Peer Review Teams will consider a range of options tailored to local conditions to help those districts raise student achievement in their Priority Schools. Academic Watch can be regarded as the academic corollary, in some ways, to the state’s current Financial Watch approach with districts that need help reorganizing and managing their finances. It is the state’s “backup plan” for schools and districts that do not improve even after other strategies and resources have been applied. While it provides for a stronger state role in analyzing the lack of progress and collaborating with districts to define new turnaround strategies, it preserves the principle of local control that lies at the heart of Washington State’s system of public education.

Basic Definitions

- **The Innovation Zone is:**
 - **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.
- **The Innovation Zone is not:**
 - Simply an effort to fix some broken schools.
 - An initiative to distribute the available resources evenly across every challenged public school.
 - A top-down, mandated state program.

The Rewards of Taking Action

- The most important goal, of course, is student achievement. In addressing the needs of its most highly challenged schools, Washington State will also be targeting its resources in the communities with the highest concentrations of poverty. Erasing the poverty and racial achievement gaps has been called the most important civil rights issue of our time.
- There is a strategic benefit in acting now. The reauthorization of No Child Left Behind will likely produce extensive federal investment in school intervention strategies. Some of these funds likely will be competitive. States with robust, transformative strategies in place – such as the Washington State Innovation Zone – will be among the readiest recipients of those competitive federal funds.

APPENDIX G

CURRENT AYP SANCTIONS

No Child Left Behind requires increased accountability for all public elementary and secondary schools, especially those that receive Title I funds. Schools and districts that do not make adequate yearly progress (AYP) face a series of specific consequences as defined in the reauthorization of the Elementary and Secondary Education Act. States are required to report the names of *all* public schools that “need improvement” based on the state’s AYP policies, regardless of whether or not they receive Title I funds. However, only districts and schools that receive Title I funds face a series of consequences until they meet AYP criteria for two consecutive years.

The following describes the actions that are required of districts and schools when they “need improvement” and receive Title I funds. Information about these sanctions and why they occurred must be made available to parents and community members in an understandable, accessible format by the beginning of the school year. This may be accomplished in part by referring them to OSPI’s Web site under the Report Card section. Teachers, principals, parents, and community members also need to be informed about AYP results, even if no sanctions occur.

DISTRICT IMPROVEMENT

Step One

Districts that do not make AYP are in a “warning” year (also known as “Year 1”). Districts that do not make AYP in the same content area (e.g., reading achievement) two years in a row in all three grade bands (elementary, middle, high) fall into Step 1 of district improvement. These districts must inform all parents of their students of this status. This information may go out through a variety of formats but must include a letter and/or email to each student’s parent or guardian. The letter must include information about why the district was identified and how parents can assist the district in its efforts to improve student achievement.

Districts in Step 1 are required to *develop or revise a district improvement plan*. The plan must be developed or revised no less than three months after being identified for improvement and implemented no later than the beginning of the next school year. The development of the plan must involve parents, school staff, and others. The district improvement plan must:

- Address the fundamental teaching and learning needs of the district’s school(s), especially the needs of low-achieving students;
- Define specific measurable achievement goals and targets for each student subgroup;
- Incorporate strategies grounded in scientifically-based research that will strengthen instruction in core academic subjects;
- Include appropriate student learning activities before school, after school, during the summer, and during any extension of the school year;
- Provide for high-quality professional development for instructional staff that focuses on improved instruction;
- Include strategies to promote effective parental involvement in the district’s schools; and
- Include a determination of why the district’s previous plan did not bring about the required increase in student academic achievement.

Districts identified for improvement must *allocate 10% of its Title I budget* to address the professional development needs of teachers who work with the student groups that have not met AYP. If a district requests technical assistance from the state, the state is required to provide the assistance, which must be supported by effective methods and instructional practices that are based on scientifically-based research.

Step Two

A district that is in Step 1 moves to Step 2 if it does not make AYP in the same content area (e.g., reading achievement) two years in a row in all three grade bands (elementary, middle, high). Again, the district must inform all students' parents of this status. The information must be provided in a written format and may also go out in alternative formats. The letter must include information about why the district was identified, how parents can assist the district in its efforts to improve student achievement, and the steps for implementing the district improvement plan.

