

Breaking Down School Labels: How Can Two Elementary Schools with Significantly Different Student Test Scores Be “Performing”?

Executive Summary

Accountability has become an important part of the nation’s public education system. Schools and districts are being held accountable for how well their students perform in a number of areas, including standardized tests. In Arizona, accountability labels are assigned by the Arizona Department of Education each year and provide information about student performance at each school. These labels have important implications for school improvement decisions.

Arizona’s accountability system labels schools as Failing to Meet Academic Standards, Underperforming, Performing, Highly Performing or Excelling, based on selected indicators of student success, such as standardized test results. Understanding the information each label provides is important to interpreting what it conveys about how well a school is doing. In this Policy Brief, a case study is used to demonstrate how two schools with dramatically different student test results can receive the same accountability label of Performing.

Accountability determinations from Washington Elementary school in the Washington Elementary district and Boulder Creek Elementary school in the Paradise Valley Unified district are closely analyzed to understand how the Education Department’s formula includes different data components to produce a label for each school. The case study illustrates that, although the most recent test data reveal these are two dramatically different schools, several factors work together to label both as Performing. These factors include the following:

- *Three different measurements that evaluate average test results over a period of three to five years. One year of test data from a single test will not explain how a school received a label of Performing.*
- *The incorporation of improvement in test scores over time, which also cannot be captured by simply looking at one year of test data. For example, test scores for Washington Elementary school showed dramatic improvement over five years in some areas while Boulder Creek showed decreases in test results over time in many areas.*

- *To receive a label that is above Performing, the formula requires a certain percentage of students to exceed the standard on Arizona’s AIMS test. While nearly 27% of students at Boulder Creek exceeded the state’s standards, that was not enough to place the school in a higher performance category. Boulder Creek met some but not all of the criteria to be an Excelling or Highly Performing school; thus it was labeled as Performing despite relatively high test results.*

Because 70% of the state’s schools have been labeled as Performing, wide disparities can be found in test results, student body makeup and school type. This means two very different schools can receive the same label, and thus a school’s label alone does not provide a complete picture of student performance over time.

Introduction and Purpose

Since the fall of 2002, the Arizona Legislature has mandated that each public school’s performance be evaluated based on a set of indicators established by the Arizona Department of Education and approved by the State Board of Education. Currently, the components fit together into a single statewide system that assigns each school an Achievement Profile, labeling it as Failing to Meet the Academic Standards, Underperforming, Performing, Highly Performing or Excelling.¹

Using a case study approach, this Policy Brief explains how two elementary schools with significantly different student academic performance on standardized tests can receive the same Performing label.² This Brief also will describe the components of Arizona’s current public school accountability formula and explain how they work together to produce a performance label for each school.

Background

In November 2000, Arizona voters passed Proposition 301, an education reform and tax initiative that, among other things, required the Arizona Department of Education to develop, refine, and implement new accountability

programs for the state’s elementary and secondary public schools. The purpose of these programs is to enable the public to evaluate and monitor the progress of the state’s public schools based on the academic performance of their students.

The basic philosophy behind education accountability is that schools and districts are responsible for ensuring the academic success of their students. It is believed that student performance on standardized tests provides a relevant and valid measure of how well a school is teaching its students and how much the students are learning.

To assess the progress of public schools, the Arizona Department of Education is required to release an annual evaluation of each school, called an Achievement Profile.³ In general, the evaluation is an assessment of student performance on two standardized tests and related academic indicators such as a school’s graduation and dropout rates. The outcome of the evaluation places a school into a performance category. These categories, or labels and associated consequences,⁴ are intended to ensure that schools and districts are held accountable for their students’ successes.

What is an Achievement Profile?

An Achievement Profile is a label that provides a general academic profile of a school based on indicators of student success, including performance on two standardized tests:⁵ the Arizona Instrument to Measure Standards (AIMS) and the Stanford-9 (SAT-9).⁶ It is determined by a complex formula that assigns one of five profiles to a school: Failing, Underperforming, Performing, Highly Performing or Excelling. A school can only be labeled Failing, however, after receiving a profile of Underperforming for three consecutive years.⁷

As a result of how Achievement Profiles are calculated, it is possible for schools with very different student academic achievement, as indicated by test results, to be labeled similarly. To demonstrate this, student test performance data from two elementary schools, Boulder Creek Elementary in the Paradise Valley Unified district and Washington Elementary in the Washington Elementary district, are examined in this Policy Brief. Following a school-to-school test score comparison, each school's Achievement Profile determination is analyzed to explain how two schools, whose students performed very differently on standardized tests, can receive the same label of Performing.

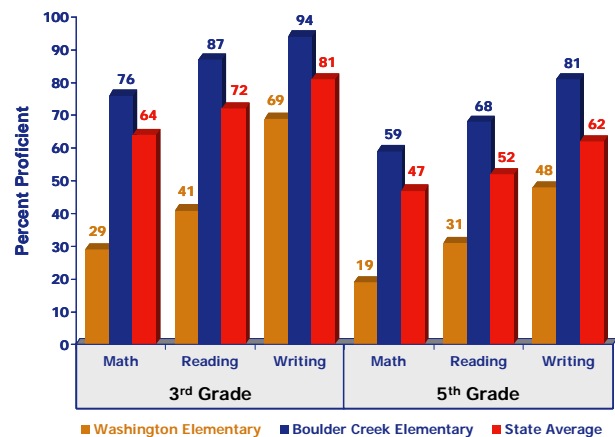
How do the test results for these two schools compare?

