
Arizona's School Accountability System *Technical Manual*



Tom Horne
Superintendent of Public Instruction
Arizona Department of Education

Volume I: AZ LEARNS Achievement Profiles

Research & Evaluation
Arizona Department of Education
Robert Franciosi, Deputy Associate Superintendent
Achieve@ade.az.gov

Published by the Arizona Department of Education, June 2005.

The Arizona Department of Education of the State of Arizona does not discriminate on the basis of race, religion, color, national origin, sex, disability or age in its programs, activities or in its hiring and employment practices. If you have questions or grievances related to this policy, please contact the Administrative Services DAS at (602)542-3186.

Table of Contents

INTRODUCTION.....	1
DEVELOPMENT OF THE ARIZONA LEARNS SYSTEM	5
WHAT’S NEW FOR THE 2003 AZ LEARNS ACHIEVEMENT PROFILES	5
STATE BOARD APPROVAL OF AZ LEARNS ACHIEVEMENT PROFILE	
METHODOLOGY	7
OVERVIEW OF THE AZ LEARNS ACHIEVEMENT PROFILE	8
GENERAL PROCESS TO PRODUCE THE ACHIEVEMENT PROFILES ..	9
THE ELEMENTARY SCHOOL MODEL (GRADES K-8).....	9
PRINCIPLES BEHIND THE USE OF TEST SCORES TO MEASURE SCHOOL PERFORMANCE	9
THE HIGH SCHOOL MODEL	9
DATA VERIFICATION AND TIMELINE	11
TIMELINE	11
CALCULATION OF AIMS BASELINE AND GROWTH SCALE POINTS	13
OVERVIEW	13
IDENTIFYING BASELINE GROUPS	13
IDENTIFYING GROWTH GROUPS	15
CALCULATION OF SUBJECT/GRADE SCALE POINTS FROM AIMS ..	19
SPECIAL CASES: NEW SCHOOLS, AND MISSING DATA	20
ADDED EVIDENCE: MEASURE OF ACADEMIC PROGRESS AND	
EXTENDED WRITING.....	21
OVERVIEW	21
MEASURE OF ACADEMIC PROGRESS	21
LONGITUDINAL MATCHING	21
MAKING ONE YEAR’S GROWTH	22
EXTENDED WRITING.....	22
CALCULATING ADDED EVIDENCE POINTS	23
GRADUATION AND DROPOUT RATES.....	25
OVERVIEW	25
GRADUATION RATES	25
DROPOUT RATES.....	27
EVALUATING THE TOTAL SCALE SCORE VALUE TO DETERMINE A	
SCHOOL CLASSIFICATION.....	30
APPLICATION OF THRESHOLD CRITERIA FOR EXCELLING AND HIGHLY	
PERFORMING SCHOOLS	33

AZ LEARNS ACHIEVEMENT PROFILE APPEALS PROCESS	36
STEP 1: DATA VERIFICATION/DATA APPEAL	36
STEP 2: APPEAL APPLICATION	36
STEP 3: APPEAL RESOLUTION	36
SMALL SCHOOL SCORE ADJUSTMENT.....	56
CRITERIA	56
METHOD	56
EXEMPTED SCHOOLS	57
OVERVIEW	57
EXTREMELY SMALL SCHOOLS	57
ALTERNATIVE SCHOOLS	57
ACCOMMODATION SCHOOLS	57
NEW SCHOOLS	57
K-2 SCHOOLS.....	57

List of Tables and Figures

Table 5.1. Baseline Groupings.....	14
Table 5.2 Number of Students	16
Table 5.3 Growth Point Groupings.....	18
Table 5.4 AIMS Scale- Point Distributions by Baseline Grouping and Growth Point Grouping	19
Figure 6.1: Summary of MAP Characteristics.....	21
Table 6.1. Determining One Year’s Growth.....	23
Table 6.2. Distribution of Elementary Added Evidence Points by Subject/Grade Combination.....	23
Table 6.3 Added Evidence Points.....	24
Figure 7.1. Graduation Rate Calculation	26
Figure 7.2. Calculating a Two Year Average for Graduation Rate	26
Figure 7.3. Calculating an Annual Dropout Rate	27
Figure 7.4. Calculating a Three-Year Average for Dropout Rate	28
Table 7.1. Target Improvements for Dropout and Graduation Rates in the High School Achievement Profile	29
Table 7.2. Point Value Outcomes	<u>34</u>
Table E8.1. Scaled Score Thresholds	<u>35</u>
Table E8.2. Scaled Score Thresholds	30
Table 8.1. Elementary School Classification Cut Points	30
Table 8.2. High School Classification Cut Points.....	31

Table 9.1. Excelling and Highly Performing Threshold Values by Grades Offered.....	34
Table 9.2. Number of Students Exceeding the Standard—All Grades.....	35
Table 11.1. An Extremely Small School ^a	43
Table 11.2. Number of Usable Third Grade Reading Tests for Two Schools.....	44
Table 11.3. Number of Students Passing AIMS, Third Grade Reading.....	45
Table 11.4. Rates for Determining UB95 for Small n and p.....	47
Table 11.5. AIMS Performance.....	48
Table 11.6. High School Classification/Non-MAP Elementary School Cut Points.....	48
Figure 12.1. Method for Evaluating Alternative Schools (358)...	52
Figure 12.2. Method for Alternative Schools (High Schools).....	52
Table 12.1. Alternative School Classification Cut Points	53
Table 13.1. AZ LEARNS Scale for K-2 Schools.....	55

1. Introduction

In November of 2001, Arizona voters approved Proposition 301 that, among other things, provided funds to the Arizona Department of Education (ADE) to develop “a system to measure school performance based on student achievement, including student performance on the AIMS test.” The legislative requirements for the accountability system are stated in section 15-241 (ARS § 15-241) of the Arizona Revised Statutes. The accountability system created to satisfy the statute is referred to as the AZ LEARNS. The school evaluation given by ADE to each school is referred to as the school’s achievement profile. In the fall of 2004, the ADE released the third year of AZ LEARNS school achievement profiles.

This manual describes the method used to generate the 2004 achievement profiles. It provides formulas, parameters, and business rules that make up the profile calculation. It also describes the AZ LEARNS process for 2004. Its intent is to document and explain the methods used and justify the policies adopted. This manual will give the ambitious user all the information required to calculate the achievement profile of his or her school.

What’s New for the 2004 AZ LEARNS Achievement Profiles

When developing modifications to the method of calculating school achievement profiles for 2004, the ADE, as in previous years, consulted with its Accountability Formula Working Group—a diverse group ranging from measurement experts, curriculum coordinators to classroom teachers. These experts volunteered their time to undertake the difficult task of advising the department on the complex issue of state-level school accountability.

For 2004, there were two major additions to the AZ LEARNS system. First, the ADE designed rubrics for calculating the achievement profiles of extremely small, alternative, and K-2 schools. These schools, totaling nearly 370, had not received profiles in the previous two years. Second, the ADE developed a protocol for reviewing appeals for schools being designated as underperforming for the third year in a row. State law requires that schools that underperforming for three consecutive years be given a designation of “failing to meet academic standards” or other designation. The ADE created an extended appeal process to determine if such schools merited the failing designation.

Other changes to the AZ LEARNS formula for 2004 were:

- **Change in Minimum Group Size.** The minimum size for a group to be included in the conventional achievement profile calculation was lowered from 30 to sixteen valid assessment scores. The minimum group size was increased

from 16 in 2002 to 30 in 2003 to make the AZ LEARNS formula consistent with the formula ADE uses to evaluate schools under No Child Left Behind (NCLB). However, upon further review, the change was deemed unnecessary. Since AZ LEARNS is a compensatory model, not a conjunctive model like the one used for NCLB, AZ LEARNS achievement profiles are less sensitive to changes in individual student scores. With a group size of fifteen the score of one student in one subgroup can cause an entire school to fail to make adequate yearly progress under the NCLB rubric. The score of one student out of fifteen would have less of an impact under AZ LEARNS.

- **Elimination of Extended Writing Scores from the Added Evidence.** In 2002 and 2003, schools were awarded points based on the percentage of students with a cumulative score of 24 or more on the extended writing portion of the AIMS test. Starting in 2004, the multiple choice sections of the writing test have been eliminated, and the writing score based solely on the students' writing samples. In order to avoid giving too much weight to students' writing performance, the AZ LEARNS formula no longer awards points for the extended writing score. AIMS writing scores are treated in AZ LEARNS exactly the same way as AIMS reading and math scores.

2. State Board Approval of AZ LEARNS Achievement Profile Methodology

The Arizona State Board of Education reviewed, commented upon, and approved changes in the AZ LEARNS achievement profile methodology, and the rubrics for evaluating extremely small, K-2, and alternative schools. The ADE provided the Board information packets that outlined the decisions regarding the formula that needed to be made. This documentation can be found online via the AZ LEARNS web site, <http://www.ade.az.gov/azlearns>.

In 2004 the following actions were taken by the Board with respect to AZ LEARNS:

- On January 26, 2004, the Board approved lowering the minimum group size (n-count) for a group to be included in the conventional achievement profile calculation.
- On August 30, 2004, the Board approved the rubrics for evaluating extremely small and alternative schools, and the opening of the application process for alternative schools. It also approved elimination of extended writing scores from the added evidence calculation. The Board also requested that the profiles for alternative schools be reported separately from other schools.
- On December 6, 2004 the Board approved the rubric for evaluating K-2 schools.

3. Overview of the AZ LEARNS Evaluation System

This section provides an overview of how AZ LEARNS achievement profiles are determined. More detailed discussions of the methodology used to determine the profiles, including descriptions of equations, algorithms, and data used are given in subsequent chapters.

Arizona law (ARS § 15-241) mandates that the Arizona Department of Education (ADE) shall compile an annual achievement profile for each public school. It specifies that the profiles of schools serving grades K-8 shall be based on:

- Arizona Measure of Academic Progress (MAP)
- Percent of students who pass AIMS.

It specifies that the profiles of high schools shall be based on:

- Drop out rate
- Graduation rate
- Percent of students who pass AIMS.

A school that serves both grades K-8 and high school receives two separate achievement profiles.

The law also calls for the ADE to use a research based methodology that shall:

- Include performance of pupils at all achievement levels
- Account for pupil mobility
- Account for the distribution of pupil achievement
- Include longitudinal indicators of academic performance.

A research based methodology is defined as “the systematic and objective application of statistical and quantitative research principles to determine a standard measurement of acceptable academic progress for each school”

The law also calls for a system of parallel achievement profiles for accommodation schools, alternative schools as defined by the Board of Education, and schools with a student count of fewer than one hundred pupils.