Districts in Step 2 of improvement are required to *implement the district improvement plan* that was developed in Step 1 by the beginning of the school year. The district must clearly address the actions that they have worked on with the state that will be implemented.

The state must continue to ensure the district is provided with technical assistance and **must** take at least one of the following corrective actions, as consistent with state law:

- Defer program funds or reduce administrative funds; or
- Institute and fully implement a new curriculum based on state and local content and academic achievement standards that includes scientifically research-based professional development for all relevant staff.

In conjunction with at least one of these actions, the state may also authorize parents to transfer their student from a school operated by the district to a higher-performing school that is not identified for improvement.

SCHOOL IMPROVEMENT

Step One

Schools that do not make AYP are in a “warning” year (also known as “Year 1”). Schools that do not make AYP in two consecutive years in the same subject and are considered to be in Step 1 of school improvement.

At the start of the school year, schools in Step 1 that receive Title I funds must notify the families of enrolled students about the *opportunity to transfer* their student to another school in the same district that is *not* identified for school improvement. Districts must use up to an amount equivalent to 20% of their Title I, Part A budget (unless a lesser amount is needed) to fund public school choice. Transportation costs (within federal parameters) must be covered by the district for families exercising this option. The school must also *develop or revise its school improvement plan*. The plan must be completed no later than three months after the school is identified for school improvement.

Step Two

Any school in Step 1 that does not make AYP in the same content area that did not make AYP the previous year advances to Step 2. The school must *continue school improvement planning* and *the district must continue to offer public school choice and must also provide supplemental educational services* to low-achieving students who are considered low-income (e.g., qualify for free/reduced lunch). Service providers must be selected from OSPI's state approved list of SES providers. Districts must use an amount equivalent to 20% of their Title I budget (unless a lesser amount is needed) to fund public school choice and supplemental educational services.

Step Three

Any schools in Step 2 that does not make AYP in the same content area that did not make AYP the previous year advances to Step 3. For schools in Step 3 (“corrective action”), *districts must select at least one of the following options* (and identify their own actions):

1. Make curriculum and instruction changes to improve student learning;
2. Appoint outside experts to work to advise the school on revising and implementing the school plan; or
3. Extend the school year or school day.

In addition, *the district must continue to offer public school choice and supplemental educational services*, and the *school must revise the school improvement plan*. The plan must include a description of the corrective action the school has instituted and how this action will lead to student success on the state's assessment measures.

Step Four

Any school in Step 3 that does not make AYP in the same content area that did not make AYP the previous year advances to Step 4. Districts are required to undertake “school restructuring” in Step 4 schools, which means a major reorganization of the school's governance. The district has one year to *prepare a restructuring plan*. Parents and teachers must be provided information that the school has entered Step 4 and provided the opportunity to comment on the proposed restructuring actions and assist in the development of the restructuring plan. The plan must be implemented no later than the beginning of the following school year, whether or not the school has move into Step 5.

The restructuring plan needs to include at least one of the following three actions:

- Replace school staff, which may include the school principal, who are relevant to the school's inability to meet standards;
- Enter into a contract with an outside entity with a demonstrated record of effectiveness, to operate the school; and/or
- Implement other restructuring activities that are consistent with the principles of restructuring.

The district must *provide technical assistance* that emphasizes (a) the importance of improving instruction by using strategies grounded in scientifically-based research so that all students achieve proficiency in the core academic subjects of reading and mathematics, and (b) the importance of analyzing and applying data in decision-making. The district must also *continue to offer public school choice and supplemental educational services* to all eligible students.

Step Five

Any school in Step 4 that does not make AYP in the same content area that did not make AYP the previous year advances to Step 5. In this step, the district must ensure that the school has *implemented the restructuring plan*. The district must also *continue to offer public school choice and supplemental educational services* to all eligible students.

APPENDIX H

ADDDITIONAL ACCOUNTABILITY DATA

The Accountability Index provides only a snapshot of school and district performance and does not include data on many other important topics. Additional information needs to be made available to educators, policymakers, and the public to provide a more complete picture of the school and district context, resources, and student outcomes. This includes data on the relative “challenge” that schools and districts face in terms of the types of students served. It also includes data for reciprocal accountability to recognize the role the community and state play in supporting education.

