

BELLINGHAM SCHOOL DISTRICT
Bellingham, Washington

MEMORANDUM

TO: Board of Directors
FROM: Dr. Greg Baker, Superintendent
DATE: January 3, 2018
SUBJECT: Ends Monitoring Report 2.1, Part 1

I am pleased to submit this Ends monitoring report (End 2.1 - Student Competence, Part 1) to the school district's board of directors. Ends 2.1, Part 1 explores student performance in reading, math and science, and compares our progress to comparable high-performing districts. Part 2 of Ends 2.1 will be presented at the February 16 board meeting and will focus on students' continuous improvement toward graduation in comparison to high-performing peer districts. Part 3 will be presented at the March 23 board meeting focusing on students' participation in post-secondary education and career preparation in comparison to high-performing districts, along with a summative conversation in anticipation of the Board's evaluation of our reporting on Ends 2.1 overall.

Introduction

The report that follows focuses on the Bellingham School District's student achievement performance levels overall on the Smarter Balanced assessment (SBA), and compares our progress to selected comparable high-performing districts. We also include additional comparison data that highlight how Bellingham students' test performance measures up to all the states who are included in the Smarter Balanced Assessment Consortium (SBAC), as well as briefly touching on the results of a recent national study that took a five-year look comparing student achievement growth in districts all over the United States, including Bellingham.

By way of reminder, the Ends 2.1 policy follows:

E - 2.1: Consistent with the district Vision and Mission, all children of the Bellingham Public Schools Community will attain high academic achievement, develop essential skills and attributes necessary for continuous growth in learning, and graduate from high school. All students will succeed and grow regardless of ethnicity, socio-economic status, English language proficiency or disabilities.

1. Every student’s achievement, skills and attributes will exceed: (a) the Washington State benchmarks and (b) similar students in comparable high performing districts, as measured by state assessments and other available data, as appropriate.
2. Every student’s achievement, skills and attributes will show continuous significant growth relative to similar students in comparable high performing districts as measured by state assessments and other available data, as appropriate.
3. Every student with a gap in achievement, skills and attributes will close the gap. For state or federally identified student populations, any gap will be eliminated and annual progress will be greater than that of similar students in comparable high performing districts.
4. All children of the Bellingham Public Schools community shall make continuous advancement toward on-time graduation or extended graduation, thereby reducing Bellingham Public Schools’ drop-out rates.
5. After graduation, student participation in post-secondary education and career preparation shall increase and exceed participation than in comparable high-performing districts.

In this Part 1 report, we focus on numbers 1-3 above. The remaining items are the subject of future Ends reports this year.

Methodology for Identifying Comparable High-Performing Districts

Ends 2.1 relies on big picture comparisons to a total of forty-three districts that were selected on the basis of being somewhat demographically similar to Bellingham. The main sections of our report focus in on three comparable high-performing districts, Bellevue, Olympia and Shoreline, that comprise a group of strategically chosen demographic peer districts that typically outperform our system on measures of student achievement and graduation rates. Table A below includes the parameters that guided the choice of comparison districts. Table B below arrays Bellingham’s key demographic data alongside of the three comparison high-performing peer districts.

Table A: Parameters to Determine Comparable District Pool

Criteria	Parameters	Low	Bellingham	High
Enrollment	0.67 above and below	3,811	11,548	19,285
% Free/Reduced Meal	no lower than 0.5 below	18	35	100
% Asian and White	no higher than 0.25 above	0	74	92
% Limited English	no lower than 0.5 below	3	7	100

Table B: Demographic Data for Bellingham and High Performing Comparison Districts

Criteria	Bellingham	Bellevue	Olympia	Shoreline
Enrollment	11,548	20,431	10,116	9,630
% Free/Reduced Meal	35	18	28	26
% Asian and White	74	76	76	68
% Limited English	7	14	2	7

▪ **Smarter Balanced Assessment Consortium Comparison**

The SBA test is given by a number of states around the United States and as such creates the possibility for us to compare how our district and state performance measures up to those other states that use the same test.