General Process to Calculate an Achievement Profile

The achievement profile for a school serving grades 3-8 consists of the following performance measures:

1. A baseline that measures the performance of students on all three sections of the AIMS (reading, writing, and mathematics) in given years.
2. A measure of improvement in aggregate student performance on the AIMS compared to the baseline years.
3. A measure of growth in individual student performance using the Stanford 9 (SAT-9) test. This is the Measure of Academic Progress (MAP).
4. A measure of whether the school made Adequate Yearly Progress (AYP) as defined by the No Child Left Behind Act of 2001. In order to comply with federal requirement that the state have an integrated accountability system, a school's AYP determination is factored into the calculation of its achievement profile.
5. In order to create the incentive for schools to improve the achievement of average and above-average students, a school cannot earn the highly performing or excelling labels unless over the previous three years the percentage of its students exceeding the standard on AIMS met specific thresholds.

The achievement profile for a high school is made up of all of the above components except for measure three. In addition, the following performance measures are used for high schools.

6. Dropout rate.
7. Graduation rate.

Schools are awarded scale score points based on their performance on measures one through four, six, and seven. Scale score points are then summed up for each school and compared to a scale that relates scale score points to the four profile labels: excelling, highly performing, performing, and underperforming. Performance measure five is then examined to determine if the school has earned the highest labels.

Principles Behind the Use of Test Scores to Measure School Performance

The Arizona Instrument to Measure Standards (AIMS) is a criterion-referenced test used by the state of Arizona to measure student performance in the areas of math, reading and

writing. In each subject area, students are grouped into performance categories based on how they performed relative to the state standard:

- Falls Far Below the Standard (FFB)
- Approaches the Standard (APP)
- Meets the Standard (Meet)
- Exceeds the Standard (EXC)

A student is deemed to have passed the test if he or she scores at meets or exceeds. The AIMS portion of the achievement profile carries an expectation that students will meet the state standards. For this reason, a year-to-year analysis of the percentage of students that fall far below the standard and meet or exceed the standard is used for determining a school classification.

The AIMS provides a measure of how well students have learned Arizona's academic standards. The MAP provides a longitudinal view of individual student progress across all grade levels. These indicators ensure that all grade levels in an elementary school share in the responsibility of moving all students forward.

Student groups vary from year to year and their performance will fluctuate from one year to the next, regardless of schooling effects. This phenomenon is known as a cohort effect. Both AIMS and MAP measures are based on a three-year rolling average to capture trends, rather than aberrations, and to account for differences in student cohorts. A three-year average mitigates these cohort effects and increases the reliability of the results. Furthermore, multiple years allows for the identification of real trends in school performance. A downward bump may be attributed to cohort effects, but a downward trend (continuous bumps) is an indication of lagging performance.

4. Data Verification and Timeline

Districts and charter holders are solely responsible for submitting the data necessary for calculating achievement profiles for their schools and for ensuring its accuracy. Because of the stakes involved and the volume and scope of the data used, the ADE considered it prudent to allow districts and charter holders to review their data before preliminary AZ LEARNS achievement profiles calculations were carried out.

From May 10, 2004, through July 7, 2004 an application to verify data was made available to districts and charter holders through the common login on the ADE web site. The primary purpose of the application was to allow districts and charter holders to correct the demographic information for individual students. In addition, a link was provided through the common login that allowed schools to download student-level testing data in order to make any necessary corrections. Schools also had the opportunity to correct data up through the close of the appeals window on October 3, 2004.

Timeline

The timeline for AZ LEARNS Achievement Profile process was:

- May 10, 2004. Opening of data verification process.
- July 7, 2004. Closing of data verification process.
- August 30, 2004. Opening of application process for alternative schools.
- September 10, 2004. Closing of application process for alternative schools.
- September 15, 2004. Preliminary release of AZ LEARNS achievement profiles for traditional schools (schools evaluated using the standard AZ LEARNS profile methodology). Opening of window for appeals submissions.
- October 4, 2004. Closing of appeals window for traditional schools.
- October 4, 2004. Preliminary release of AZ LEARNS achievement profiles for extremely small schools. Opening of window for appeals submissions.
- October 12, 2004. Preliminary release of AZ LEARNS achievement profiles for alternative schools. Opening of window for appeals submissions.
- October 15, 2004. Public release of AZ LEARNS achievement profiles for traditional schools.

- October 15, 2004. Closing of appeals window for extremely small schools.
- October 22, 2004. Closing of appeals window for alternative schools.
- November 3, 2004. Public release of AZ LEARNS achievement profiles for extremely small and alternative schools.
- December 10, 2004. Preliminary release of AZ LEARNS achievement profiles for K-2. Opening of window for appeals submissions.
- December 17, 2004. Closing of appeals window for K-2 schools.
- January 6, 2005. Public release of AZ LEARNS achievement profiles for K-2 schools.

5. Calculation of AIMS Baseline and Growth Scale Points

In the AZ LEARNS profile, schools are awarded scale score points based on student performance on the AIMS. Points are determined by performance in the baseline years, and improvement in student performance since then. The baseline years for most subjects, grades, and schools are the years 2000 and 2001 with the following exceptions:

1. For high school mathematics, the baseline year is 2001.
2. For schools that opened in 2001, the baseline years for all grades and subjects are 2001 and 2002.
3. For schools that opened in 2002, the baseline year for all grades and subjects is 2002.

The same criteria apply if school began offering a new grade. For example, if Gila Monster Elementary began offering fifth grade in 2002, 2002 would be the baseline year for that grade.

Identifying Baseline Groups

Equation Used to Calculate Baseline Scores

The following equation is used to calculate a baseline for every subject/grade combination offered by a school.

$$\text{Baseline Points} = \frac{\# \text{ Students passing AIMS in baseline years}}{\# \text{ Students tested in baseline years}}$$

The baseline is rounded to the nearest hundredth e.g. .675 = .68; .672 = .67.

Data Used

A student's score is excluded from the baseline calculation if any of the following criteria are met:

- Student received no score on the test.
- Student was not English proficient. A student was considered not proficient if he/she indicated on the AIMS test sheet that he/she was an English learner AND the student was classified as an English learner for three years or less. This condition applies only to grades three through eight. Students were included in

the calculation of achievement profiles for high schools regardless of English proficiency.

- The student received a non-standard accomodation on the test which changed the validity of the results.
- The student tested out of level.
- The student did not start the year at the school (Startyr = N).

The baseline calculation for high schools includes all students in grades ten through twelve who have taken the AIMS.

The baseline cut points used for the 2004 achievement profiles are the same used for the previous year. Please refer to the technical manual for the 2003 profiles (release date February 17, 2004) for a description of how the baseline cut points were established. The baseline groups for each subject/grade combination are given in table 5.1.

Table 5.1. Baseline Groupings

Grade	Subject	Baseline Grouping 1	Baseline Grouping 2	Baseline Grouping 3	Baseline Grouping 4	Baseline Grouping 5	Baseline Grouping 6
3	Math	0% - 26%	27% - 40%	41% - 56%	57% - 71%	72% - 82%	83% - 100%
3	Reading	0% - 46%	47% - 59%	60% - 73%	74% - 84%	85% - 91%	92% - 100%
3	Writing	0% - 54%	55% - 67%	68% - 79%	80% - 89%	90% - 94%	95% - 100%
5	Math	0% - 11%	12% - 21%	22% - 36%	37% - 52%	53% - 66%	67% - 100%
5	Reading	0% - 31%	32% - 44%	45% - 60%	61% - 75%	76% - 85%	86% - 100%
5	Writing	0% - 25%	26% - 38%	39% - 53%	54% - 68%	69% - 79%	80% - 100%
8	Math	0% - 1%	2% - 5%	6% - 12%	13% - 22%	23% - 34%	35% - 100%
8	Reading	0% - 25%	26% - 37%	38% - 51%	52% - 66%	67% - 77%	78% - 100%
8	Writing	0% - 18%	19% - 28%	29% - 42%	43% - 56%	57% - 68%	69% - 100%
H.S.	Math	0% - 3%	4% - 8%	9% - 19%	20% - 33%	34% - 47%	48% - 100%
H.S.	Reading	0% - 28%	29% - 42%	43% - 58%	59% - 73%	74% - 83%	84% - 100%
H.S.	Writing	0% - 16%	17% - 25%	26% - 39%	40% - 53%	54% - 66%	67% - 100%

A school is awarded baseline scale points for each subject/grade combination it offers that meets the minimum group size. The condition of a minimum group size of sixteen usable scores is applied to the baseline years. That is, a subject/grade combination must have an average of sixteen usable in the baseline years. Usable scores are student test scores not excluded by the above criteria. The number of points awarded for each subject grade is equal to that subject/grade's baseline grouping.

Example. In the baseline years, 66 percent of the students in Gila Monster Elementary passed math portion of the third grade AIMS. This value places the subject/grade in baseline grouping four. Gila Monster Elementary has earned four baseline scale points for this particular subject/grade.

Identifying Growth Groups

Determining a school's growth points for each subject/grade combination is based on student movement out of the falls far below (FFB) category and student movement into the meet/exceeds (M/E) category. The growth points earned are determined by the difference between average percentages of the most recent three years and the baseline percentages. The ADE considers a school to have made positive change if the three-year average percent of students that fall into the M/E category is higher than the baseline percentage or if the three-year average percent FFB is lower than the baseline. The ADE considers a school to have made negative change if the three-year average percentage of students in the M/E category is lower than the baseline or if the percentage of students in the FFB category is higher than the baseline percentage. The ADE expects schools to increase the percentage of students that meet the standards over time, and decrease the percentage that falls far below the standards over time.

Equations Used to Calculate Growth Points

Growth points for a subject/grade are calculated in the following seven steps.

Equation 1

$$\text{Baseline FFB} = \frac{\# \text{ Students FFB in baseline years}}{\# \text{ Tested in baseline years}}$$

Equation 2

$$\text{Baseline M/E} = \frac{\# \text{ Students M/E in baseline years}}{\# \text{ Tested in baseline years}}$$

Equation 3

$$3\text{-Yr. Avg. FFB} = \frac{\# \text{ Students FFB in most recent three years}}{\# \text{ Tested in most recent three years}}$$

Equation 4

$$3\text{-Yr. Avg. M/E} = \frac{\# \text{ Students M/E in most recent three years}}{\# \text{ Tested in most recent three years}}$$

Equation 5

$$\text{Change FFB} = 3\text{-Yr. Avg. FFB} - \text{Baseline FFB}$$

Equation 6

$$\text{Change M/E} = 3 - \text{Yr. Avg. M/E} - \text{Baseline M/E}$$

Equation 7

$$\text{Growth Points} = \text{Change M/E} - \text{Change FFB}$$

All values are rounded to the nearest ten-thousandth e.g. .67556 = .6756; .67221 = .6722.