The AIMS test is a statewide assessment used to measure a student's proficiency of Arizona's academic content standards in reading, writing and mathematics. Since the 1999-2000 school year, the AIMS test commonly has been used to compare student performance from school to school in Arizona. On AIMS, student performance is generally expressed in one of four categories: Falls Far Below the Standard (FFB),

Approaches the Standard, Meets the Standard and Exceeds the Standard. A student who has demonstrated proficiency in a subject area is considered to have met the academic standard in that subject. A school's overall performance then can be expressed as the percentage of students who meet the standard in each subject and grade combination.⁸

An examination of the 2004 AIMS data for these two schools reveals wide gaps in their student proficiency levels, as seen in Figure 1.⁹ In third-grade math, for example, 76% of valid Category 1¹⁰ students are considered to have met the state's standard at Boulder Creek compared with 29% at Washington Elementary. Likewise, in third-grade reading, 87% of students are considered proficient at Boulder Creek while 41% achieved proficiency at Washington Elementary.

Figure 1. 2004 AIMS Proficiency by Subject and Grade



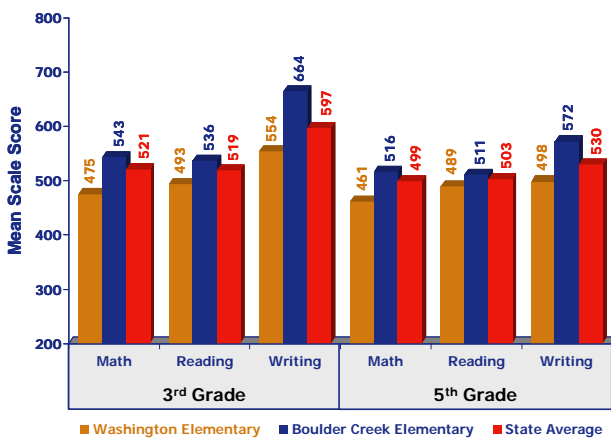
Source: Arizona Department of Education

A student's success also can be measured by the magnitude of performance, not only on whether the proficiency standard is met. In general, the higher the average score on an assessment, the more successful a student is thought to be. For these two schools, the higher the average score on the AIMS test, referred to as a mean scale score, the better the students have performed in

that subject and grade. A student is considered proficient in a subject and grade if a scale score of 500 or higher is achieved.

Similar to the differences in the proficiency levels in Figure 1 are differences between the two schools in the mean scale score across all subjects and grades (Figure 2). In Washington Elementary, only third-grade writing results show an average score above 500 while Boulder Creek averages above 500 in all subjects and grades, and the mean scale scores are significantly higher in most areas. The largest difference between the two schools can be seen in third-grade writing, which shows a 110-point difference in the mean scale score.

Figure 2. 2004 AIMS Mean Scale Score by Subject and Grade



Source: Arizona Department of Education

It is also important to note that in most subjects and grades, the average scale scores at Washington Elementary are below 500, which is below passing. In contrast, Boulder Creek is above passing in all subjects and grades.

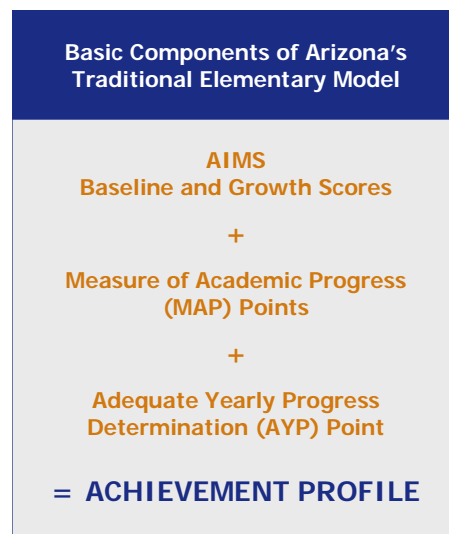
How is it then that although the students at Boulder Creek performed significantly better than the students at Washington Elementary, both schools received Performing labels? The answer lies in the components of the

Achievement Profile formula and how they work together to produce a school classification.

What are the Basic Components of Arizona's Accountability Model?

The law currently requires that three major components be included when determining the Achievement Profile of a traditional public elementary school:¹¹

Figure 3. Components of the Traditional Elementary Achievement Profile Model¹²



AIMS. As a trend model, the Achievement Profile calculation uses up to five years of AIMS data to show a performance trend for the school's students. The beginning year for AIMS data is the 1999-2000 school year unless a school opened after that date. The AIMS component of the Achievement Profile formula has two parts: baseline student scores and student performance growth. They are used to assess past student performance and whether student performance has changed over time.

MAP. The Measure of Academic Progress (MAP) results provide general information about individual student performance on a national assessment¹³ from one year to the next.