As the data on page 1 of the attached data set reveal, Washington State is at or near the top-performing state over all when compared with other SBAC states. Understanding where Washington falls as a state provides a better context for understanding what it means to exceed state percentages. Our district’s students tend to outperform our state average, and in fact outperform many of the other states that administer the Smarter Balanced exam. As a district, Bellingham students scored above the average in Washington state on all dimensions of the test, with the exception of third grade English language arts and math. Third graders’ performance this past spring has us exploring a variety of factors we believe may be influential. In a later section of this report, we go deeper into our third grade data and look at a comparison between the test information and the data from the third grade report cards generated by our Bellingham teachers.

▪ **2017 Proficiency Percentiles -- Bellingham Compared to 43 Comparable Districts**

Pages 2-3 of the data set show the proficiency bar charts of the forty-three districts in the larger comparable set, plus Washington state. The dark blue highlighted line shows where Bellingham is positioned versus the three comparison high-performing districts (highlighted in red) in terms of student achievement on the Smarter Balanced test in both ELA and math, grades 3-8 in 2017.¹ As the bar chart illustrates, student populations from Bellevue, Shoreline and Olympia, the three districts in our comparison set, achieve at a higher rate than Bellingham students overall. Bellingham students scored at the 82nd percentile in English language arts, which ranked us #12 out of the forty-three comparable districts in 2017 English language arts (page 2). As the board knows, the three high-performing comparable districts we have relied on for our “high bar” comparisons are Bellevue (ranked #1 overall), Shoreline (#2) and Olympia (#4). In mathematics, Bellingham students scored at the 84th percentile, which ranked us #11 out of the forty-three comparable districts (see page 3 of the data set). Bellevue, Shoreline and Olympia ranked #1, #3 and #4 respectively in this same group.

¹ We include only limited high school testing data in our comparisons due to its instability. With a number of students’ families opting them out of taking the state test in high school, the data picture we are left with (in our district as well as in the comparison group of districts) is an incomplete picture of student performance levels.

We know historically that income levels are highly correlated with test scores. The choice to compare our district to Bellevue, Shoreline and Olympia, all districts with smaller percentages of low income students, puts us always in the likely position of comparing unfavorably. Given this, we have also included an analysis of how Bellingham's students compare with the group of districts that are within +/- 5% of us in terms of the percent of students eligible for free/reduced price meals. Page 4 arrays these comparable districts by percent of students eligible for free/reduced price meals. By definition, Bellingham falls in the exact middle of this group of comparable districts when arrayed this way, with half of the districts within +5% and half within -5% of Bellingham's free/reduced percentage. Page 5 arrays these same districts by the overall average of students meeting standard. As the graphic reveals, Bellingham students' performance is slightly above middle when arrayed by percent meeting standard, suggesting that by comparison within this +/-5% group, our students performed slightly better on average for 2017.

▪ **District Smarter Balanced Scores Compared to Comparable High Performing Peers**

2.1.1 (Exceed state benchmarks). In the data set provided for the board, we also present overall comparisons against the state benchmarks on the first three years of the Smarter Balanced test to the selected high-performing comparable group of school districts, namely Bellevue, Shoreline and Olympia. Several different indicators of achievement are tracked and presented. These include:

- ELA in grade bands 3-5 and 6-8;
- math in grade bands 3-5 and 6-8;
- science in grade bands 3-5, and 6-8.

Pages 6-10 of the data set compare percentages of Bellingham students who *met* standard on the state test with percentages of students from the other three comparable high performing districts, as well as with the state overall. Pages 11-15 compare the percentages of Bellingham students who *exceeded* standard to our comparison group. Bar graphs include student overall comparisons, and comparisons on the four subgroup populations we have typically tracked: Hispanic/Latino students; English language learners, students with individualized education plans (IEPs); and students eligible for free and reduced price meals. We have included the 2015 and 2016 data alongside the 2017 data to give at least a visual nod to how these scores are trending over this three-year period.

Bellingham students overall, as well as within examined subgroups, tended to outperform the state averages on the Smarter Balanced test once again this year, and underperform those students from our high performing comparable district set (Bellevue, Olympia, Shoreline). Scores for students in Bellingham were generally down slightly across the board in the overall picture (page 6), as well as within subpopulations of students we track including Hispanic students (page 7), students receiving English language learner support (page 8), students with individualized education plans (page 9), and students who are from low income households (page 10). We also noted these across-the-board decreases were true for our highest-performing comparable district, Bellevue, as well as for many student subcategories in Shoreline, Olympia and for Washington state overall.