Data Used

A student's score is excluded from the growth point calculation if any of the following criteria are met:

- Student received no score on the test.
- Student was not English proficient. A student was considered not proficient if he/she was an English learner AND the student was classified as an English learner for three years or less. This condition applies only to grades 3-8. Students were included in the calculation of achievement profiles for high schools regardless of English proficiency.
- The student received a nonstandard accommodation on the test.
- The student tested out of level.
- The student did not start the year at the school (Startyr = N).
- The growth calculation for high schools includes all students in grades ten through twelve who have taken the AIMS.

Example. The following example demonstrates how growth points are calculated. Table 5.2 shows four years of AIMS scores for a single subject and grade for a hypothetical school that has been open since 2000.

Table 5.2. Number of Students

Year	FFB	A	M	E	Total
2000	25	25	25	25	100
2001	20	30	25	25	100
2002	15	35	30	30	110
2003	10	35	30	35	110
2004	10	25	35	40	110

The following equations show the steps used to calculate the growth points given the test scores in the above table.

Equation 1

$$\text{Baseline FFB} = \frac{25 + 20}{100 + 100} = .2250$$

Equation 2

$$\text{Baseline M/E} = \frac{50 + 50}{100 + 100} = .5000$$

Equation 3

$$\text{3 - Yr. Avg. FFB} = \frac{15 + 10 + 10}{110 + 110 + 110} = .1061$$

Equation 4

$$\text{3 - Yr. Avg. M/E} = \frac{60 + 65 + 75}{110 + 110 + 110} = .6061$$

Equation 5

$$\text{Change FFB} = .1061 - .2250 = -.1189$$

Equation 6

$$\text{Change M/E} = .6061 - .5000 = .1061$$

Equation 7

$$\text{Growth Points} = .1061 - (-.1189) = .2250$$

The growth group cut points used for the 2004 achievement profiles are the same used for the previous year. Please refer to the technical manual for the 2003 profiles (release date February 17, 2004) for a description of how the cut points were established. The growth group cut points are given in table 5.3.

Table 5.3. Growth Point Groupings

Grade	Subject	Growth Point Grouping 1	Growth Point Grouping 2	Growth Point Grouping 3	Growth Point Grouping 4	Growth Point Grouping 5	Growth Point Grouping 6
3	Math	<-1.75%	-1.74% - 4.18%	4.19% - 10.12%	10.13% - 16.05%	16.06% - 21.98%	21.99% >
3	Reading	<-5.41%	-5.40% - -0.90%	-0.89% - 3.60%	3.61% - 8.11%	8.12% - 12.61%	12.62% >
3	Writing	<-9.23%	-9.22% - -5.00%	-4.99% - -0.77%	-0.76% - 3.46%	3.47% - 7.69%	7.70% >
5	Math	<-1.61%	-1.60% - 4.11%	4.12% - 9.83%	9.84% - 15.56%	15.57% - 21.28%	21.29% >
5	Reading	<-15.16%	-15.15% - -10.46%	-10.45% - -5.77%	-5.76% - -1.07%	-1.06% - 3.62%	3.63% >
5	Writing	<-8.18%	-8.17% - -3.44%	-3.43% - 1.29%	1.30% - 6.02%	6.03% - 10.76%	10.77% >
8	Math	<-7.99%	-7.98% - -1.94%	-1.93% - 4.11%	4.12% - 10.17%	10.18% - 16.22%	16.23% >
8	Reading	<-5.86%	-5.85% - -0.81%	-0.80% - 4.24%	4.25% - 9.29%	9.30% - 14.34%	14.35% >
8	Writing	<-10.24%	-10.23% - -5.92%	-5.91% - -1.61%	-1.60% - 2.70%	2.71% - 7.02%	7.03% >
H.S.	Math	<-5.81%	-5.80% - -1.60%	-1.59% - 2.61%	2.62% - 6.83%	6.84% - 11.04%	11.05% >
H.S.	Reading	<-10.50%	-10.49% - -6.10%	-6.09% - -1.70%	-1.69% - 2.71%	2.72% - 7.11%	7.12% >
H.S.	Writing	<10.72%	10.73% - 15.32%	15.33% - 19.92%	19.93% - 24.52%	24.53% - 29.12%	29.13% >

A school is awarded growth scale points for each subject/grade combination it offers that meets the minimum group size. The condition of a minimum group size of sixteen usable scores is applied only to the baseline years. That is, a subject/grade combination must have an average of sixteen usable in the baseline years. The minimum group size condition is not applied to subsequent years. Usable scores are student test scores not excluded by the above criteria. The number of points awarded for each subject grade is equal to that subject/grade's growth grouping.

Example. Gila Monster Elementary has obtained 5.15 growth points in third grade mathematics. This is equivalent to growth grouping three. Gila Monster Elementary has earned three growth scale points in third grade math.

Calculation of Subject/Grade Scale Points from AIMS

The total scale score points derived from AIMS performance are calculated by adding the baseline points awarded to the growth points awarded. A 70 percent weight is given to the school's strongest scale value (baseline group or growth point group) and a 30 percent weight to the other scale value. Table 5.4 shows the scale points earned per subject/grade for all combinations of baseline and growth group scale points.

Example. In third grade mathematics, Gila Monster Elementary has earned four baseline group scale points and three growth groups scale points. Because it has earned more scale points for its baseline group, the baseline scale points are given a 70 percent weight and the growth group scale points are given a 30 percent weight. Thus, the total scale points earned for third grade math are $(.7 \times 4) + (.3 \times 3) = 3.7$.

Example. In mathematics, Desert Mountain Vista High School has earned two baseline group scale points and five growth group scale points. Because it has earned more scale points for its growth group, the growth scale points are given a 70 percent weight and the baseline group scale points are given a 30 percent weight. Thus, the total scale points earned for third grade math are $(.3 \times 2) + (.7 \times 5) = 4.1$.

The total scale score points derived from AIMS for all subject/grade combinations for a school are summed and added to the scale score values for other performance measures.

Table 5.4. AIMS Scale- Point Distributions by Baseline Group and Growth Point Group

	Growth Point Group 1	Growth Point Group 2	Growth Point Group 3	Growth Point Group 4	Growth Point Group 5	Growth Point Group 6
Baseline Group 1	1.0	1.7	2.4	3.1	3.8	4.5
Baseline Group 2	1.7	2.0	2.7	3.4	4.1	4.8
Baseline Group 3	2.4	2.7	3.0	3.7	4.4	5.1
Baseline Group 4	3.1	3.4	3.7	4.0	4.7	5.4
Baseline Group 5	3.8	4.1	4.4	4.7	5.0	5.7
Baseline Group 6	4.5	4.8	5.1	5.4	5.7	6.0

Special Cases: New Schools, and Missing Data

If a school is missing AIMS test data for a subject/grade combination it offers, it receives zero scale points for that subject/grade for the achievement profile calculation for the test year in which the data are missing. In the subsequent years, the following rules approved by the State Board of Education are used for calculating baseline and growth scale points.

1. If a school has one year of missing data for the baseline years, its baseline is calculated using the single year of baseline data it does have.
2. If a school has two years of missing data for the baseline years it receives zero baseline scale points. Growth is calculated from a base of zero, but growth scale points only receive 30 percent weight.

Example. Gila Monster Elementary is missing data for fifth grade writing for both 2000 and 2001. Its growth points for that subject/grade combination put it in growth group five. Thus it receives $5 \times 0.3 = 1.5$ AIMS scale points for fifth grade writing.

1. If a school is missing one year of data used in the three-year rolling average to calculate growth points, its growth points are calculated using two years of data.
2. If a school is missing two years of data used in the three-year rolling average to calculate growth points, it receives 0 growth scale points, and its baseline scale points are given a weight of 30 percent.

Example. Gila Monster Elementary is missing data for fifth grade reading for both 2002 and 2003. Its baseline for that subject/grade combination put it in baseline group two. Thus it receives $2 \times 0.3 = 0.6$ AIMS scale points for fifth grade writing.

New Schools

Once a school has been operational for three test administrations, the school receives an achievement profile utilizing the updated formula changes passed by the Board during 2003, with the exception of a one-year baseline analysis rather than a two-year baseline analysis. The new school is evaluated using three years of data; a school must only have one overlapping year of test data to be evaluated for an achievement profile.

Example. Desert Mountain Vista High School opened in 2002 and is still currently in operation. Its achievement profile would be calculated with the following methodology:

1. Baseline grouping AIMS data (single year) - 2002
2. Growth point group AIMS data (three year average) - 2002, 2003, 2004.

6. Added Evidence: Measure of Academic Progress (MAP)

Overview

The Added Evidence portion of an elementary profile relies on the Arizona Measure of Academic Progress (MAP), which uses longitudinal Stanford 9 test scores to provide a measure of student academic growth.

Measure of Academic Progress

Unlike traditional measures of achievement, such as percentile ranks that mark achievement at one point in time, MAP measures growth over time. A measure of the progress made over a school year is obtained through linking individual student test scores from one year to the next. Traditional indices of achievement, such as a comparison of schools' percentile ranks calculated at a point in time, are highly correlated to student demographic variables. As a result, the same schools consistently score at the top and bottom of the percentile rank listings. With MAP as a measure of school effectiveness, schools traditionally seen as low performing, by way of a percentile rank, may show remarkable gain with the students they have had an opportunity to teach.

Figure 6.1: Summary of MAP Characteristics

MAP Key Features:
• Utilizes Stanford 9 stanine scores • Captures individual student growth over time • Accounts for mobility • Includes only those students a school has had an opportunity to teach • Focuses on <i>all</i> students

Longitudinal Matching

In order to determine the number of students that made One Year's growth (OYG), individual student records are matched from test year to test year. The process of matching Stanford 9 student records between test years involves the use of four pieces of student level data. Perfect student matches were made if all of the following are true:

- The student attended the same school during two years of consecutive testing
- The student had the same first and last name
- The student's gender was the same for both years
- And the birth date of the student was identical for both years.

However, due to inconsistencies in some of the student information (e.g. if a student misspelled his or her name, or if a mistake was made on any other piece of necessary information) matching student records that are not picked up in a perfect match involved a multi-stage process. For example, if a student spelled his name "Frank Adams" in 2001 and misspelled it as "Grank Adams" in 2002, Frank's records would not be matched in the initial pass. If all other pertinent pieces of information were identical, the second phase of the matching process would in all likelihood link his records. In addition to letter changes or letter transposition, the matching routine could also handle transposed birth dates. For example, if Frank Adams had also indicated that he was born 09/10/90 in 2001 and 90/10/09 in 2002 his record would likely be linked. Typically, close to 90 percent of student test scores can be matched between any two test years.