Until 2005, the national test administered in Arizona during the spring semester in grades two through nine was the SAT-9.¹⁴ The MAP component is currently weighted to comprise up to 40% of a school's total possible points and is based on the percentage of students that demonstrate one year's academic growth on the SAT-9 test.

AYP: The third component of the elementary model is the addition of an Adequate Yearly Progress (AYP) point. The AYP determination is calculated using a federal accountability formula and evaluates whether a school meets certain performance objectives each year. The AYP component accounts for just one point in the total formula regardless of the progress made by students at a school.¹⁵

The three formula components work together to produce a label for each elementary school in Arizona. To simplify the complexity of the formula, this Brief organizes the calculation into five steps and uses actual data from two schools to illustrate how the formula operates.

Five Steps to a Performing Label: Evaluating the Use of Test Data in the Achievement Profile Formula¹⁶

AIMS Component

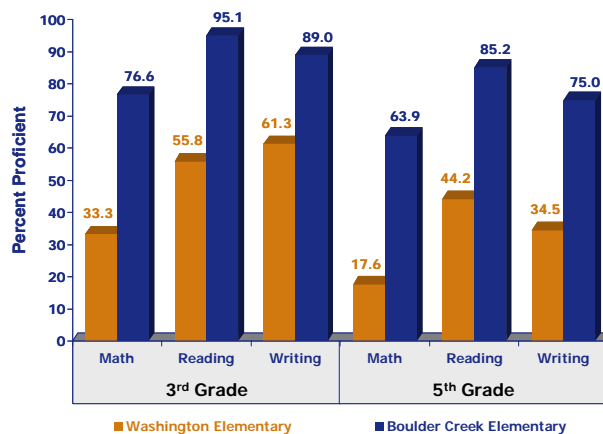
The AIMS component of the formula examines performance for each subject and grade and uses a combination of three, four or five years of data, depending on how many years of test data are available for a school. Two portions of the AIMS component are evaluated under this model: baseline AIMS scores and growth on AIMS over time. It is important to note that the AIMS results used in the actual calculation differ from those that are publicly reported. This portion of the calculation uses Category 1 AIMS results; however, unlike the publicly reported AIMS data shown in Figure 1 and Figure 2, it

does not include students who are considered mobile. This is a unique selection criteria used only for accountability purposes.¹⁷ Therefore, comparisons between publicly reported AIMS data and the AIMS data used in the Achievement Profile calculation will differ, in some cases dramatically.

Step 1. AIMS Evaluation: Determining Baseline Proficiency

To assess whether AIMS results have improved over time, a starting point for each school must be evaluated. This starting point, called the *baseline*, is calculated by using the percentage of students meeting or exceeding the standard in reading, writing, and mathematics for the baseline year(s).¹⁸ The baseline percentages for Washington and Boulder Creek elementary schools are calculated using the average percentage of proficient students from the 2000 and 2001 school years (Figure 4).

Figure 4. Baseline Proficiencies (2000 and 2001)



As shown in Figure 4 above, there are wide gaps in the baseline scores for these schools in every subject and grade evaluated by the formula. Third-grade reading, for example, shows an average of nearly 56% proficient in the baseline years for Washington Elementary compared with more than 95% proficient for Boulder Creek. In fifth-grade math, the proficiencies are nearly 18% and 64%, respectively. In this particular piece of the formula, Boulder Creek

has much higher proficiency rates in the baseline years, but it is important to note that this is only one part of the overall determination.

Step 2. AIMS Evaluation: Measure of Growth

To assess growth, the most recent three years of AIMS data are averaged and compared with the AIMS results in the baseline years. The amount of growth then is measured by a combination of two factors:

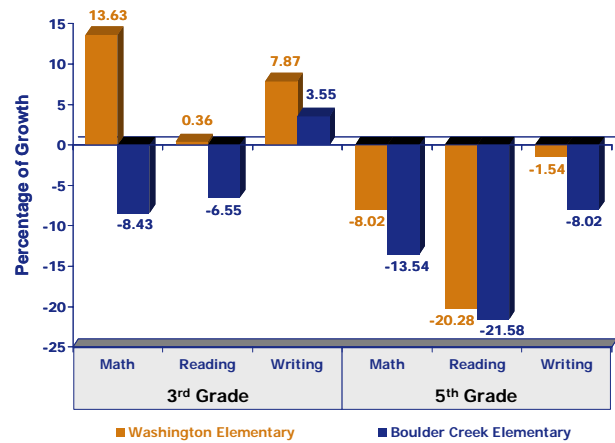
1. an increase in the percentage of students who are meeting or exceeding the standard.
2. a decrease in the percentage of students who fall far below (FFB) the standard, the lowest performance group.

These two measures of student performance and growth are calculated using the percentage change in math, reading and writing by comparing the average results of recent AIMS data to the baseline years. Positive growth is considered to be an increase in the percentage of students who reach proficiency and a decrease in the percentage of students who fall far below the standard. The total percentage of growth then is calculated by adding together the percentages of change in each performance category. Thus, a decrease in the percentage of students in the FFB category is added to an increase in the percentage of students who are proficient.