A similar picture emerged when we looked at the percent of students exceeding standard, comparing Bellingham with our high bar group of comparable districts. Percent of Bellingham students exceeding standard was down overall (page 11), as well as for subgroups of Hispanic students (page 12), students receiving English language learning supports (page 13), and low income students (page 15). Trends for students with individualized education plans were more mixed (page 14), with slight decreases in the elementary grade band in English language arts and science, and flat scores or slight increases in the middle level grade band.

Why these score decreases in year-to-year comparisons? Normal fluctuations, or a negative trend? Are our students in the cohort that tested this year less capable, or worse yet, learning less, than the cohort that tested last year, or the year before? It is possible that the decreases we see in this year's data resulted from differences between the 2017 cohort of students and those from the prior two years. Further, as a district, we have not mandated a heavy emphasis on "test prep" as a high priority in our schools, which makes this test truly a singular springtime event for our students. This fact also may have some bearing on how we measure up to other districts in the region. We also recognize that we need to continue to work to support our teachers' full understanding of the state standards, and what it means to teach to those standards. This is a part of our ongoing professional development work across disciplines.

We also explored some alternate hypotheses for the decreases we saw this year. One possible additional explanation for the drops is that the SBAC test is still new and, as such, different versions of the test are still rolling out. Psychometricians do the best they can at establishing equivalent versions of standardized tests. However, comparisons of the 2017 results with last year's scores showed a decline in English language arts scores in every one of the fourteen states that rely on the SBAC test. This is considered by at least one expert to be highly unusual. Edward Haertel, professor emeritus at the Stanford Graduate School of Education and an internationally recognized psychometrician, is on record stating that uniform declines such as those seen this year are unusual and raise questions about the test itself. "The fact that 14 out of 14 (SBAC) states show changes in the same direction pretty much confirms that these are not merely chance fluctuations. The obvious question that arises is whether there's some reason the numbers for this year versus last year are noncomparable," he said. "It's possible there was in fact some slight overall decline in students' proficiency and the test results are accurate, but as a psychometrician, I'd want to work hard at eliminating rival hypotheses before settling on that conclusion."²

We also took a deeper look into the third grade data, as it was unusual to see our students' performance drop below state averages in that specific cohort. Several teachers and principals in the district who observe the students in the testing setting have wondered about the nature of the test itself, and whether taking the test on the computer adds layers of complexities to understanding what third grade students actually know and are able to do. Page 16 in the data set provides a comparison of Bellingham third graders' SBA math scores with teacher report card marks for mathematics from the final report card given last spring to those same third graders.

² Fensterwald, J. (2017, September 27) "Average scores flat in third year of California's Common Core-aligned tests." Retrieved from URL <https://edsources.org/2017/no-growth-statewide-in-third-year-of-smarter-balanced-test-scores-caaspp/587973>.

To equate these two different measures and allow us to do a comparison, we established a mid-point that constitutes “meeting standard” on both measures. The first chart on page 16 contrasts students eligible for free and/or reduced price meals with non-eligible peers, and juxtaposes teacher report card marks for both groups with the SBA averages for both groups. As the graphic reveals, teacher marks in some but not all schools tended to be slightly lower than what SBA scores would have predicted for both groups, and while there is still a gap between the free/reduced eligible students and non-eligible peers in most schools, that gap fluctuates in size between the teacher-given marks and what shows up on SBA. The second chart does the same comparison with the Hispanic subgroup compared to white peers. Again here our teachers assigned marks that were in some cases slightly lower than the SBAC scores, but in other cases nearly exactly aligned, or slightly higher than the corresponding test scores. The gaps between these groups of students in terms of the teacher marks virtually disappears in a few of our schools, and is significantly smaller than what the test scores revealed for other schools.