The multistage approach to matching has limitations, however. If the student's first or last name varied by more than two letters from one year to the next, the records cannot be matched. Also, students that do not provide their birth dates or indicate their gender in one of the years cannot be matched. In rare instances, false matches can occur between two different students in different test years. This is most likely to occur when students are related, attending the same schools in successive years and have similar first names, identical last names and are a year apart in age. Given the high proportion of test scores that are appropriately matched, the above stated limitations are of minor concern.

Making One Year's Growth

MAP is used to measure individual student growth. Student Stanford 9 test scores are linked from one year to the next and growth on the test is calculated. One Year's growth (OYG) is broadly defined as attaining the same level of relative achievement from year to year, while learning more difficult material. OYG is determined by examining a student's stanine score: if one attains the same stanine score or a higher stanine score relative to the previous year, that student has made one year's growth. There are two exceptions to this approach (see table 6.1):

- Students who begin in stanine nine and move to stanine eight will make OYG.
- Those that stay in stanine one from one year to the next will not make OYG.

For example, a student who begins at the fifth stanine (50th percentile) as a third grader and maintains a fifth stanine score as a fourth grader has achieved OYG. This is the minimum growth that is expected for any student who remains at a school for an

academic year. When all students achieve OYG, schools ensure that no students are falling behind from one school year to the next.

Table 6.1. Determining One Year's Growth

Stanine 2001	Stanine 2002	Result
5	5	OYG
6	7	OYG
9	8	OYG
5	4	Did not Accomplish OYG
1	1	Did not Accomplish OYG

Calculating Added Evidence Points

Added evidence points are calculated by determining the percentage of students making one year's growth for the most recent three years for the whole school (reading and mathematics). The percentage of students making one year's growth is compared to the number of subject grade combinations in the school to determine the number of points earned.

Table 6.2. Distribution of Elementary Added Evidence Points by Subject/Grade Combination

% Total Added Evidence	<u>Subject/Grade Combinations</u>								
	1	2	3	4	5	6	7	8	9
90% +	3	5	8	10	12	15	17	20	22
80%-89%	2.25	3.75	6	7.5	9	11.25	12.75	15	16.5
70%-79%	1.5	2.5	4	5	6	7.5	8.5	10	11
60%-69%	0.75	1.25	2	2.5	3	3.75	4.25	5	5.5

Example. The table below shows hypothetical data used to determine added evidence points for the 2003 achievement profiles.

Table 6.3. Added Evidence Points

Year	MAP Math # Making OYG	MAP Read # Making OYG	Number in MAP Analysis
2002	100	105	300
2003	110	100	320
2004	105	100	310
Total	315	305	930

The formula for calculating the percent total added evidence (PTAE) is:

$$\text{PTAE} = \frac{(\text{Total students making OYG reading and math})}{(\text{Total students in MAP analysis})}$$

For the numbers in the above table, the percent total added evidence would be:

$$\text{PTAE} = \frac{315 + 305}{930} = .67.$$

Assuming the school is has a grade configuration of K-6 and meets the minimum n-count for evaluation for all subjects and grades, it would have six total subject/grade combinations. Reading from table 6.2, 67 percent total added evidence and six subject/grade combinations would earn the school 3.75 added evidence points.

7. Graduation and Dropout Rates

Overview

The graduation and dropout rates are important complements to the high school model used in the revised achievement profiles. Graduation rates indicate the success of students in meeting course requirements and achieving passing grades in subject areas not covered by the AIMS test. Dropout rates are included in the high school model as a measure of student persistence and to ensure that AIMS results reflect the largest percentage of the student population possible.

Graduation and dropout rates are used solely in the calculation of a high school profile. High school status was granted to any school that reported data in grade ten for each of the relevant school years.

In the calculation of a high school achievement profile, the two rates operate independently of one another; in other words, the point value outcome is a reflection of how the school performs in each of the categories. When one of the pieces of data is missing, a school cannot meet the target for that category.

The rounding of dropout and graduation rate percentages worked in such a way that the decimal was only taken out to four places, regardless of the value of the fifth. In other words, the percentages, expressed at this point as decimals, were divided by 1000, then rounded based on one decimal place, such that 0.5 percent was rounded up and 0.4 percent was rounded down. Lastly, the number was divided by ten to produce a percentage and one decimal place.

For example, if, in 2001, a school reported a total enrollment of 90 and a dropout count of 25, this translates into a baseline dropout rate of 0.27777778. To round this into a percent, the calculation divided this number by 1000, which equals 277.7, then rounded to 278, then divided by ten: 27.8 percent. The final percent is compared to the cutoff points in each category.

Graduation Rates

The Graduation Rate is a five-year, longitudinal measure of how many students graduate from high school. By examining a cohort of students who began high school at the same time, the graduation rate assesses how many students actually complete high school within a five-year period (see figure 7.1).

Figure 7.1. Graduation Rate Calculation

Graduation Rate	=	$\frac{\text{Number of Cohort members who graduated after five years}}{\text{Original Cohort Membership} + \text{Transfers In} - \text{Transfers Out} - \text{Deceased}}$	X 100
--------------------	---	---	-------

For graduation rate data, four years of data were used:

Graduation rate for the cohort class of 2000, which represents the baseline rate and is used as a reference point in order to measure increases from year to year

-and-

A three-year average of the cohort classes of 2001, 2002 and 2003. The three-year average is calculated by taking the total number of combined five year graduates and dividing by the total number of students in the combined cohort classes (see figure 7.2).

Figure 7.2. Calculating a Three Year Average for Graduation Rate

2 Year Average	=	$\frac{2001 + 2002 + 2003 \text{ 5 Year Graduates}}{\text{Combined number of students in three cohorts}}$	X 100
-------------------	---	---	-------

Cut points for the graduation rate portion of the profile were determined by examining a distribution of school graduation rates in Arizona and comparing them to the state mean. After analyzing the state distribution of scores and the guidelines in the achievement profile legislation, the following rates were targeted as cut point values for graduation rate:

73.5 percent: The state mean when school size is controlled for by averaging the rates of all schools in Arizona.

89.5 percent: As stated in A.R.S.§15-241, the highest cut point for a 5-year graduation rate is at least 90 percent.

In order for a school to meet the target for their 5-year graduation rate achievements, incremental gains must be made from one year to the next. These gains are evaluated by comparing the three-year average rate to the baseline rate (see table 7.1). A school can meet the target for graduation rate in multiple ways, depending upon the value of its three-year average rate:

- If the three-year average is 89.5 percent or greater, the target is automatically met.

- If the three-year average is less than 89.5 percent, the baseline rate is used as a reference point and incremental gains must be made from year to year. The gains required to meet the target were derived from the statewide distribution of graduation rates.
- If the baseline rate is greater than or equal to 73.5 percent, the difference between the three-year average and the baseline rate must be greater than or equal to 0.5 percent.
- If the baseline rate is less than 73.5 percent, the difference between the three-year average and the baseline rate must be greater than or equal to 1.5 percent.

Dropout Rates

The Dropout Rate is an annual measure of how many students drop out of a school during a twelve-month reporting period. It is expressed as the proportion of students who dropped out during the year to the total number of students that enrolled in the school over the course of the year (see figure 7.3). The state of Arizona defines a dropout as a student who was enrolled in a school at any point during the year, was not enrolled at the end of the year and did **not**:

Graduate or complete high school

-or-

Transfer to another qualified educational facility

-or-

Die

Figure 7.3. Calculating an Annual Dropout Rate

$\text{Dropout Rate} = \frac{\text{Number of students who dropped out}}{\text{Number of students who were enrolled during the school year} \times 100}$

For dropout rate data, two values were used:

- Dropout rate for the 2001 school year, which represents the baseline rate and was used as a reference point in order to measure increases from year to year

-and-

- A three-year average of the 2002, 2003, and 2004 school years. The three-year average is calculated by taking the total number of combined dropouts and dividing by the total number of combined students served (see figure 7.4).

Figure 7.4. Calculating a Three-Year Average for Dropout Rate

$$\text{3 Year Average} = \frac{\text{Total number of students who dropped out in 3 years}}{\text{Total number of students who were enrolled during 3 years}} \times 100$$

Cut points for the drop out rate portion of the profile were determined by looking at a distribution of school dropout rates in Arizona and comparing them to the state mean. After analyzing the state distribution of scores and the guidelines in the legislation, the following rates were targeted as cut point values for dropout rate:

- **9.4 percent:** The state mean when school size is controlled for by averaging the rates of all schools in Arizona.
- **6.0 percent:** As stated in A.R.S.§15-241, the lowest cut point for dropout rate is less than 6 percent.

In order for a school to meet the target for their annual dropout rate achievements, incremental decreases must be made from one year to the next. These decreases are evaluated by comparing the three-year average rate to the baseline rate (see table 7.1). A school can meet the target for dropout rate in multiple ways, depending upon what the three-year average rate is:

- If the three year average for the annual dropout rate is 6.0 percent or less, the target is automatically met.
- If the three year average is greater than 6.0 percent, the baseline rate is used as a reference point and incremental gains must be made from year to year. The gains required to meet the target were derived from the statewide distribution of dropout rates.
- If the baseline rate is less than or equal to 9.4 percent, the difference between the three year average and the baseline rate must be greater than or equal to 0.5 percent.
- If the baseline rate is greater than 9.4 percent, the difference between the three-year average and the baseline rate must be greater than or equal to 1.5 percent.

Table 7.1. Target Improvements for Dropout and Graduation Rates in the High School Achievement Profile

Baseline Dropout Rate*	Target**	Baseline Graduation Rate*	Target**
< or = 9.4%	0.5% Decrease	> or = 73.5%	0.5% Increase
> 9.4%	1.5% Decrease	< 73.5%	1.5% Increase

*Recall the baseline rates are taken from the 2000-2001 academic year

**Meeting the target is met assessed by calculating the difference between the average rate and the baseline rate. Recall that, for the annual dropout rate, a three-year average is used and for the graduation rate, a two-year average is used.

A total of two points is awarded for the combination of dropout and graduation rate portion of an achievement profile. If both targets are met, two points are awarded; if one of the targets is met, one point is given. If neither of the targets is met, no points are awarded. The target could not be met if data were missing in a category.

Table 7.2. Point Value Outcomes

	Point Value
Graduation and Dropout Targets Met	2
Graduation or Dropout Target Met	1
Neither Graduation or Dropout Target Met	0

8. Evaluating the Total Scale Score Value to Determine a School Classification

The tables below show the total number of scale score points that schools must earn in order to receive a given classification. A school may receive up to six AIMS scale points for each of its subject/grade combinations that are evaluated; up to 22 added evidence scale points if it is an elementary or middle school; and up to two graduation/drop out scale points if it is a high school. A school receives one scale point if it made adequate yearly progress (AYP).