With this methodology, it is important to note that a decrease in FFB is given as much emphasis as an increase in the percentage of students that meet the standard. In other words, because of a decrease in the percentage of students in the FFB category, a school may show positive growth even though there is no increase in the percentage of students meeting the standard on AIMS over time. The total growth

percentages are shown in Figure 5.¹⁹

Figure 5. Growth Percentages by Grade and Subject



In Figure 5, the combined increases and decreases in proficiencies are represented as either positive or negative growth. Negative growth represents a higher negative gain than positive gain. For example, in fifth-grade reading, Boulder Creek showed a decrease in the percent proficient coupled with an increase in the percentage of students in the FFB category. Both of these changes are considered negative change and are therefore expressed in terms of negative growth. Here, a higher negative growth value represents more negative growth. Likewise, a higher positive growth value represents more positive growth.

In Figure 5, Washington Elementary shows higher growth percentages in all subjects and grades. In third-grade math for example, Washington shows a +13.63% change and Boulder Creek shows a -8.43% change. Also, as seen in fifth-grade math results, while neither school shows a positive increase, Washington Elementary shows a smaller decrease: -8.02% compared with Boulder Creek’s -13.54%. Boulder Creek is showing negative changes in all subjects and grades but third-grade writing, while only half of Washington’s subjects and grades show negative change.

Awarding AIMS Baseline and Growth Points

After the baseline proficiencies and percentages of growth are calculated, points are awarded in each subject and grade. Points for baseline scores are awarded based on the percentage of students who are proficient in the baseline years. The higher the percentage, the more points awarded for the baseline scores. Likewise, growth points are awarded based on the percentage of growth shown over time; the higher the percentage of total growth, the more points awarded. There are six possible points in each of the categories: six points for the baseline proficiency and six points for the growth.²⁰

However, this methodology lends an inherent advantage to growth outcomes. If a school has a high baseline proficiency level, it does not have as much room to show improvement as a school with a much lower baseline proficiency level. To mitigate this advantage, the application of a 70/30 “float weight”²¹ is utilized: a 70% weight is given to the school’s strength (baseline or growth) and a 30% weight is assigned to the other category.

Table 1 provides a detailed summary of the baseline and growth points received by Washington Elementary and Boulder Creek in

each subject and grade. The 70/30 weight already has been applied. As shown below, Boulder Creek shows higher baseline points than Washington in both third and fifth grades. Since the baseline points are awarded based on average student performance in 2000 and 2001, this means that the students at Boulder Creek performed better on the AIMS test in those years. Washington Elementary, however, was awarded more growth points in both grades, an indication that there was greater improvement in test scores over time than at Boulder Creek.

As seen in Table 1, point values assigned for each school differ overall. Washington Elementary was awarded more than twice the points for growth as Boulder Creek: 11.8 compared with 5.3, respectively. On the other hand, Boulder Creek’s strength is in its baseline AIMS scores: 19.4 points, a value nearly four times the 5.2 baseline points that Washington Elementary was awarded. In this case, Washington Elementary is making much more positive growth than Boulder Creek in student achievement on the AIMS test. This is one of the key reasons why looking at one year of test data will not assist in understanding why a school received a Performing label; trends over time must be examined.

Table 1. AIMS Points Received by Subject and Grade

	Washington Elementary				Boulder Creek Elementary			
	Baseline Proficiency	Baseline	Growth	Total	Baseline Proficiency	Baseline	Growth	Total
Math	33.3%	0.6	2.8	3.4	76.6%	3.5	0.3	3.8
Reading	55.8%	0.6	2.1	2.7	95.1%	4.2	0.3	4.5
Writing	61.3%	0.6	4.2	4.8	89.0%	1.2	3.5	4.7
3rd Grade Totals		1.8	9.1	10.9		8.9	4.1	13
Math	17.6%	1.4	0.3	1.7	63.9%	3.5	0.3	3.8
Reading	44.2%	1.4	0.3	1.7	85.2%	3.5	0.3	3.8
Writing	34.5%	0.6	2.1	2.7	75.0%	3.5	0.6	4.1
5th Grade Totals		3.4	2.7	6.1		10.5	1.2	11.7
Grand Totals		5.2	11.8	17		19.4	5.3	24.7

Step 3. The Measure of Academic Progress Component

The Measure of Academic Progress (MAP) results provide general information about individual student performance on a national assessment from one year to the next. In Arizona, the administered norm-reference test was the SAT-9 test, which was administered during the spring semester each year, in grades 2 through 9 until 2005.²²

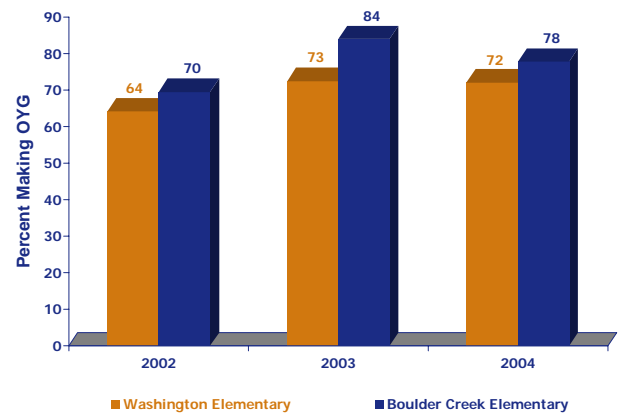
MAP currently is weighted to comprise up to 40% of a school's total possible points and is dependent on the percentage of students who showed one year's academic growth on the SAT-9 test. It is calculated based on a comparison of individual students' scores from one year to the next and is expressed as a percentage of students making what is considered to be the equivalent of one year of academic growth (OYG).²³