These data raise further questions for us. Does one give the test score, which is a snapshot in time, more credence than the teachers’ collective assessment of student progress, based on their daily observations over the course of an entire school year? In cases where the average of teacher marks was below the test scores, does this signify that our teachers are grading students even more stringently than the exam? The bottom line is that we don’t know for sure what is underneath the declines this year, but with teacher marks that appear to be generally slightly lower than the test scores, it adds credence to the idea that the test wasn’t far off the mark in terms of accurately assessing overall student performance. With now two years of slight drops in SBA scores in many categories, we continue to delve into the questions this raises about the test, about learning and about how the test represents what kids actually know and are able to do.

Finally, a bit about progress that we can report on at the high school level. As we have reported in previous years, because of the relatively high numbers of students opting out of the SBAC test at the high school level, we are certain that the results reported by the state are not representative of the actual performance abilities of our high school population. Other measures have more credibility, however, such as the data we have on students meeting standard on high school proficiency exams. Page 17 of the data set shows the results of the Grade 10 Biology End of Course exam, and the Grade 11 English language arts Smarter Balanced test. In year over year comparison with 2016 data, 2017 results show proficiency levels for Bellingham students are well above state averages and rising overall on both exams, as well as for Hispanic and low income subgroups. Biology results put Bellingham students’ performance higher than either Shoreline or Olympia in 2017 for all students, as well as for subgroups of Hispanic students and low income students. Hispanic and low income subgroups of Bellingham students also scored higher proficiency levels on the English language arts exam than did those same subgroups in Bellevue.

2.1.2 (Show continuous significant growth).

The growth discussion for this report is significantly less robust than in last year’s report due to the fact that the Office of the Superintendent of Public Instruction (OSPI) still has the student growth data for districts embargoed, subject to internal review. Given that fact, we report on what we know about growth at this point in time, and we are happy to accommodate the board

with more information about this aspect of student progress over the course of future meetings, dependent on when OSPI releases the data for public view.

Student growth percentile is an indicator that shows how a given population of students grew compared with the expected growth overall from the state. The median state growth score sits precisely at the 50th percentile; this represents the overall midpoint of growth scores across all districts in Washington. So, when we compare ourselves on growth, we like to see percentiles both above 50% (doing better growth-wise than the state) and as high or higher compared to our high-performing comparable district peers.

The bubble graph included in the data set on page 18 offers a three-dimensional view of how we compare to the group of 43 districts that factors in both SGP and income level along with proficiency rates. The y-axis represents proficiency on the state test, the x-axis factors in SGP, and the circle size represents the magnitude of each district's population of students eligible for free or reduced meals. For 2017, we know that Bellingham students' scores (average of grades 4-8) were slightly above the state average (green center circle) in percent meeting standard in English language arts and math, and slightly below the state average in growth. While we are unable to show the data on the high-performing comparable districts due to the embargo, we do know that our student growth scores were lower overall than those in the comparison group. We do see a bright spot in that Bellingham's low income student population appeared to score comparably to Bellevue in math, however the low income subpopulation scored below all three of our high-performing comparable districts in English language arts.

2.1.3 (Close achievement gaps).

The final section of the report focuses on the degree to which achievement gaps between subgroups is shrinking or closing. Pages 19-21 display the proficiency gaps for three of the subgroup populations we track annually—students who are low income versus those who are not, students receiving special education services versus those not receiving services, and Hispanic students versus white students. These are presented again in grade bands (3rd-5th and 6th-8th) and show the difference between the sub group population and their counterpart comparison group.

The trends we aim for in this data are reductions in year-over-year comparisons. So anywhere the 2017 bar (dark blue bars on the graphic) is smaller than the 2016 bar (gray bar on the graphic) indicates a reduction in the gap separating the two subgroups being compared. We've also included the 2015 data to support a longer term look at trends in this data. Each graph also includes a "candy cane" column (red horizontal stripes) that shows the relative percentage of each subgroup for each of the four districts and the state. As we've reported in previous years, this is an area where we've seen significant gaps that separate subgroups, and has been a strong focus of our equity work as a system, trying to ensure greater and greater opportunities for all students regardless of income, race, or learning handicap.