For each school, the applicable scale score cut points for classification determinations are calculated by multiplying the number of subject/grade combinations evaluated for AIMS scale score points at that school by the base classification cut points. The base classification cut points for elementary schools are given in the first column of table 8.1 and those for high schools are given in the first column of table 8.2.

Table 8.1. Elementary School Classification Cut Points

	Subject/Grade Combination 1	Subject/Grade Combination 3	Subject/Grade Combination 6	Subject/Grade Combination 9
Underperforming	< 4	< 12	< 24	< 36
Performing	4	12	24	36
Highly Performing	4.6	13.8	27.6	41.4
Excelling	5.4	16.2	32.4	48.6

Table 8.2. High School Classification Cut Points

	Subject/Grade Combination 1	Subject/Grade Combination 3
Underperforming	< 3.2	< 9.6
Performing	3.2	9.6
Highly Performing	5	15
Excelling	5.4	16.2

Example. Gila Monster Elementary, a K-6 school, has six subject/grade combinations (reading, math, and writing for grades three and five) evaluated for AIMS scale score points. Thus, the scale score thresholds for Gila Monster Elementary are:

Table 8.3. Scale Score Thresholds

	#Subject/Grade Combinations	Base Cut Points	Scale Points
Underperforming			<24.0
Performing	6	4.0	=24.0
Highly Performing	6	4.6	=27.6
Excelling	6	5.4	=32.4

Example. Kangaroo Rat Elementary, a K-6, school only has five subject/grade combinations evaluated. One subject/grade combination failed to meet the minimum sixteen valid score count for the baseline years. Thus, the scale score thresholds for Kangaroo Rat Elementary are:

Table 8.4. Scale Score Thresholds

	#Subject/ Grade Combinations	Base Cut Points	Scale Points
Underperforming			<20.0
Performing	5	4.0	=20.0
Highly Performing	5	4.6	=23.0
Excelling	5	5.4	=27.0

9. Application of Threshold Criteria for Excelling and Highly Performing Schools

To ensure continued focus on improving the academic achievement of all students as they reach their absolute levels of attainment, including those students currently demonstrating proficiency in Arizona's academic standards on AIMS, threshold criteria are applied to determine excelling and highly performing schools. These threshold criteria are based on the three-year average percentage of students in the exceeds-the-standard category on AIMS (reading, writing, and mathematics) in a particular school. Conceptually, these threshold criteria serve as parameters to establish distinct boundaries around the excelling and highly performing achievement profile classifications. Schools must not only receive a total scale value that places them into either excelling or highly performing, but must also meet the requisite percentage of students in the exceeds-the-standard category on AIMS to be designated as either an excelling or highly performing schools. The application of threshold criteria for excelling and highly performing schools results in the following scenarios:

1. A school that receives a total scale value that places it in the excelling classification and meets the requisite percentage of students in the exceeds category on AIMS necessary for an excelling classification will be designated an excelling school.
2. A school that receives a total scale value that places it in the excelling classification and does not meet the requisite percentage of students in the exceeds category on AIMS necessary for a excelling classification, but did met the requisite percentage of students in the exceeds category on AIMS necessary for a highly performing classification will be designated as a highly performing school.
3. A school that receives a total scale value that places it in the excelling classification and does not meet either the requisite percentage of students in the exceeds category on AIMS necessary for the excelling classification or the highly performing classification will be designated as a performing school.
4. A school that receives a total scale value that places it in the highly performing classification and meets the requisite percentage of students in the exceeds category on AIMS necessary for an excelling classification will be designated as a highly performing school.
5. A school that receives a total scale value that places it in the highly performing classification and meets the requisite percentage of students in the exceeds category on AIMS necessary for a highly performing classification will be designated as a highly performing school.

6. A school that receives a total scale value that places it in the highly performing classification and does not meet either the requisite percentage of students in the exceeds category on AIMS necessary for an excelling classification or highly performing classification will be designated a performing school.
7. A school that receives a total scale value that places it in the performing classification will be designated as a performing school, regardless if the school meets the requisite percentage of students in the exceeds category on AIMS necessary for an excelling classification or highly performing classification.
8. A school that receives a total scale value that places it in the underperforming classification will be designated as an underperforming school, regardless if the school meets the requisite percentage of students in the exceeds category on AIMS necessary for an excelling classification or a highly performing classification.

Utilizing a three-year average of the percentage of students in the exceeds-the-standard category on AIMS the ADE set the thresholds for excelling and highly performing achievement assessed at a particular school. In order to establish thresholds for excelling and highly performing schools, schools were rank ordered by the three-year average percentage of students exceeding the standard on AIMS. Then, the threshold for highly performing was set at the 75th percentile rank of schools with students exceeding the standard; the threshold for excelling was set at the 90th percentile rank of schools with students exceeding the standard. (Please see table 9.1).

Table 9.1. Excelling and Highly Performing Threshold Values by Grades Offered

Subject Grade Combinations	School Type (Serving grades)	Highly Performing	Excelling
3	3 or 5	22.6%	28.7%
3	8	6.5%	10.7%
6	3 and 5	30.9%	38.8%
6	5 and 8	19.9%	31.7%
9	3, 5 and 8	19.3%	25.7%
3	High School	9.3%	12.7%

Example. The following table shows distribution of AIMS scores for Gila Monster Elementary. The numbers are for all grades in Gila Monster Elementary for which the AIMS is administered.

Table 9.2. Number of Students Exceeding the Standard—All Grades

Subject	Reading		Mathematics		Writing	
Year	# Exceeding	#Tested	# Exceeding	#Tested	# Exceeding	#Tested
2002	25	100	24	100	23	100
2003	24	105	23	105	22	105
2004	26	99	25	99	24	99
Total	75	304	72	304	69	304

The percent of students exceeding the standard is then:

$$\text{Percent Exceeding} = \frac{75 + 72 + 69}{304 + 304 + 304} = 23.6\%.$$

Since Gila Monster serves both grades 3 and 5, it must meet thresholds of 30.9 percent of students exceeding the standard to be classified as highly performing, and 38.8 percent to be classified as excelling. Gila Monster Elementary meets neither threshold.

10. AZ LEARNS Achievement Profile Appeals Process

Procedure and Timeline

In accordance with A.R.S. §15-241, school administrators are allowed the opportunity to appeal an achievement profile classification on behalf of the schools for which they are responsible.

Step 1: Data Verification/Data Appeal

The first step in completing the AZ LEARNS appeals process requires all schools to review and verify all data in order to confirm its accuracy. The data verification took place utilizing the AZ LEARNS/Adequate Yearly Progress (NCLB) Application through the common logon located at the ADE's web site. Data verification took place May 10 through July 7, 2004. In the application, schools/districts were asked to verify:

- **SAIS/Student Details Demographic Data**, which included student's full name, grade enrolled, SAIS number, date of birth, gender, racial background/ethnic group, English language learner Status (Y/N), special education program membership (Y/N), student's school ID.
- **Pre-printed test label information collected**, which included start year (Y/N), number of years in school, number of years in district, 504 Accommodation status, level of English proficiency (ELL or FEP), number of years classified as ELL or FEP, ELL program enrollment.
- **School information** used to determine if a school was to be evaluated for AZ LEARNS.

It is important to note that districts and charter holders are solely responsible for verifying information for their schools. If a district or charter holder does not verify the information for its schools, the ADE assumes the schools on file and the data available are correct as listed.

Step 2: Appeal Application

School administrators that chose to appeal an achievement profile must have completed the appeal application, which is accessible via the common logon during the specified appeal timeframe in order to indicate the exact issues of appeal. Appeals are only accepted through the website application. Appeals sent to ADE via email, fax, or mail/delivery are not accepted.

Schools are able to appeal achievement profile in two categories: data (statistical) and non-data (substantive) reasons - schools are not limited to one category and are able to appeal in both if necessary.

Schools that appeal based on statistical arguments may argue that their data are inaccurate. This may include (but is not limited to) data that are missing, miscoded, or invalid. Any of the data verified, as listed above, may be included in a statistical argument. Schools that appeal based on substantive arguments may argue that mitigating factors, outside of the school's control, negatively affected the quantity or quality of test data in any of the years in which data were collected. This includes circumstances that affected test conditions, test scores, and performance levels.

School administrators that chose to appeal an achievement profile must clearly articulate the issues they believe merit an appeal through the appeal application. School administrators must submit evidence that the issues they believe merit an appeal directly resulted in a *significant* decrease in student academic achievement as demonstrated on AIMS or SAT-9 and/or an impact on other indicators used in the formula (Adequate Yearly Progress (AYP); the Measure of Academic Progress (MAP)-for elementary schools; or graduation/dropout rate-for high schools). The evidence must be submitted to ADE at the time the appeal is submitted. Failure to provide this evidence results in the appeal not being granted. Evidence that is submitted after the appeal deadline closes is not considered. Once appeals are submitted through the common login, the school/district/charter receives an email verifying that the appeal has been received.

The ADE, if necessary, may request that a school administrator provide additional information or evidence to assist in the appeals process. Only those requests for additional information that are provided during the specific timeframe are included in the appeals process. Requests submitted after the specified timeframe were excluded from the appeals process. Unsolicited additional information submitted after the appeal deadline is not accepted.

Step 3: Appeal Resolution

After all appeals are submitted and the appeal window closes, the ADE processes the appeals. Appeals are addressed categorically, not necessarily in the order received so the fact that a school submitted its appeal during the first day of the appeal window does not mean it will necessarily receive a decision first during the resolution process. The appeal process is implemented in four stages.

Stage 1 – Statistical Appeals Processed. All statistical appeals need to be supported with compelling evidence. For example, if the school is disputing the number of test scores used in the analysis because some scores were excluded due to coding errors, the school must be explicit as to how the number in the analysis is not accurate. Simply stating, “number of students ADE used in the analysis doesn’t match district’s count” is not compelling; ADE needs to know *why* the numbers are different; meaning that particular students need to be identified as miscoded or missing. **Note: Statistical**

appeals that appeal the formula used to calculate the achievement profile are not accepted.

Stage 2 – Site Visits for Third Year Underperforming Schools. Schools that are labeled underperforming for a third consecutive year are entitled to a site visit to determine if the designation was warranted. These site visits are important as the third year underperforming designation merits an alternative designation of failing to meet the academic standards in accordance with statute. After the statistical review of appeals takes place, schools designated underperforming for a third consecutive time are scheduled to receive that visit. Teams from the research and evaluation and state intervention sections of the ADE visit each of the third year underperforming schools to gather additional supporting data for the appeal as well as gather information related to the school’s Arizona school improvement plan (ASIP) as required by statute. All information gathered from the site visits is taken to the substantive appeal committee for use in the third stage of the appeal process. Only third year underperforming schools receive site visits as part of the appeals process.