In general, if a student falls into the same relative range of performance from one grade to the next, that student is considered to have made OYG. For example, a student is considered to have made OYG if, in fourth grade, he falls within the same range of scores as he did in third grade.²⁴

Average MAP results for the most recent three years are illustrated in Figure 6. As shown, schools differ noticeably in the percentage of students who have demonstrated OYG. In

2002, 64% of students in Washington Elementary demonstrated OYG while 70% of students in Boulder Creek made OYG. In 2003, the difference in percentages is wider, 73% at Washington Elementary and 84% at Boulder Creek, an 11% difference.

Figure 6. Three Year Summary of MAP Percentages²⁵



Awarding MAP Points

The points for this component of the model are assigned based on the overall average of three years of MAP data. As shown in Table 2 below, the total numbers of students are added together in order to calculate a three-year average percentage of students attaining OYG. Washington Elementary shows an average percentage of 70% while Boulder Creek shows 77%. The overall points that are assigned to each school for this portion of the formula are based on these three-year average percentages.

Table 2. Three-year Average MAP Scores

	Washington Elementary			Boulder Creek Elementary		
	# Making OYG	# in Analysis	% Making OYG	# Making OYG	# in Analysis	% Making OYG
2002	415	646	64%	338	486	70%
2003	542	748	72%	473	563	84%
2004	566	785	72%	472	607	78%
Grand Totals	1,523	2,179	70%	1,283	1,656	77%

In this example, both schools received 7.5 points for the MAP portion of the calculation. This is because the average OYG over the course of three years was within the range of 70% to 79% (see Figure 7).²⁶

Figure 7. Point Outcomes for MAP²⁷

	Points Awarded	% OYG
	15.00	90 – 100
	11.25	80 – 89
7.5 Pts Awarded →	7.50	70 – 79
	3.75	60 – 69
	0	0 – 59

Step 4. The Adequate Yearly Progress Component

The Adequate Yearly Progress (AYP) component of Arizona’s accountability model is derived from a separate calculation developed by the state to comply with the federal accountability legislation, No Child Left Behind. This determination is complex and takes into account several indicators, including the AIMS assessment scores, percentage of students assessed and attendance rates. Although the details of this calculation need not be addressed, the impact on Arizona’s Achievement Profile model is important.²⁸ The AYP determination works as an “all or nothing” component, leading to a Pass or Fail for each school.

The AYP component of the model gives a school one point if AYP is passed and no points if it is not. Since both schools received one point for making AYP, it had no impact on the formula’s result.

Step 5. Overall Achievement Profile Designation

Schools are given points for each of the foregoing components. Then, the points are totaled and compared with a set of point ranges within each Achievement Profile. By

evaluating the AIMS baseline and growth groupings for all six subjects and grade combinations, the schools receive points for each subject and grade. These points are added together to make up a total AIMS point value. For example, as previously seen in Table 1, Washington Elementary received a total of 3.4 points and Boulder Creek received 3.8 points in third-grade math.

In addition to the AIMS points, MAP and AYP points also are added to arrive at a total point value for the entire school (see Table 3). In this case study, Washington Elementary received a total of 25.5 points and Boulder Creek received 33.2 when point outcomes for all the components were added together.

Table 3: Total Point Outcomes by Grade and Category

	Washington Elementary	Boulder Creek Elementary
Grade 3	10.9	13.0
Grade 5	6.1	11.7
MAP	7.5	7.5
AYP	1	1
% Exceeds	7.60%	26.80%
Total Points	25.5	33.2

Percent Exceeds Thresholds for Highly Performing and Excelling Labels

Additionally, there is an evaluation of the percentage of a school’s students who exceed the AIMS test standard over three years. For a school to move out of the Performing category and into the Highly Performing or the Excelling category, a school must show a required percentage of students not only meeting but exceeding the standard across all subjects and grades over three years.²⁹

After points are added for each school, the point total is compared with a range of values associated with the different labels (see Table 4). To be Performing, each school needed at least 24 points (23.5 due to rounding); to be

an Excelling school, each needed at least 32.4 points and the required percent of students exceeding the standard on AIMS. For these schools, a 38.80% was required in order to be an Excelling school.³⁰ Neither school showed the required percentage of students exceeding the standard.

Table 4. Point Requirements for a K-5 Elementary School³¹

Total Points	Label Designation	Required Percent of Students Exceeding the Standard on AIMS
<24	Underperforming	
24 to 27.5	Performing	
27.6 to 32.3	Highly Performing	30.90%
32.4 to 52	Excelling	38.80%

Despite the differences observed in AIMS points and overall points for each school, both Washington Elementary and Boulder Creek received a Performing Achievement Profile. As seen in Table 4, a K-5 school must receive a total of 24 points or more (23.5 due to rounding) to earn a Performing label. To receive a label of Highly Performing or Excelling, a K-5 school must receive more than 27.5 points and meet the required percentage of students exceeding the standard on the AIMS test over three years.

Washington Elementary is Performing because it received enough points to be Performing but not enough points to be placed in a higher category. Boulder Creek is Performing because it had enough points to be Excelling, but the percentage of students exceeding the standard on AIMS was not met. As seen in Figure 8 on Page 11, these two schools earned a Performing label by taking two very different paths.

Summary and Conclusion

The evaluation categories, or labels, relay certain messages to the public and to a school’s staff and student body about the quality of a

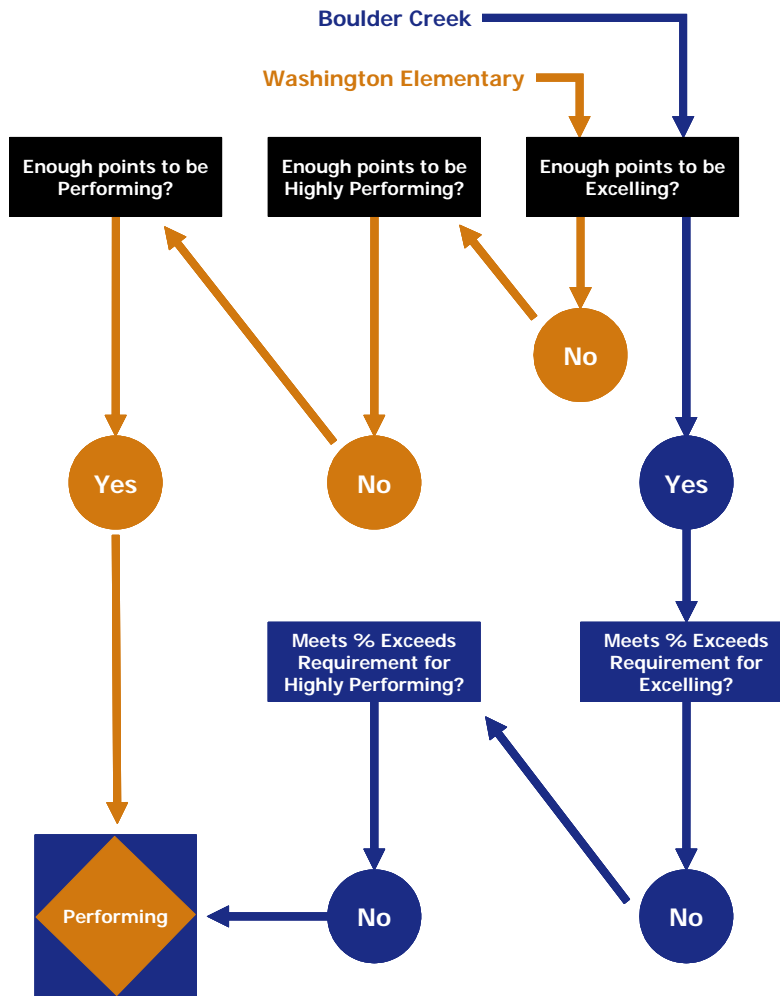
school. For example, if a school does not meet the required performance expectations, a label of Underperforming is given. This not only reveals a need for improvement, but impacts the public’s perception of a school and sends a message to the community about how well a school and its students are doing.

This case study compared two schools that received a Performing label. While the 2004 AIMS data appeared dramatically different, with one school seeming to significantly outperform the other, the formula treated the data in such a way as to classify both schools in the Performing category. There are two main reasons for this: 1) Washington Elementary, while showing lower AIMS scores than Boulder Creek, showed greater improvement over time; and 2) Boulder Creek, receiving enough points to be labeled Excelling, did not have enough students exceeding the standard on the AIMS test to have received a label other than Performing.

Applying a similar comparison across all Performing schools may reveal similar results. In 2004, nearly 70% of all public schools evaluated under this model were Performing,³² a distribution that encompasses schools with wide disparities in student performance, student body makeup and school type. Thus, a school’s label may not give a complete picture about how well a school is educating its students.

Examining labels over time is also an important exercise in determining how schools are evaluated from year to year.³³ Although this Policy Brief offers insight into what a Performing label means in 2004, it does not provide information on how Achievement Profiles can be compared across time. In the future, ThinkAZ will examine major changes that have occurred in the formula’s evolution and how those changes have impacted school labels during the formula’s three-year history.

Figure 8. Two Different Paths to a Performing Label



thinkAZ is an objective, non-partisan research institute dedicated to providing thorough and accurate information concerning key public issues in Arizona. ThinkAZ is a non-profit, Section 501(c)(3) corporation.