The non-low income to low income gaps (page 19) fluctuated depending on grade level band and content area. We again noted a slight year-over-year increase in the gap in science in both grades 3-5 and 6-8. In grades 3-5, this gap was slightly larger in ELA, and flat in math but

smaller relative to Bellevue. In grades 6-8 the gaps in both ELA and math decreased for this population and also put us in a better position than Bellevue regarding the size of these gaps this year.

With regard to proficiency gaps between students with and without an IEP (page 20), we see the continuation of a positive trend overall, with gaps shrinking for both 3-5 and 6-8 grade bands in ELA and science. In math, the grades 3-5 gap was slightly larger and grades 6-8 trended smaller. The gaps for Bellingham students with IEPs were also significantly smaller than our high performing comparable districts in several areas including 3-5 and 6-8 math, 6-8 science, and comparable to both Bellevue and Shoreline in 3-5 ELA. We also noted that the gaps for students with IEPs from our highest-performing comparable district, Bellevue, increased in all areas except 6-8 ELA, while many of the gaps for our students with IEPs narrowed.

Comparing gaps in proficiency between white and Hispanic students (page XX), Bellingham achievement gaps are notably smaller at the elementary level (3-5) across all three content areas, but larger at the middle level in all three content areas. While our gaps were slightly lower than Bellevue's in some areas, we see that the performance gaps of our Hispanic students relative to white peers overall tends to be larger than high-performing comparable districts, and higher than state averages.

These persisting gaps show that we continue to have a long way to go toward our goal of ensuring that ALL students are enabled and supported to perform at standard. This remains also true for our state overall, and for the three comparable high-performing peer districts. That said, we were glad to see continued progress this year in closing achievement gaps in the area of special education, and this appears to be trending in the right direction.

Five-Year National Growth Comparison

Finally, we are including and commenting on some additional, newly available data for this report this year. Comparing and reporting on year-over-year test score averages versus other districts in the neighborhood gives a less than satisfactory longer-term view of how we are doing as a system. The typical student cohort fluctuations show that as a district we are down or up in any given year; these fluctuations ebb and flow depending on the year and the cohort. The obvious goal is for the district to show general improvement in these measures over time. This kind of a longitudinal look that compares us with other districts across the county has not been available until just recently. We've learned that our district was included in a national, longitudinal study of student proficiency and growth that took a five-year look at a large number of districts and assessed which districts seem to add the most value to students' learning and growth over time³. The study analyzed standardized test scores from roughly 45 million students

³ Riordan, S. (2017). Educational opportunity in early and middle childhood: Variation by place and age. (Working Paper No. 17-21). Retrieved from Stanford University Center for Education Policy Analysis website: <https://cepa.stanford.edu/sites/default/files/wp17-12-v201712.pdf>

to describe the structure of educational opportunity (aka, learning) in over 11,000 school districts. Bellingham was included as one of the examined districts. For each school district, the authors constructed two measures: the average academic performance of students in grade 3 and the within-cohort growth in test scores from grade 3 to 8. What this study was attempting to show was, taking into account where students scored as third graders, how much growth occurred for those same students over this five-year period.

Our third grade cohort of students that was included in the study began their schooling about a half year above grade level compared to peers around the country. Over the course of five years, the study found that Bellingham students on average made 6.1 years of growth, adding more than a year's worth of growth beyond what would have been predicted on average. As the graphic we'll show in the board meeting suggests, this puts Bellingham's overall student performance (in terms of proficiency and growth) at the top of school districts with similar poverty levels around the United States.⁴ The opportunity to see how we stack up in this national, longitudinal study is both gratifying and adds fuel to our resolve to continue to work to improve the district for all students.

Concluding Statement

We offer this Ends 2.1 (Part 1) monitoring report, in combination with the remaining parts of the report that will follow, as evidence of a reasonable interpretation of Ends 2.1 that aligns with our vision, mission and outcomes, and is supported by data that demonstrates progress toward achievement of these Ends. Further, we hope this report serves as a useful tool in support of the board's ability to regularly review our ends to ensure they remain relevant and inspire meaningful work throughout the organization and community.

⁴ Badger, E. and Quealy, K. (2017, December 5) "How effective is your school district? A new measure shows where students learn the most." Retrieved from URL <https://www.nytimes.com/interactive/2017/12/05/upshot/a-better-way-to-compare-public-schools.html>