Stage 3- Substantive Appeals Processed. Substantive appeals are resolved in a committee process. All committee members represent a diverse background to ensure that appeals are considered from multiple perspectives. Among those perspectives are those of principals, teachers, school administrators, department administrators, researchers, and Title I representatives.

Once the committee is assembled, the appeals are evaluated utilizing an appeals rubric approved by the State Board of Education that evaluates the argument presented and whether or not the evidence provided to support the argument is compelling. The appeals rubric consists of a three-tiered system for appeal evaluation:

1. Initial review of the appeal to determine its merit.
2. Review of the evidence provided.
3. Committee recommendation.

Initial Review. The substantive appeal rubric provides for three categories that apply during the initial review. Each appeal is classified into the categories based on the information provided in the appeals.

1. ***Data Calculation Discrepancies.*** Appeals of this nature reference when schools attempted to compare data details provided in the common logon with their data sets and found different results with *their* numbers. If in the appeal, the school/district provides information of data discrepancies and those data discrepancies are actual and not a result of the school or district’s inability to replicate the achievement profiles formula, the appeal is deemed as passing the initial review. At this stage, it is important for the school or district to be

specific in their claim of differences so that the ADE can adequately determine if the data sets were in fact different.

2. ***Mitigating Factors Outside the School's Control.*** Appeals of this nature reference when the school indicates significant issues that affected test scores, administration, etc. such as teacher attrition; environmental issues/events; adverse testing conditions; school/community/ emergency/crisis; etc. If a school provides information detailing a *significant* event that impacted test scores, which was clearly outside the school's control the appeal is deemed as passing the initial review.
3. ***Implementation of the School Improvement Plan.*** Appeals in this category discuss how the school is actively, consistently, and reliably implemented the school improvement plan; that the priorities of the solutions team had been addressed; the school improvement plan had been revised and updated to address assessed ongoing needs; professional development that supported the targeted goals had been planned and implemented. If a school provides information highlighting their school improvement plan, it may accentuate the other two components on the rubric. However, per statute, a successful implementation of the school improvement plan alone cannot change a classification of a school. Additionally, only third-year underperforming schools may include information about their ASIP in their appeal.

Review of Evidence. Once the appeals progresses through the first tier of the rubric, *initial review*, the evidence provided to support the appeal is evaluated. In this tier, three determinations are possible:

1. ***Compelling evidence:*** in this area, the school or district adequately provided information that leads the committee to conclude that had the circumstance been different, the achievement profile would have been different as well. For example, if a school presents an argument that their data are different and is able to provide specific and relevant data supporting their claim, they are deemed as providing compelling evidence. If a school has a special circumstance that affected a certain grade and are able to demonstrate that the specific grades test scores suffered, they are deemed as providing compelling evidence. Again, if a school provides compelling evidence highlighting their school improvement plan, it may accentuate the other two components on the rubric (data calculations/mitigating factors). However, per statute, a successful implementation of the school improvement plan alone cannot change a classification of a school.
2. ***Not compelling evidence:*** appeals are categorized in this area when they are able to provide information that a significant issue *could* have impacted the school's performance but did not provide detailed, specific information as to specific outcomes that hindered the school's performance. For example, if a school has high teacher attrition mentioned as a special circumstance, the committee might

accept that appeal during the initial review. However, during the evidence stage, the committee needs to know when the teachers left the school, what grades were affected, and if test scores suffered in that grade (as compared to other grades or prior years). If no specific information is presented, other than there was attrition, that evidence is deemed as not compelling. Another example of this type of outcome is in the case of a school mentioning that they had high teacher attrition in certain grades during a specific year. However, if researching the test scores finds that the grade in question had higher scores in that year than other grades/classes and other test years despite the attrition, the evidence would not be compelling.

3. ***Not applicable evidence:*** if an appeal submitted, makes it through the initial review, and presents evidence that is not linked in any way to the performance of the school, the evidence is deemed not applicable. For example, some appeals describe data discrepancies as a problem yet support their data with evidence that argued the AZ LEARNS methodology (which is not eligible for appeal). In cases such as this, the evidence is deemed inapplicable. If the evidence does not directly support the claim made in the appeal, it is deemed not applicable.

Committee Recommendation. Once the appeal and evidence are reviewed, the committee arrives at a decision as to the outcome of the appeal. There are three possible outcomes:

1. ***Appeal granted and AZ LEARNS determination changed.*** In these cases, the appeal successfully made it through the initial review and evidentiary stages. It is determined that the points needed to change classifications would have been earned by the school had the special circumstance/data discrepancy not occurred. Therefore, the classification for the school is changed.
2. ***Appeal granted and AZ LEARNS determination remains the same.*** In these cases, the appeal successfully made it through the initial review and evidentiary stages. However, it is determined that the criteria needed to change classifications were not earned by the school had the special circumstance/data discrepancy not occurred. For example, a school provided information and evidence that their AYP point was not accurately included in the calculations and the committee determined the school provided information to prove they earned the 1-point set aside for AYP in the AZ LEARNS formula. However, the school needed 5 points to get from underperforming to performing. Therefore, the one point earned was not enough to change their designation, so their determination remains the same.
3. ***Appeal Denied.*** In these cases, the appeal did not successfully make it through the initial review and evidentiary stages. Therefore, the classification for the school remains the same.

Appeal Resolution Notes. If a school submits both a statistical and substantive appeal, the statistical appeal is evaluated first. Only after the statistical arguments have been exhausted and site visits taken place (where applicable) is the appeal sent to the substantive committee for evaluation. Schools need to be certain to provide all information/support when submitting an appeal; late information to support the appeal is not accepted (unless ADE specifically asks for additional information as noted above or the information is gathered in a site visit where applicable).

Stage 3 – Notification of Result Sent to Schools

Once all appeals are resolved, notifications are sent to the schools that filed appeals. The contact person of record for the school receives an email from the ADE with directions as to how to access appeal information via the common logon when the appeal is processed. Schools are notified before the final public release of the achievement profiles as to the outcome of the appeal process. All appeals are final.

.

11. AZ LEARNS Achievement Profiles for Extremely Small Schools

In 2004, the ADE published profiles for extremely small schools for the first time. Extremely small schools are defined as schools in which more than one-third of the AIMS subject-grade combinations cannot be evaluated due to having too few test scores in the baseline years. When AZ Learns achievement profiles were first issued in 2002 the State Board of Education determined that extremely small schools would not receive an achievement profile using the conventional AZ LEARNS methodology, and that ADE should develop an alternate method for evaluating these schools. The ADE worked with its advisory groups to develop such a method. The greatest concern regarding small schools was the impact one or two students in a small group would have on the number of growth points earned by a school. Consequently, the calculation of growth points is excluded from the profile for extremely small schools. Confidence intervals are used to mitigate the impact of small group size in the calculation of baseline points. Schools with too few students in their MAP analysis are not awarded scale score points for MAP and have their profile calculated using an alternate performance scale.

General Process to Produce an Achievement Profile for an Extremely Small School

The method for calculating an achievement profile for an extremely small school is as follows:

1. AIMS scale score points are calculated using a baseline score only. The baseline is calculated by aggregating test scores backward across years starting from the current year. Aggregation is carried out until a group size of 32 is attained or the year 2000, whichever comes first. The percent of students passing AIMS in this aggregate group is then compared to the baseline grouping scale for the relevant subject/grade. The group then receives baseline points based upon which baseline group it is in.
2. Added evidence points are calculated as for all other schools. Exception: Schools with less than 16 students in the MAP analysis will not receive added evidence points. Instead, their total scale score points will be evaluated against a separate scale to determine a label.
3. Graduation/dropout points are calculated as for all other schools.
4. The application of the percent exceeding thresholds for highly performing and excelling schools are calculated as for all other schools.
5. The point scales for school labels are the same as for all other schools.

6. Because of the uncertainty of measurement associated with small sample sizes and the high stakes of school designations, schools initially determined to be underperforming receive a “second look.” Instead of determining baseline groups based on the mean percent of students passing AIMS, alternate baseline groups for these schools are determined using the upper bound of a 95 percent confidence interval around the mean. If a school initially determined to be underperforming moves to a higher classification due to the “second look,” that school receives a performing label.

Definition of an Extremely Small School

An extremely small school is defined as a school in which more than one-third of its subject/grade combinations cannot be evaluated because the average number of usable test scores in the baseline years of 2000-2001 is below the minimum group size (N-size or N-count) of 16.

Example. The following table shows the number of usable test scores over the past four years for Brine Shrimp Middle School. The school serves only 8th grade. Usable test scores are defined in chapter five.

Table 11.1 An Extremely Small School

Subject	2000	2001	2002	2003	2004
Math	15	15	12	15	20
Reading	15	17	12	15	20
Writing	14	14	12	15	20

The minimum group size condition is applied by looking at the average group size in the baseline years of 2000-2001. If the average group size is below 16 then that subject/grade combination is not evaluated using the conventional AZ LEARNS methodology. In practice, the number of usable test scores in the baseline years must add up to 32 or greater. By this condition, reading can be evaluated ($15 + 17 = 32$), but writing and math cannot. Since more than one-third of the subject grade combinations for this school cannot be evaluated, this school meets the definition of an extremely small school.

NOTES.

1. Group size for non-baseline years is not taken into account when determining if a school is extremely small.
2. State law (ARS 15-241) defines a small school as having less than 100 students. In the past a school must have had an average daily membership (ADM) of less

than 100 to be considered extremely small. However, there are a significant number of schools with an ADM of greater than 100 yet fail to make the minimum group size standard. The evaluation method outlined here is applied to all schools in which more than one-third of their subject/grade combinations cannot be evaluated, regardless of ADM.

Calculation of AIMS Scale Score Points

Subject/grade combinations with less than 16 students in the baseline years cannot be evaluated using the conventional AZ LEARNS method. The baseline-growth method for calculating scale score points is too sensitive to changes in the scores of individual students for very small groups. Consequently, an alternate method was been developed to calculate AIMS scale score points for the AZ LEARNS achievement profile.

The method is as follows:

1. For each subject/grade combination, an evaluation group is created by aggregating usable test scores across years starting in the most current year. Scores are added to the group until: 1) A group size of 32 is reached, or 2) Data for all available years is reached.
2. For each evaluation group, the percentage of students who passed AIMS is calculated.
3. The percentage passing for each evaluation group is compared to the AZLEARNs cut points for baseline groups.
4. Each subject/grade combination is awarded scale score points based on the baseline group it is in.

Example. Table 11.2 below shows the number of usable tests for third grade reading for two different schools.

Table 11.2. Number of Usable Third Grade Reading Tests for Two Schools

School	2000	2001	2002	2003	2004	Total Group Size
Javelina	14	15	13	12	15	40
Gila Monster	4	5	6	6	5	26

Applying the method described above, for each school we form an evaluation group by counting backward from the most current year until either we achieve a group size of 32

or use all the available data. The years of data used for each school are shown in **boldface**. For Javelina this results in a group of 40 students from the years 2002, 2003, 2004. For Gila Monster, we must use all the years of available data back to 2000.