Data should be provided over time at multiple levels (school, district, state) so one can better understand the changes taking place and how they compare to the other levels. Looking at several years of data at a time recognizes that natural fluctuations occur from year to year. The data can be color-coded to show trends. Separate data sets should be produced for each school type (elementary, middle, high, K-12) so comparisons can be made for similar schools.

Data for the variables in the table below should be reported electronically on the state Web site. In some cases, only district or state data are shown. The data fall in eight categories: student and staff characteristics, financial information (district revenues and expenditures, community support), state support, student outcomes, and NAEP results (when applicable). Most are currently available—those in *italics* are not currently reported or collected by the state. In some cases, the data are available but not yet reported (e.g., data for individual grades are not aggregated across all grades in a school or district). Those that are not yet collected should be done so when funding is made available to collect and report the data.

RESULTS FOR	SCHOOL			DISTRICT			STATE		
	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09
STUDENT CHARACTERISTICS									
# of students (October count)									
% low income (eligible for subsidized meal)									
% English language learners (ELL)									
% with disabilities (in special education)									
<i>% mobile</i>									
<i>% gifted</i>									
% migrant									
<i>Challenge Index¹</i>									
<i>Challenge Decile²</i>									
% African American									
% American Indian									
% Asian									
% Pacific Islander									
% Hispanic (not White)									
<i>% multiracial</i>									
<i>% non-white</i>									
% White									
STAFFING CHARACTERISTICS									
Number of teachers (FTEs)									
Student/teacher ratio (Oct count/FTEs)									
<i>Student/counselor ratio (Oct count/FTEs)</i>									
<i>Average class size of core classes³</i>									
% teachers with MA or greater									
Avg. teacher experience (in years)									
% teachers National Board certified									
<i>% teachers with < 5 years experience</i>									
<i>% teachers in same school > 3 years</i>									
<i># years principal in current position</i>									

RESULTS FOR SCHOOL YEAR	SCHOOL			DISTRICT			STATE		
	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09
FINANCES – DISTRICT REVENUES									
Total funding per FTE pupil (unadjusted for student need or cost of living)									
% State funding									
% Local funding									
% Federal funding									
% Other funding									
Ending total fund balance, in % of total revenues at end of school year									
FINANCES – DISTRICT EXPENDITURES									
Total expenditures/FTE student									
Total expenditures/FTE student, adjusted for student need & cost of living ⁴									
% of total spent on district administration									
% of total spent on school administration									
% of total spent on teaching									
% of total spent on teaching support ⁵									
% of total spent on maintenance and operations									
% of total spent on transportation									
% of total spent on other activities ⁶									
FINANCES – COMMUNITY SUPPORT									
Total levy amount per FTE pupil									
Tax rate rank (out of 295, with 1 being the highest tax rate)									
General fund levy rate (per \$1,000 of assessed property value) ⁷									
Maximum general fund levy lid									Varies from 24% to 34%
% of general fund levy lid requested									
General fund levy assessed value per FTE pupil									
General fund levy amount per FTE pupil									
% voted "Yes" on general fund levy (requires >50% to be approved) ⁸									
% voted "Yes" on transportation levy ⁸									
% voted "Yes" on capital levy ⁸									
% voted "Yes" on debt service levy ⁸									
STATE SUPPORT (MOST RECENT DATA AVAILABLE IN THE YEAR)									
Student-teacher ratio rank (out of 50 states & DC; 1 is lowest ratio)									
Total revenues/pupil rank, unadjusted (out of 50 states & DC; 1 is highest)									
Total revenues/pupil rank, adjusted for cost of living (out of 50 states & DC; 1 is highest) ⁴									
Total revenues/pupil rank, adjusted for cost of living & student need (out of 50 states & DC; 1 is highest) ⁴									
Percent difference in total revenues/pupil (adj. for cost of living and student need) from national average ⁴									
Percent difference in per capita income (adjusted for cost of living) from national average ³									
Percent difference in minimum teacher salary (adjusted for cost of living) from national average ⁴									
STUDENT OUTCOMES									
RESULTS FOR SCHOOL YEAR	SCHOOL			DISTRICT			STATE		
	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09
Reading % meeting standard, ⁹ All students									
African American									
American Indian									
Asian									
Pacific Islander									
Hispanic (not White)									
Multiracial									
White									
Low income									
Not low income									
ELL									
Special education									
Reading Learning Index, ⁹ All students									
African American									
American Indian									
Asian									
Pacific Islander									
Hispanic (not White)									
Multiracial									
White									
Low income									
Not low income									
ELL									
Special education									

RESULTS FOR SCHOOL YEAR	SCHOOL			DISTRICT			STATE		
	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09
Writing % meeting standard, ⁹ All students									
African American									
American Indian									
Asian									
Pacific Islander									
Hispanic (not White)									
<i>Multiracial</i>									
White									
Low income									
Not low income									
ELL									
Special education									
Writing Learning Index, ⁹ All students									
African American									
American Indian									
Asian									
Pacific Islander									
Hispanic (not White)									
<i>Multiracial</i>									
White									
Low income									
Not low income									
ELL									
Special education									
Math % meeting standard, ⁹ All students									
African American									
American Indian									
Asian									
Pacific Islander									
Hispanic (not White)									
<i>Multiracial</i>									
White									
Low income									
Not low income									
ELL									
Special education									
Math Learning Index, ⁹ All students									
African American									
American Indian									
Asian									
Pacific Islander									
Hispanic (not White)									
<i>Multiracial</i>									
White									
Low income									
Not low income									
ELL									
Special education									
Science % meeting standard, ⁹ All students									
African American									
American Indian									
Asian									
Pacific Islander									
Hispanic (not White)									
<i>Multiracial</i>									
White									
Low income									
Not low income									
ELL									
Special education									
Science Learning Index, ⁹ All students									
African American									
American Indian									
Asian									
Pacific Islander									
Hispanic (not White)									

RESULTS FOR SCHOOL YEAR	SCHOOL			DISTRICT			STATE		
	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09
<i>Multiracial</i>									
<i>White</i>									
<i>Low income</i>									
<i>Not low income</i>									
<i>ELL</i>									
<i>Special education</i>									
% meeting stand. in 3 subj., ¹⁰ All students									
African American									
American Indian									
Asian									
<i>Pacific Islander</i> ¹¹									
Hispanic (not White)									
<i>Multiracial</i>									
<i>White</i>									
<i>Low income</i>									
<i>Not low income</i>									
<i>ELL</i>									
<i>Special education</i>									
<i>WLPT Composite, K-12 average level</i> ¹²									
<i>% Level 1 (no or limited English)</i>									
<i>% Level 2 (intermediate English)</i>									
<i>% Level 3 (advanced English)</i>									
<i>% Level 4 (exit)</i>									
<i>% ELL for 5 or more years</i>									
On-time graduation rate, All students									
Male									
Female									
African American									
American Indian									
Asian									
<i>Pacific Islander</i> ¹¹									
Hispanic (not White)									
<i>Multiracial</i>									
<i>White</i>									
<i>Low income</i>									
<i>Not low income</i>									
<i>ELL</i>									
<i>Special education</i>									
Extended graduation rate, All students									
African American									
American Indian									
Asian									
<i>Pacific Islander</i> ¹¹									
Hispanic (not White)									
<i>Multiracial</i>									
<i>White</i>									
<i>Low income</i>									
<i>Not low income</i>									
<i>ELL</i>									
<i>Special education</i>									
Unexcused absence rate, All students									
African American									
American Indian									
Asian									
<i>Pacific Islander</i> ¹¹									
Hispanic (not White)									
<i>Multiracial</i>									
<i>White</i>									
<i>Low income</i>									
<i>Not low income</i>									
<i>ELL</i>									
<i>Special education</i>									
Reading participation rate, ¹¹ All students									
African American									
American Indian									
Asian									

RESULTS FOR SCHOOL YEAR	SCHOOL			DISTRICT			STATE		
	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09	2006-07	2007-08	2008-09
<i>Pacific Islander</i> ¹¹									
Hispanic (not White)									
<i>Multiracial</i>									
White									
Low income									
<i>Not low income</i>									
ELL ¹²									
Special education									
Math participation rate, ¹³ All students									
African American									
American Indian									
Asian									
<i>Pacific Islander</i> ¹¹									
Hispanic (not White)									
<i>Multiracial</i>									
White									
Low income									
<i>Not low income</i>									
ELL ¹¹									
Special education									
Federal AYP status/Step (if applicable)									
College eligible rate, ¹⁴ All graduates									
<i>Male</i>									
<i>Female</i>									
<i>African American</i>									
<i>American Indian</i>									
<i>Asian</i>									
<i>Pacific Islander</i>									
<i>Hispanic (not White)</i>									
<i>Multiracial</i>									
<i>White</i>									
<i>Low income</i>									
<i>Not low income</i>									
ELL									
<i>Special education</i>									
<i>Other High School Outcomes</i>									
<i>% 9th graders passing algebra I</i>									
<i>% classes at AP/IB level</i>									
<i>% taking at least one AP/IB course</i>									
<i>% enrolled in dual enrollment program</i> ¹⁵									
<i>Average SAT (total of 3 subjects)</i>									
<i>% graduating seniors taking SAT</i>									
<i>Post-Secondary Outcomes</i>									
<i>% enrolling in CTC or 4-year college</i>									
<i>% taking remedial math courses</i> ¹⁶									
<i>% taking remedial English courses</i> ¹⁶									
<i>% completing one year of college credit</i> ¹⁷									
NAEP OUTCOMES (WHEN APPLICABLE) ¹⁸									
Grade 4 Reading, All Students (percent Proficient or Advanced)									
Grade 4 Reading, All Students (difference from US average)									
Grade 8 Reading, All Students (percent Proficient or Advanced)									
Grade 8 Reading, All Students (difference from US average)									
Grade 4 Math, All Students (percent Proficient or Advanced)									
Grade 4 Math, All Students (difference from US average)									
Grade 8 Math, All Students (percent Proficient or Advanced)									
Grade 8 Math, All Students (difference from US average)									

¹ The sum of the percent low income, ELL, special education, and mobile minus the percent gifted.

² A grouping of the Challenge Index that divides schools and districts into 10 roughly equal groups based on the number of students assessed (i.e., each decile has approximately 10% of the students assessed). Those with a decile of 1 have the lowest challenge; those with a decile of 10 have the highest challenge.

³ Elementary is average for regular classes; secondary is average of all English, math, and science classes.

⁴ Cost of living adjustments reflect differences in relative purchasing power compared to the national average. Student need adjustments recognize that some students need more help and cost more to educate (e.g., those who are learning English, have a disability, and are low income).

⁵ Includes guidance/counseling, learning resources, extracurricular, pupil management/safety, health/related services, and payments to other districts.

⁶ Includes food services, data/information services, and other expenditures.

⁷ Also known as Maintenance and Operations (M&O) or general excess levy.

- ⁸ Most recent vote; other conditions may apply to be approved (e.g., minimum turnout percentage). If missing, no levy was submitted for a vote. Bonds require a 60% Yes vote, but levies require only a majority vote.
- ⁹ Includes results for all assessed grades from both the regular state tests and alternate assessments (WAAS).
- ¹⁰ Reading, writing, and math. Schools that do not assess students in all three of these subjects do not have results reported.
- ¹¹ Combined with Asian results.
- ¹² Based on a combined score for reading, writing, listening, and speaking.
- ¹³ Percent assessed in all assessed grades; ELL students are exempted from the math assessment in their first year of attending a U.S. public school.
- ¹⁴ Percentage of graduating students who earned the number of credits that meets the HEC Board requirements for entry into a 4-year public higher education institution in Washington state.
- ¹⁵ Includes Running Start, Tech Prep, and any college courses offered within the high school system.
- ¹⁶ Percent of total students enrolled in WA public institutions (community/technical colleges and 4-year universities); does not include data from private higher education institutions or those outside Washington state.
- ¹⁷ Percent of students from the previous year's graduating class who complete at least one year of college credit.
- ¹⁸ The National Assessment of Educational Progress is given regularly only in grades 4 and 8 in reading and math; the first set of results is the percentage of students who scored at either the Proficient or Advanced levels, while the second set of results is the percentage point difference from the U.S. average scoring at the Proficient or Advanced levels.

APPENDIX I

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 has been assisted by a number of Washington state advisors. This diverse set of advisors reviewed the work that had been done to date, identified data that could be used to examine schools in the Struggling tier prior to their designation as a Priority school, and discusses numerous technical issues related to the proposed index and recognition system. Members of the advisory group are:

Dr. Karen Banks, Shelton SD (District Improvement Facilitator)
Ms. Maggie Bates, Hockinson SD (Acting Superintendent)
Ms. JoLynn Berge, OSPI (Federal Policy and Grant Administrator)
Dr. Phil Dommes, 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)
Mr. Lile Holland, Washington Association for Learning Alternatives (Exec. Director)
Dr. Feng-Yi Hung, Clover Park SD (Assessment/Evaluation Director)
Mr. David Iseminger, Lake Stevens SD (School Board)
Ms. Randi Ivancich, Bainbridge Island SD (Teacher Specialist for Assessment)
Dr. Nancy Katims, Edmonds SD (Assessment Director)
Dr. Bill Keim, ESD 113 (Superintendent)
Dr. Dennis Maguire, Pasco SD (Associate Superintendent of Instruction)
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)
Dr. Lorna Spear, Spokane SD (Executive Director for Teaching and Learning)
Dr. Alan Spicciati, Highline SD (Chief Accountability Officer)
Ms. Holly Williams, Evergreen SD (School Board)

Other stakeholders are involved as well. OSPI staff working in the areas of assessment and data collection, student information, federal programs and accountability, and state assistance for schools and districts are involved in the discussions. Multiple interactive briefings across the state generated more feedback. In addition, a working group that focuses on System Performance Accountability provides feedback on issues related to the state accountability plan. Various members and staff of the SBE attend meetings of this group as well. Members of the working group (other than SBE members) are:

Mack Armstrong, Washington Association of School Administrators (Assistant Executive Director)
Mike Bernard, Association of Washington Business
Phil Brockman, Ballard High School, Seattle Public Schools (Principal)
Marc Cummings, Battelle (Public Affairs Director)
Karen Davis, Washington Education Association (Government Relations)
Larry Ehl/Caroline King, Partnership for Learning
Roger Erskine, Professional Educators Standards Board (Board member)
Edie Harding, State Board of Education (Executive Director)

Bob Harmon, OSPI (Assistant Superintendent, Special Programs and Federal Accountability)
Mary Alice Heuschel, Superintendent, Renton School District
Glenn Johnson, Washington Association of School Administrators (Cashmere School District Superintendent)
George Juarez, Superintendent, Othello School District
Bruce Kelly, ESD 113 (Content Specialist for Math and Science)
Gary Kipp, Association of Washington School Principals (Executive Director)
Janell Newman, OSPI (Assistant Superintendent, District and School Improvement and Accountability)
Don Rash, Association of Washington School Principals (Director of Middle Level Programs and Assessor/Mentor & Intern Programs)
Martha Rice, Washington State School Directors Association (President, Yakima)
Marilee Scarbrough, Washington State School Directors Association (Policy Director)
Ben Soria, Washington Association of School Administrators (Yakima School District Superintendent)
Ted Thomas, Washington State School Directors Association (Past President, Longview)
Anne Walker, Wiley Elementary School, Richland School District (Teacher)
Steven Warren, Centralia Middle School, Centralia School District (Principal)
Bill Williams, Washington State PTA (Executive Director)

Members of OSPI's National Technical Advisory Committee reviewed the details about the proposed index and provided feedback. Member of this committee are:

Dr. Patricia Almond, Professor, University of Oregon
Dr. Peter Behuniak, Professor in Residence, University of Connecticut
Dr. Richard Duran, Professor, University of California–Santa Barbara
Dr. George Engelhard, Professor, Emory University
Dr. Robert Linn, Professor Emeritus, University of Colorado and UCLA/CRESST
Dr. William Mehrens, Professor Emeritus, Michigan State University
Dr. James Popham, Professor Emeritus, University of California–Los Angeles
Dr. Joseph Ryan, Professor Emeritus, Arizona State University

Finally, four accountability experts located in Washington DC but not associated with the federal government reviewed early drafts of this document and provided helpful feedback.