ENDNOTES:

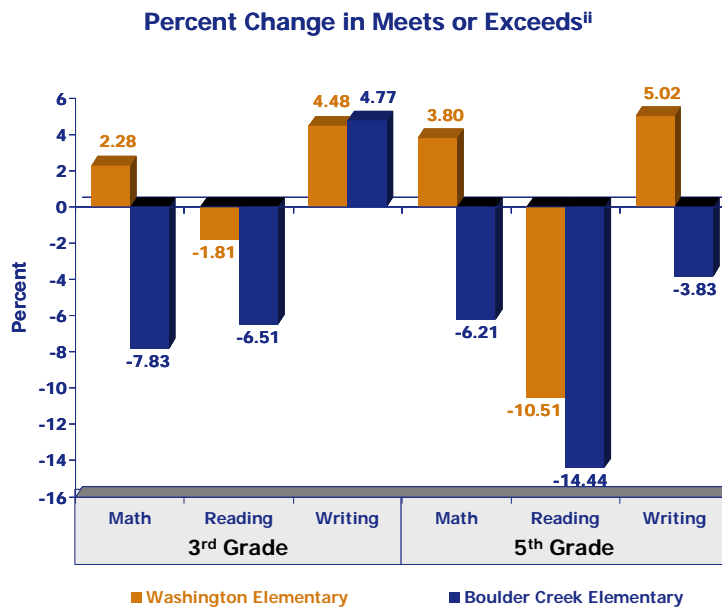
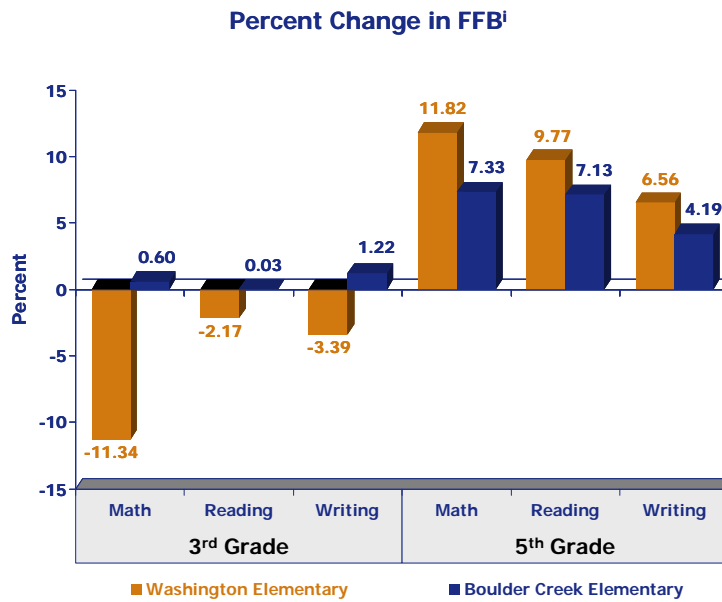
- ¹ Arizona Revised Statute A.R.S. § 15-241.G
- ² It is also important to understand that dramatic differences will not only be seen in the Performing category. Likewise, schools that appear to be very similar in terms of student performance can be awarded different Achievement Profiles.
- ³ Arizona Revised Statute A.R.S. § 15-241.A

-
- ⁴ In addition to the evaluation by the Arizona Department of Education, ARS § 15-241.J also mandates consequences for schools that do not meet performance objectives. If a school does not meet the expected level of performance, it must first develop an improvement plan within 90 days; then it must hold a public meeting in the school and provide written notification to all residents living in the attendance boundary. Failure to submit an improvement plan may result in withheld Classroom Site Fund monies generated by Proposition 301.
- ⁵ These requirements refer to the traditional K-8 Achievement Profile and are legislatively mandated in Arizona Revised Statute A.R.S. § 15-241.D, paragraphs 1 and 2.
- ⁶ Results from these tests are used in the calculation of the 2004 Achievement Profiles, which are discussed here. In the spring of 2005, Arizona administered a new test, which included AIMS and the Terra Nova.
- ⁷ Arizona Revised Statute A.R.S. § 15-241.N
- ⁸ A “subject and grade combination” refers to the fact that until 2005, the AIMS test was administered in grades 3, 5, 8 and in high school. Based on school configurations, some schools serve more AIMS grades than others. For example, a school that serves kindergarten through eighth grade has three AIMS grades (3, 5 and 8) within which three tests are administered (math, reading and writing). This nine subject and grade combination school has more subjects and grades than a K-5 school, which serves two AIMS grades and has six subject and grade combinations.
- ⁹ The AIMS data used in this section are publicly available from the Arizona Department of Education’s Web site at <http://ade.az.gov/researchpolicy/AIMSResults/> (accessed on November 29, 2004).
- ¹⁰ The state of Arizona publicly reports valid Category 1 and valid Category 2 students. Valid Category 1 students are students who took the test under normal circumstances and are, at the elementary level, considered to be proficient in English. This selection criteria excludes students who have not sufficiently acquired the English language, and it excludes students who are cognitively disabled and require special accommodations when testing. Arizona Department of Education, *Definition of Cat1and2 AIMS 02 04*. (2002), 1.
- ¹¹ Arizona has developed several different evaluations for its public schools. The variations have been developed as alternatives to evaluating schools with different sets of characteristics such as high schools, alternative schools and small schools. For example, there is a high school evaluation that takes into account dropout and graduation rates. For the purposes of the Brief, the analysis will remain focused on the traditional K-8 evaluation.
- ¹² Arizona Department of Education, “The ABCs of Arizona’s School Accountability,” 2004, 2, www.ade.az.gov/azlerans/puttingpieces/.
- ¹³ This assessment, called the Stanford-9, is a nationally normed test. A norm-referenced test rank orders student scores to provide information about how students’ results compare on a national scale.
- ¹⁴ The Stanford-9 assessment was administered in Arizona from 1997 through 2004. In the spring of 2005, Arizona adopted the Terra Nova, a different nationally norm-referenced assessment. Results from the Terra Nova are not discussed here.
- ¹⁵ As Adequate Yearly Progress AYP contributes a fixed one point value, the weight of this component will vary due to the number of subject and grade combinations that are evaluated in a school. For example, a one point value in a nine subject and grade combination school (K-8 configuration) will carry less weight than a high school with only three subject and grade combinations.
- ¹⁶ All data used in this section have been supplied by Research and Evaluation at the Arizona Department of Education and may not all be readily available for public consumption. Achievement Profile data received on November 29, 2004.
- ¹⁷ Arizona Revised Statute A.R.S. § 15-241.F states that the methodology for the Achievement Profile evaluation must “account for pupil mobility.”
- ¹⁸ Depending on the number of years a school has been in operation, the baseline either may be a percent proficient achieved by students in the 2001-2002 school year or an average percent proficient for the two earliest years, 1999-2000 and 2000-2001, or 2000-2001 and 2001-2002. In the case study provided in this Brief, both schools use an average of 2000 and 2001 school years for the baseline proficiency scores.
- ¹⁹ For the purposes of this Brief, the percentages already have been added together. To view a detailed breakdown of the improvement attained in each performance category, see Appendix 1.

-
- ²⁰ It is important to note that the points are assigned based on overall state averages in those subjects and grades. It is possible to receive points when showing a negative change in AIMS results if the statewide trend in that subject and grade is also negative.
- ²¹ In third-grade math, for example, the 70/30 weight would be applied in the following fashion: based on the school's baseline and growth scores, Washington Elementary was awarded 2 points for the baseline and 4 points for the growth. Rather than simply add these values together, a weight is applied such that 70% of the total points for this subject and grade combination is derived from the stronger area. In this particular case, the stronger area is the growth, as 4 is higher than 2. The 4 is weighted 70% and the 2 is weighted 30%. When these values are multiplied and added together, the equation is: $(4) \times (70\%) + (2) \times (30\%) = 2.8 + 0.6 = 3.4$. Washington Elementary has received 3.4 total points for third-grade math after the application of the 70/30 weight.
- ²² The SAT-9 assessment is a norm-referenced test that provides information about how, on a national scale, individual students compare with one another. Beginning in 2005, the state administered the Terra Nova as its norm-referenced assessment.
- ²³ It is important to note that, in many cases, not all students can be included in the analysis. Longitudinal matching of students is dependent on student stability and accurate coding so that individual students can be matched across years.
- ²⁴ The scores that are compared from one year to the next are called Stanine scores. Stanine scores range from 1 to 9 and represent a range of percentile ranks. Stanines are useful for comparing the relative range of a student's performance.
- ²⁵ These results are derived by combining both the math and reading subject areas.
- ²⁶ This MAP point scale is for a school with six subject and grade combinations. The point scale for MAP varies depending on the number of subject and grade combinations a school has.
- ²⁷ Arizona Department of Education, "Arizona's School Accountability System Technical Manual Volume I: Arizona LEARNs Achievement Profiles," 2003, 29, <http://ade.az.gov/azlearns/>.
- ²⁸ For more detailed information on how Adequate Yearly Progress status is determined, download Arizona's Accountability Workbook, available at <http://ade.az.gov/azlearns/workbook.asp>.
- ²⁹ While the required percentages may vary from school to school based on the grades offered, the two schools in this example have the same requirements.
- ³⁰ It is important to note that ARS 15-241 allows an appeal process through which schools may provide additional documentation of improvement, though the data examined may not show this success according to how the formula evaluates a school. A successful appeal may result in a different Achievement Profile. For example, if an Underperforming school provides documentation that results in a successful appeal, a Performing label may be granted. In this brief, neither of the two schools submitted an appeal.
- ³¹ Arizona Department of Education, "Arizona's School Accountability System Technical Manual Volume I: Arizona LEARNs Achievement Profiles," 2003, 36 and 42, <http://ade.az.gov/azlearns/>.
- ³² Achievement Profiles for all are schools are available at <http://ade.az.gov/profile/publicview/AZLEARNSSchoolList.asp?Year=2005>. Distribution information contained herein was accessed November 10, 2004.
- ³³ Three years of Achievement Profile determinations for each school can be found on a school's report card as published annually by the Arizona Department of Education. . School report cards can be accessed at www.ade.az.gov/srds/.

Appendix 1

Detailed Breakdown of Growth Percentages



ⁱ Positive change in the FFB category is expressed by a decrease in the percentage of students that fall far below the standard. Negative values represent positive growth.

ⁱⁱ Positive change in the Meets or Exceeds categories is expressed by an increase in the percentage of students meeting or exceeding the standard. Here, positive values represent positive growth.