The next step is to examine the percent of students passing in the groups we have formed. Table 3 shows the number of students who passed in the groups from Table 11.2. Again for Javelina we consider only the most recent three years while for Gila Monster we consider five years.

Table 11.3. Number of Students Passing AIMS, Third Grade Reading

School		2000	2001	2002	2003	2004	Total	% Pass
Javelina	# Pass	NA	NA	8	8	5	21	53
	# Tested	NA	NA	13	12	15	40	
Gila Monster	# Pass	2	3	5	5	5	20	77
	# Tested	4	5	6	6	5	26	

For Javelina, 53 percent of the students in the evaluation group have passed AIMS. Comparing this to the cut points for baseline groups (table 5.1) we see that this puts third grade reading for Javelina in baseline group two, earning the school two scale score points.

For Gila Monster, 77 percent of students in the evaluation group have passed AIMS. Comparing this to the cut points for baseline groups (Table 5.1) we see that this puts third grade reading for Gila Monster in baseline group four, earning the school four scale score points.

Calculation of an Achievement Profile for an Extremely Small School

In addition to AIMS scale score points, schools may also earn scale score points via added evidence, their graduation and dropout rates, and their current AYP status. The methods used to calculate scale score points earned by extremely small schools for these performance measures is the same as the methods used for other schools.

- A school may receive up to six scale score points (calculated using the method described above) for each subject/grade combination it serves.
- A school may receive one point if it has made adequately yearly progress (AYP) under the methodology mandated by the No Child Left Behind Act.
- If a school is an elementary school, it may receive added evidence points based on its performance on the Measure of Academic Progress (MAP).

- If a school is a high school it may receive up to two points based on its graduation and dropout rates.
- The total points earned by a school are added up and compared to the school classification scale to determine a school's preliminary classification.
- In order to be classified as a highly performing or excelling school, a school must meet the thresholds for percentage of students exceeding the standard.

“Second look” for extremely small schools

Because of the high-stakes consequences of being designated as an underperforming school, and because of the uncertainty of measurement involved with small sample sizes, it is prudent to give extremely small schools a “second look” if they face the possibility of receiving an underperforming classification. If the preliminary designation of an extremely small school is underperforming, then the AIMS scale score points for that school are recalculated. For each subject/grade combination, the upper bound of the 95-percent confidence interval is used to calculate to which baseline group the school belongs. If the recalculated points move the school into a higher classification, the school receives a performing classification.

Let p =the percent of students in a group passing the AIMS and n =the number of students in the group. Then the equation for the upper bound of the 95 percent confidence interval (UB95) is:

$$UB95 = p + 1.96\sqrt{p(1-p)/n}.$$

As can be seen from the equation, the confidence interval depends upon the percent of students who passed the test, and the number of students tested. Thus, the confidence interval will differ among grades, subjects, and schools.

The equation is an approximation of the confidence interval for a binomially distributed variable. It uses the standard normal distribution and is sufficiently accurate if the group size and percentage of students passing are large enough. For small values of n and small p , a more accurate estimate of the confidence interval was made using statistical tables that provide confidence intervals for a binomially distributed variable.¹ The tables were applied using the rules given in table 11.4.

¹ Mansfield, Edwin. 1991. *Statistics for Business and Economics, 4th Edition*. New York: W.W. Norton and Company. 280-284.

Table 11.4. Rules for Determining UB95 for Small n and p.

If $n \geq 0$ and $n < 8$, and
 $p \geq 0$ and $p < .04$, then
UB95=.42;
 $p \geq .04$ and $p < .10$, then
UB95=.50;
 $p \geq .10$ and $p < .20$, then
UB95=.60;

If $n \geq 8$ and $n < 10$, and
 $p \geq 0$ and $p < .04$, then
UB95=.37;
 $p \geq .04$ and $p < .10$, then
UB95=.45;
 $p \geq .10$ and $p < .20$, then
UB95=.55;

If $n \geq 10$ and $n < 12$, and
 $p \geq 0$ and $p < .04$, then
UB95=.33;
 $p \geq .04$ and $p < .10$, then
UB95=.41;
 $p \geq .10$ and $p < .20$, then
UB95=.52;

If $n \geq 12$ and $n < 16$, and
 $p \geq 0$ and $p < .04$, then
UB95=.27;
 $p \geq .04$ and $p < .10$, then
UB95=.35;
 $p \geq .10$ and $p < .20$, then
UB95=.47;

If $n \geq 16$ and $n < 20$, and
 $p \geq 0$ and $p < .04$, then
UB95=.24;
 $p \geq .04$ and $p < .10$, then
UB95=.32;
 $p \geq .10$ and $p < .20$, then
UB95=.44;

If $n \geq 20$ and $n < 24$, and
 $p \geq 0$ and $p < .04$, then
UB95=.21;
 $p \geq .04$ and $p < .10$, then
UB95=.29;
 $p \geq .10$ and $p < .20$, then
UB95=.42;

If $n \geq 24$ and $n < 30$, and
 $p \geq 0$ and $p < .04$, then
UB95=.18;
 $p \geq .04$ and $p < .10$, then
UB95=.27;
 $p \geq .10$ and $p < .20$, then
UB95=.38;

If $n \geq 30$ and $n < 40$, and
 $p \geq 0$ and $p < .04$, then
UB95=.15;
 $p \geq .04$ and $p < .10$, then
UB95=.23;
 $p \geq .10$ and $p < .20$, then
UB95=.36;

Example. Kangaroo Rat Middle School, which serves grades seven and eight, has earned a total of ten scale score points: seven via its students' performance on AIMS, two points via MAP, and one point for making AYP. This is less than the twelve points needed to be a performing school. The following table shows the school's performance for grade eight.

Table. 11.5. AIMS Performance

Subject	Percent proficient	Scale score points	Number of test scores	Upper bound of confidence interval %	Second look scale score points
Math	6	2	20	29	5
Reading	31	2	20	51	3
Writing	29	3	20	51	4
Total		7			12

As the table shows, the upper bound of the confidence interval for math is 29 percent proficient, which earns the school five scale score points instead of two for the subject. The school also earns one additional scale score point in reading and writing. This gives the school a total of $12+2+1=15$ scale score points. This is a sufficient number of points for the school to be designated highly performing, however the school is designated as performing. The second look policy only enables a school facing an underperforming determination to move into the performing category.

Classification Scales for Extremely Small Schools

The classification scales used for extremely small schools are the same used for other schools, with one exception. If an extremely small school has less than sixteen students in its MAP analysis, then the school does not receive added evidence points and instead is classified using the scale in table 11.6. It is likely that extremely small schools will have very few or no students matched across years for the MAP analysis. Consequently, the MAP analysis, and thus the points earned for an achievement profile, would be sensitive to the performance of one or two students. In order to avoid this problem, elementary schools with less than sixteen students in the MAP analysis do not receive added evidence points and are evaluated against a separate scale.

**Table 11.6. High School Classification/Non-MAP
Elementary School Cut Points**

	Subject/Grade Combination 1	Subject/Grade Combination 3
Underperforming	< 3.2	< 9.6
Performing	3.2	9.6
Highly Performing	5	15
Excelling	5.4	16.2

12. AZ LEARNS Achievement Profiles for Alternative Schools

In 2004, the ADE published profiles for alternative schools for the first time. Alternative schools are defined as schools that meet the Board-approved definition as schools whose sole and clearly-stated mission is to serve specific populations of at-risk students. Alternative school status is granted by application to the ADE. A.R.S. §15-241 makes an allowance for a “parallel” evaluation method for alternative schools. When AZ Learns achievement profiles were first issued in 2002 the Board determined that alternative schools would not receive an achievement profile using the conventional AZ LEARNS methodology, and that ADE should develop an alternate method for evaluating these schools.

General Process to Produce Achievement Profiles for Alternative Schools

The small number and wide variety of alternative schools makes it difficult to develop reliable parameters for measures of school performance. Consequently, a single method of evaluating alternative schools was created that employed performance measures available to the majority of schools. In plainer terms, it would be too inaccurate to develop separate methods for large alternative high schools, small alternative high schools, alternative middle schools, and small alternative middle schools. Given the small number of schools in each category, benchmark performances and expectations would be very unreliable. Consequently, ADE developed a single rubric to evaluate all alternative schools.

The method for calculating an achievement profile for alternative schools is as follows:

1. AIMS scale score points are calculated using a baseline score only. The baseline is calculated by aggregating test scores backward across years starting from the most current year. Aggregation is carried out until a group size of 32 is attained or the year 2000, whichever comes first. The percent of students passing AIMS in this aggregate group is then compared to the baseline grouping scale for the relevant subject/grade. The group then receives baseline points based upon the baseline group it is in. See chapter eleven on extremely small schools for an explanation of this calculation.
2. Added evidence points are not included in the achievement profile for alternative schools. Because most alternative schools below the high school level are quite small and have highly mobile populations, there would be little or no students in the Measure of Academic Progress analysis. Consequently, the MAP analysis would be highly volatile and inaccurate.

3. Points based on an alternative school's dropout rate are calculated as for all other schools.
4. As with other schools, alternative schools will receive one scale score point if they make AYP for the current year.
5. Alternative schools will only receive determinations of performing and underperforming.

Because of the uncertainty of measurement associated with small sample sizes and the high stakes of school designations, schools initially determined to be underperforming receive a "second look." Instead of determining baseline groups based on the mean percent of students passing AIMS, an alternate baseline group for these schools is determined based on the upper bound of a 95 percent confidence interval around the mean. If a school initially determined to be underperforming moves to a higher classification due to the "second look," that school will receive a performing designation. See chapter eleven on extremely small schools for the details of this calculation.

Definition of an Alternative School

The following is the definition of an alternative school as approved by the Board of Education in 2002. There are currently 138 schools that have been granted alternative school status.

1. A school operated by a school district must have adopted a mission statement that clearly identifies its purpose and intent to serve a specific student population (please see criterion three) that will benefit from an alternative school setting. A charter school must be expressly chartered to serve a specific student population that will benefit from an alternative school setting.
2. The educational program and related student services of the school must match the mission or charter of the school.
3. The school must intend to serve students exclusively in one or more of the following categories:
 - Students with behavioral issues (documented history of disruptive behavior)
 - Students identified as dropouts
 - Students in poor academic standing who are either severely behind on academic credits (more than one year) or have a demonstrated pattern of failing grades
 - Pregnant and/or parenting students

- Adjudicated youth
4. Any school offering secondary instruction for academic credit used to fulfill Arizona State Board of Education graduation requirements (in part or in full) must offer a diploma of high school graduation.

Calculation of an Achievement Profile for an Alternative School

In addition to AIMS scale score points, alternative schools also earn scale score points via their dropout rates and their current AYP status. The methods used to calculate scale score points earned by alternative schools for these performance measures are the same as the methods used for other schools.

Figures 12.1 and 12.2 give a graphical summary of the method for evaluating alternative schools at the elementary and high school levels.

- A school may receive up to six scale score points for each subject/grade combination it serves. This is represented by the far left column in each figure.
- A school may receive one point if it has made adequately yearly progress (AYP) under the methodology mandated by the No Child Left Behind Act. This is represented by the column second-from-left in each figure.
- If a school is a high school it may receive one point based on its dropout rate. This is the third-from-left column in Figure 12.2.
- The total points earned by a school are added up and compared to the school classification scale to determine a school's preliminary classification—the final column in Figures 12.1 and 12.2.

Figure 12.1. Method for Evaluating Alternative Schools (358)

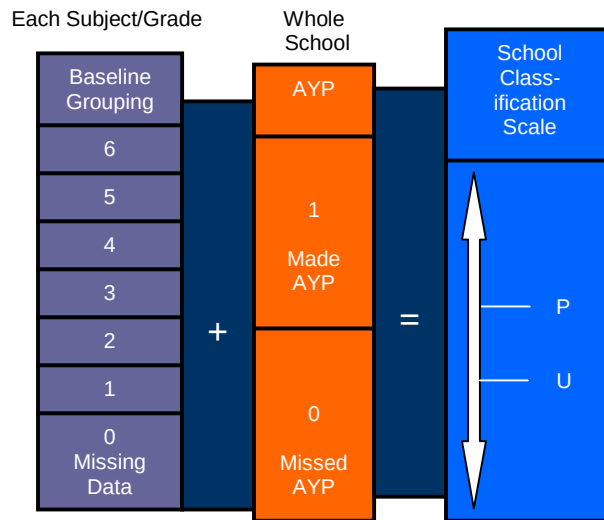
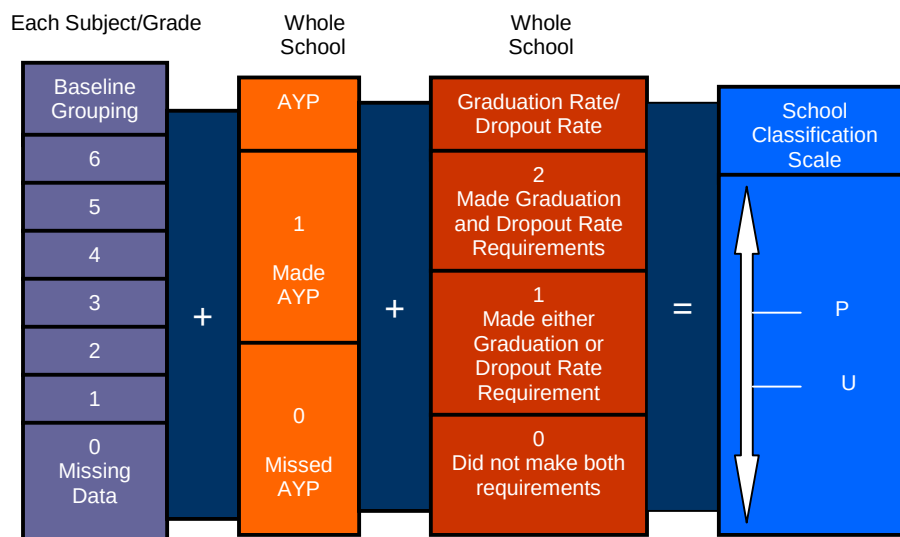


Figure 12.2. Method for Alternative Schools (High Schools)



Classification Scales for Alternative Schools

The following classification scales are used for alternative schools. Separate classification scales are used for high schools and elementary schools due to extra points being available to high schools from the dropout rate. The scale score points are roughly

equivalent to expecting an alternative school to attain baseline group two in all subject/grade categories. This implies that the school performance of students on AIMS would lie between the 10th and 24th percentiles of schools in the state in 2000-2001.

Table 12.1. Alternative School Classification Cut Points

		Subject/Grade Combination 1	Subject/Grade Combination 3	Subject/Grade Combination 6	Subject/Grade Combination 9
358 Schools	Underperforming	< 2.0	< 5.0	< 11.0	< 16
	Performing	2.0	5.0	11.0	16
High Schools	Underperforming	<2.0	<6.0	<12.0	<18.0
	Performing	2.0	6.0	12.0	18.0

13. AZ LEARNS Achievement Profiles for K-2 Schools

When AZ Learns achievement profiles were first issued in 2002 the Board determined that K-2 schools would not receive an achievement profile using the conventional AZ LEARNS methodology, and that ADE should develop an alternate method for evaluating these schools. In 2004, the ADE published profiles for K-2 schools for the first time. K-2 schools are schools that serve only grades kindergarten through second grade. Since the AIMS is not administered at any of the grades served by these schools, the AZ LEARNS profiles are based solely on the performance of the schools' second graders on the Stanford-9 test.

The method of calculating the profile for these schools is straightforward:

1. The average percentile ranks on the reading and math portions of the test are calculated for the most current year for a school's second graders.
2. The average percentile ranks for the school are added together, and
3. Compared to a scale to determine the school's classification.

Example. In 2004, the average percentile rank for second graders in Gila Monster Elementary was 52 for math and 48 for reading. The two averages summed together yield $52 + 48 = 100$ points. This is sufficient for Gila Monster to be a performing school. (Please see the scale on the next page.)

Setting the scale

The cut point for performing schools is set at the sum of 68—equivalent to the expectation that students in a performing school average in the 34th percentile in reading and math. It is equal to one standard deviation below the state average. In other words, underperforming schools are in the bottom 16 percent of schools compared to performance of second graders statewide.

The cut point for highly performing schools was set at the sum of 130—equivalent to the expectation that students in a highly performing school average in the 65th percentile in reading and math. It is equal to one-half standard deviation above the state average. In other words, highly performing schools are in the top 31 percent of schools in reading and math compared to performance of second graders state.

The cut point for excelling schools was set at the sum of 150—equivalent to the expectation that students in a highly performing school average in the 75th percentile in reading and math. It is equal to one standard deviation above the state average. In other

words, highly performing schools are in the top 16 percent of schools in reading and math compared to performance of second graders state.

Table 13.1 illustrates the cut points for the K-2 schools.

Table13.1. AZ LEARNS Scale for K-2 Schools

Points	Achievement Profile
<68	Underperforming
68	Performing
130	Highly Performing
150	Excelling

14. Small School Score Adjustment

Criteria

Some schools are eligible for low score outlier removal as a result of small overall size. The following criteria are used to determine which schools should receive this adjustment:

- The school has an average daily membership (ADM) of 99 students or fewer and the school has AIMS scores for ninety-nine students or fewer.
- Or, if the school did not report ADM but has valid AIMS scores for 99 students or fewer.

Method

Small school adjustments were made to schools that met these criteria. The following steps were taken in adjusting the scores of qualified schools:

1. Low and high score outlier identification was conducted in each of the subjects of reading, writing and math for grades 3, 5, 8 and high school respectively. A decision was reached to remove low scores only. It was reasoned that statistically low outlier scores represented chance occurrences of poor individual level test performance. Removing low outlier scores from a distribution of test scores was particularly important for schools with small size because it negated the impact that a very low score(s) might have on the overall picture of the school's performance. These scores were identified by employing a bivariate regression model where each school's mean scale score was regressed on a two-category indicator for the school:

Where 1 = the student was tested in that school, and
0 = the student was not tested in the school.

2. A unique equation was estimated for each school in order to produce a statistic that would address the question about the affect that any one student's score *within* that school would have on the small school's overall mean scale score. This equation took the following form:

Mean scale score_{*i*} = α + b School Dummy_{*i*}

In this application, the regression coefficient plus or minus the intercept represented the school's mean scale score.

3. The DFFIT statistic was computed for each score within the school and saved to a separate data file for later analysis. The DFFIT statistic assesses the change that a particular case will have on the predicted value i.e., the small school's mean scale score when it has been deleted from the regression equation (Belsey et al. 1980)². The following equation was used:

$$\text{DFFITS}_i = t_i \sqrt{\frac{h_i}{1-h_i}}$$

where t_i = studentized residual
and h_i = leverage value

4. Once the DFFITS were estimated for each case, the standard deviation of the statistic was computed *within* each school. A cut-point was set such that any student with a DFFIT score greater or less than two standard deviations above or below the within school mean DFFIT could be considered an outlier relative to the distribution of test scores of that school.
5. Students were assigned an identification number so that their observation could be flagged if it fell above or below this school specific cut-point. Scores that had been identified statistically as outliers and were at the low end of the school's test score distribution were eliminated from the datasets that were used to compute the final performance label.

² Belsley, D.A., Kuh, E. and Welsch, R.E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley.

Appendix VI: Substantive Appeal Rubric for AZ LEARNS

Evaluation Criteria	Initial Review (Please check the applicable option)			Review of Evidence Provided
Data Calculation Discrepancies i.e., school attempts to compare data details with their data sets and gets different numbers	Not applicable	ADE data are accurate and calculations are correct.	Data does not match that of ADE. School submits evidence of discrepancies and provides additional data.	☐ Compelling evidence ☐ Not compelling evidence ☐ Not applicable evidence
Special Circumstances Outside the Control of School/District Administration or Management i.e., school indicates significant teacher attrition; environmental issues/events; adverse testing conditions; school/community emergency/crisis	Not applicable	Special circumstances that were outside of the school's control, were not a substantial cause of the overall school performance.	School had a situation that was unavoidable and outside of the school's control and hindered the test administration or student performance. This situation resulted in adverse data for the year(s) in question.	☐ Compelling evidence ☐ Not compelling evidence ☐ Not applicable evidence
Policy/Methodology Issues i.e., school disagrees with use of two year baseline	The ADE will not accept/review appeals related to policy/methodology			

Team Decision AZ LEARNS Substantive Appeal
 (Results represent group consensus regarding appeal)

Reason Reviewed	Initial Review			Review of Evidence		
Data Calculation	N/A	Correct data/calcul	Data does	N/A	Compelling	Not compell

Discrepancies		ation	not match		evidenc e	ing evidenc e
Special Circumstances	N/ A	Did not cause overall performan ce	Adve rse result based on situati on	N/ A	Compel ling evidenc e	Not compell ing evidenc e
Policy/Methodo logical Issues	The ADE will not accept/review appeals related to policy/methodology.					

Please indicate appropriate response(s) by checking within the box(es) provided.

Committee Recommendation: Granted Denied

Final Appeal Decision: Granted Denied

Comments:

Appeal Result: