
Arizona's School Accountability System *Technical Manual*

Volume II: Adequate Yearly Progress

Prepared by Robert Franciosi and Jennifer Regalado
Education Research Associates

Research & Policy
Arizona Department of Education
Ildiko Laczko-Kerr, Deputy Associate Superintendent

Tom Horne
Superintendent of Public Instruction
Arizona Department of Education

Published by Arizona Department of Education February 17, 2004.

Table of Contents

Introduction.....	1
Overview of the NCLB Evaluation System.....	3
Schools to Be Evaluated	3
Proficiency Standards	3
Percentage of Students Assessed	8
Applicable Subgroups.....	8
Additional Indicators of School Performance.....	9
Putting It All Together	9
Data Verification	11
Timeline	12
Meeting the Annual Measurable Objectives for Proficiency.....	14
Calculation.....	14
Data used.....	18
Special rule.....	19
Meeting the Standard for Number of Students Tested.....	20
Calculation.....	20
Data used.....	20
Applicable subgroups.....	21
Special rules	23
Other Indicators of School Performance.....	24
Attendance Rate	24
Graduation Rate	25
Calculation of Adequate Yearly Progress for K-2 Schools	26
Meeting the Annual Measurable Objectives for Proficiency for K-2 Schools.....	26
Attendance Criteria for K-2 Schools.....	27
Summary of AYP Determination for K-2 Schools	27
Determining Adequate Yearly Progress for School Districts and Charter Holders	29
Differences between District and School AYP Evaluation Methods.....	29
The Adequate Yearly Progress Appeals Process	32
Step One: Data Verification.....	32
Step Two: Appeal Application.....	32
Step Three: Appeal Resolution.....	34

List of Tables and Figures

Table 1.1 Comparison of Arizona’s Accountability Systems.....	2
Table 2.1. Calculation of Performance Starting Points.....	4
Table 2.2 Starting Points for State Performance Standards	5
Table 2.3 Annual Measurable Objectives (AMOs) and Intermediate Goals	6
Figure 2.1 Intermediate Goals: Grade 3 Reading	7
Table 2.4. Categories Evaluated Under NCLB for a K-5 Elementary School.....	10
Table 4.1. Rules for Determining UB99 for Small n and p.	16
Table 5.1. Allocation of Students with “Other/Blank” Ethnicity	22
Table 7.1. AYP Determination for K-2 Schools	28

Introduction

The federal No Child Left Behind Act of 2001 (NCLB) requires states to establish an accountability system to evaluate the performance of local public schools and school districts, including charter schools. Specifically, states are required to:

1. Institute performance standards for reading/language arts, mathematics, and science.
2. Develop and administer tests to measure whether students meet these standards. By the 2005-06 academic year, states must give tests in reading/language arts and math for grades three through eight. By the 2007-08 academic year, states must also administer a test to evaluate student performance in science in elementary (grades 3-5), middle (grades 6-8), and high school (grades 10-12).
3. Establish a timeline to ensure that all students are proficient according to state standards by 2013-2014.
4. Create a statewide accountability system to evaluate school progress in meeting the goals of the timeline, and issue report cards informing parents of school performance.

In 2001 Arizona voters also approved Proposition 301 that among other things called for a state accountability system for public schools. In 2001, Arizona also had in place state standards and a test to measure whether students met them: Arizona's Instrument to Measure Standards (AIMS).

Since the passage of NCLB and Proposition 301 the staff of the Arizona Department of Education has worked with scholars, school officials ranging from superintendents to teachers, and members of the public to develop an accountability system that fulfills the requirements of both laws. The result is a system that consists of two linked components. Arizona LEARNS was created to comply with Proposition 301. Its primary focus is on longitudinal change through time of student performance as measured by AIMS and the Stanford 9 tests. The system created to comply with NCLB, commonly referred to as Adequate Yearly Progress (AYP), provides a single-year snapshot of school performance as measured by AIMS. Table 1.1 provides a brief comparison of the two accountability systems.

The State of Arizona's complete plan to meet the requirements of NCLB is contained in the workbook submitted to the U.S. Department of Education. The workbook is available at <http://www.ade.az.gov/azlearns/workbook.asp>.

Table 1.1 Comparison of Arizona's Accountability Systems

NCLB	Arizona LEARNS
Required by federal law	Required by state law
One-year snapshot of student performance	Longitudinal examination of student performance
Components of evaluation • AIMS scores • Percent students assessed • Attendance/Graduation rates	Components of evaluation • AIMS scores • Measure of Academic Progress • Graduation/dropout rates • AYP
Labels schools on a yes/no system	Labels schools on a graded scale: • Failing to meet academic standards • Underperforming • Performing • Highly performing • Excelling

Overview of the NCLB Evaluation System

This section provides an overview of the determination of adequate yearly progress. More detailed discussions of the methodology used to determine AYP, including descriptions of equations, algorithms, and data used are given in subsequent chapters.

The No Child Left Behind Act requires that every public school and district in a state—as well as the state itself—be evaluated on three measures:

1. Progress toward meeting the goal of 100 percent proficiency in state standards;
2. Percentage of students assessed; and
3. An additional measure of school performance. NCLB mandates that for high schools this indicator be the graduation rate. States may select an alternative indicator for elementary schools. Arizona, along with many other states, has chosen attendance rate for the other indicator for elementary schools.

If an entity—school, district, or state—passes on all three measures, then it is deemed to have made adequate yearly progress (AYP).

Schools to Be Evaluated

Extremely small schools, new schools, and schools that only offer grades K-2 are not evaluated under Arizona Learns. These exemptions are not granted under NCLB, all schools must receive an AYP determination. Similarly, although the state's system for school accountability, Arizona LEARNS, allows alternative schools to be evaluated under different criteria, NCLB requires *all* public schools in the state to be given an AYP designation based on the same criteria.

Proficiency Standards

NCLB requires that every student in Arizona meet state standards in reading/language arts and mathematics—that is, pass AIMS—by the year 2013-2014. To further this goal, the state must set annual

**Setting of Baselines,
Annual Measurable
Objectives, and
Intermediate Goals**

measurable objectives (AMOs) for each grade and subject evaluated. The annual measurable objectives describe the yearly growth in the fraction of students passing AIMS that is necessary for Arizona to reach the 100 percent requirement by 2013-2014. The AMOs are then used to set intermediate goals. To make AYP an entity must reach the intermediate goals for every subject in each grade it offers. If an entity fails to reach an intermediate goal, it still may be deemed to have made adequate yearly progress if it satisfies the safe harbor provisions that will be described later.

The Arizona Department of Education established the starting points, annual measurable objectives, and intermediate goals in the manner specified by the No Child Left Behind Act. To determine the baselines for each subject/grade combination, all schools in Arizona were ranked in descending order according to the percentage of students passing AIMS for that subject and grade. Then, cumulative enrollment was calculated adding upward from the bottom of the list of schools. The baseline was then set to be equal to the fraction of students passing AIMS for that grade and subject in the school where the cumulative enrollment was equal to 20 percent of state enrollment for that grade. The data used for this calculation were AIMS results for the spring of 2002. As required by NCLB, students with invalid scores such as English language learners and special education students who received nonstandard accommodations were included in the setting of the baselines.

Table 2.1 provides a hypothetical example of how the baselines were established. In this case, we assume there are only eight schools in the state that offer third grade.

Table 2.1. Calculation of Performance Starting Points

Grade	Subject	School	Percent pass	Enrollment	Cumulative percent of total state enrollment
3	Math	1	100	10	100
		2	75	40	95
		3	70	30	75
		4	61	30	60
		5	55	20	45
		6	48	30	35
		7	32	20	20
		8	15	20	10

These eight schools are ranked in descending order by the percentage of their students who passed the AIMS for third grade math (fourth column). The third grade enrollment for each school is given in the fifth column. Starting from the bottom of the list, enrollment is summed until the total equals 20 percent of the state's total enrollment for that grade. In table 2.1 this point is

reached at School Seven, where the cumulative sum equals forty students ($40/200 = 0.20$). The percent of students passing for School 7 (32 percent) is then taken as the starting point for the state for third grade math.

Table 2.2 provides the starting points for each of the subjects and grades evaluated in 2003. These served as the AMOs for the first year of AYP determination (2003).

Table 2.2 Starting Points for State Performance Standards

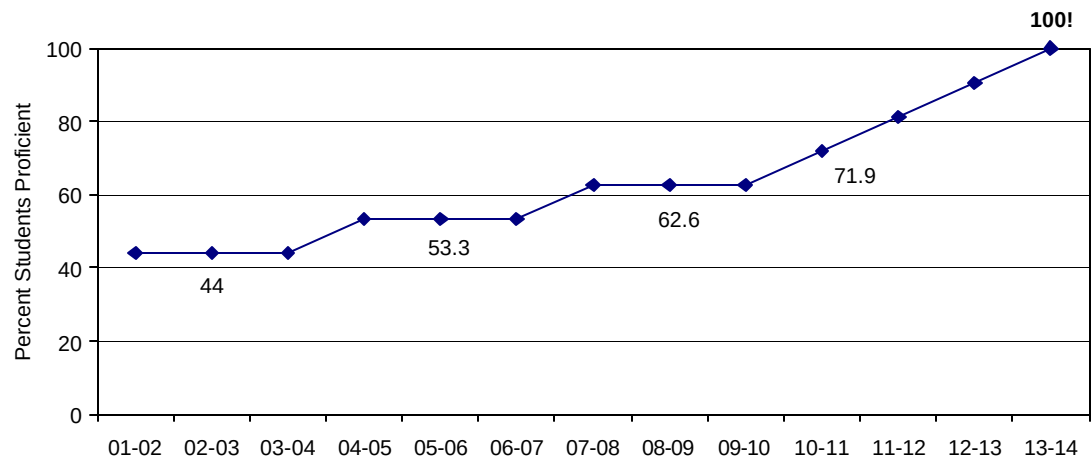
Subject/Grade	Reading	Mathematics
Grade 3	44	32
Grade 5	32	20
Grade 8	31	7
High School	23	10

The annual measurable objectives were calculated as six equal percentage-point increments from the 2002 starting point to the 2014 goal of 100 percent. The AMO for third grade reading, for example, is 9.3 percentage points ($[100-44]/6$). The AMOs cover three-year increments through 2010, and one-year increments thereafter. This leads to a stepwise increase in the intermediate goals until 2010, followed by a linear increase until 2014 (see table 2.3). Figure 2.1 shows an example using third grade reading for the increase in the intermediate goals.

Table 2.3 Annual Measurable Objectives (AMOs) and Intermediate Goals

Grade 3	Inter- mediate Goals	Reading AMO	Reading Proficiency (percent)	Math AMO	Math Proficiency (percent)
2004-05	1	9.3	53.3	11.3	43.3
2007-08	2	9.3	62.6	11.3	54.6
2010-11	3	9.3	71.9	11.3	65.9
2011-12	4	9.3	81.2	11.3	77.2
2012-13	5	9.3	90.5	11.3	88.5
2013-14	6	9.3	100	11.3	100
Grade 5	Inter- mediate Goals	Reading AMO	Reading Proficiency (percent)	Math AMO	Math Proficiency (percent)
2004-05	1	11.3	43.3	13.3	33.3
2007-08	2	11.3	54.6	13.3	46.6
2010-11	3	11.3	65.9	13.3	59.9
2011-12	4	11.3	77.2	13.3	73.2
2012-13	5	11.3	88.5	13.3	86.5
2013-14	6	11.3	100	13.3	100
Grade 8	Inter- mediate Goals	Reading AMO	Reading Proficiency (percent)	Math AMO	Math Proficiency (percent)
2004-05	1	11.5	42.5	15.5	22.5
2007-08	2	11.5	54.0	15.5	38.0
2010-11	3	11.5	65.5	15.5	53.5
2011-12	4	11.5	77.0	15.5	69.0
2012-13	5	11.5	88.5	15.5	84.5
2013-14	6	11.5	100	15.5	100
High School	Inter- mediate Goals	Reading AMO	Reading Proficiency (percent)	Math AMO	Math Proficiency (percent)
2004-05	1	12.8	35.8	15	25
2007-08	2	12.8	48.6	15	40
2010-11	3	12.8	61.4	15	55
2011-12	4	12.8	74.2	15	70
2012-13	5	12.8	87.0	15	85
2013-14	6	12.8	100	15	100

Figure 2.1 Intermediate Goals: Grade 3 Reading



The reasons for setting all annual measurable objectives (and corresponding intermediate goals) in this stepwise manner were:

- 1.) The ADE completed a grade-level articulation of Arizona's Academic Content Standards in 2003. The progressive setting of annual measurable objectives and corresponding intermediate goal allows schools the necessary time to align these grade-level standards with school curricula/resources and implement these standards via instruction.
- 2.) The ADE is developing new assessments for grades four, six, and seven for reading and mathematics, as well as a science assessment to be administered on an annual basis in grades three, five, eight, and high school as mandated by NCLB. The progressive setting of annual measurable objectives and intermediate goals allows schools the opportunity to effectively prepare students for these assessments.
- 3.) Currently, the academic performance of several disaggregated student subgroups is below (in some cases, far below) the state's starting points in reading and mathematics. Many schools and districts have initiated scientifically based research programs and other instructional practices to assist students in these groups. In addition, the ADE has implemented a comprehensive K-3 reading program designed to have all students proficient in the state's reading standards by the third grade. By setting

**Refinements to AMO
Determination**

the state's annual measurable objectives and corresponding intermediate goals in a progressive manner, schools, districts, and the state are given the necessary time to effectively implement these programs and initiatives, giving students in this circumstance an opportunity to catch up with the aggregated student population as represented by the state's starting points.

There are two additional steps taken when determining if a school has met the AMO for a specific subject and grade. First, rather than comparing the actual percentage of students who are proficient to the AMO, a 99 percent confidence interval is calculated to estimate the percent proficient. If the upper bound of this confidence interval is above the AMO, the school is deemed to have met the objective.

Second, if a school fails to meet the objective after the confidence interval is applied, it may still be deemed to have met the AMO if it meets the safe harbor provision. Safe harbor is a two-part test that requires schools to demonstrate sufficient progress over the previous year in the percentage of students meeting the standard *and* meet a threshold set by the Arizona Department of Education for an additional indicator. Both of these refinements will be discussed in more detail later.

Percentage of Students Assessed

In order for a school, district, or the state to make adequate yearly progress it must assess 95 percent of its students for each subject in every grade offered, including each applicable subgroup. Students count as assessed if they had a valid score for AIMS or the alternative assessment for the severely disabled, AIMS-A.

All the students enrolled for the day (high school) or the first day of the week the test was given (elementary) represent the population to be assessed.

Applicable Subgroups

In addition to assessing 95 percent of its students and meeting the intermediate goals for all subject/grade combinations it encompasses, an entity must also meet the same objectives for every applicable subgroup within each subject/grade combination. NCLB specifies the following subgroups be evaluated: the five major ethnic groups—Hispanic, White, African-American, Asian-Pacific Islander, and Native American—English Language Learners (ELL), students with disabilities, and students from low-income families.

Additional Indicators of School Performance

NCLB requires that an additional indicator be used for AYP determinations. The law mandates that a four-year graduation rate be used for high schools, but allows state to select the standard schools must meet. The performance goal for the high school graduation rate was set at 70.5 percent, the state average graduation rate for 2001. To make adequate yearly progress, a high school must have a four-year graduation rate of 70.5 percent, or show a 1 percentage-point improvement in the graduation rate over the previous year.

NCLB allows states to select the additional indicator used for elementary schools. Arizona has chosen to use the school-wide attendance rate. The performance goal for the attendance rate was set at 94 percent, the implicit expectation for school attendance rates set by the state's school finance laws in A.R.S. § 15-902 A. To make AYP, elementary schools must have a school-wide attendance rate of 94 percent, or show a 1 percentage-point improvement in the attendance rate over the previous year.

Putting It All Together

Table 2.4 provides an example of how the three performance measures—proficiency in state standards, percentage of students assessed, and an additional indicator—are combined to determine whether a school has made AYP. The example given is for a K-5 school. The school is evaluated based on student performance on AIMS reading and mathematics tests for grades 3 and 5, the percentage of students evaluated for each test and attendance rates. Since our example is an elementary school, all the combinations for which a typical elementary school would be evaluated under NCLB are provided; there are 73 separate combinations examined.

NCLB requires that schools be evaluated using a conjunctive model. That is, to make AYP, a school must meet the performance objective in *every* category in which it is evaluated. For example, if the school in table 2.4 fails to meet the objective in any one of the cells in the table, it fails to make AYP.

Table 2.4. Categories Evaluated Under NCLB for a K-5 Elementary School

Grade	Third				Fifth			
Subject	Math		Reading		Math		Reading	
Subgroup	Met 95% tested?	Met AMO?	Met 95% tested?	Met AMO?	Met 95% tested?	Met AMO?	Met 95% tested?	Met AMO?
All students	Yes/No	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
African American	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
Asian-Pacific Islander	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
Hispanic	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
Native American	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
White	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
Special Education	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
English Language Learner	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
Low Income	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
Met Other School wide Indicator: Attendance Rate?					Y/N			

Data Verification

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 schools' data before preliminary AYP calculations were carried out. From August 14 through August 27, 2003 an application to verify school level data was made available to districts and charter holders through the common logon on the ADE web site. The application displayed all the data that were to be used for the calculation of adequate yearly progress. In addition, a link was provided through the common logon that allowed schools to download student-level testing data in order to make corrections and facilitate the data verification process.

All schools in districts or under a charter holder that expected to receive an AYP determination were required to verify data and additional school information, which included:

Elementary Schools (K-8)

- Attendance rate (ADM)
- Enrollment on April 28, 2003 (the first day of testing for elementary schools)
- Number of students tested (includes AIMS and AIMS-A students)
- AIMS results
- Data disaggregated by subject

High Schools (9-12)

- Attendance (ADM)
- Enrollment on April 28 and 29, 2003 (dates of AIMS testing for high schools)
- Number of students tested (includes AIMS and AIMS-A students)
- AIMS results
- Graduation rate
- Data disaggregated by subject

It is important to note that districts and charter holders were solely responsible for verifying information for their schools. If a district or charter holder did not verify the information for its schools through the verification process, the ADE assumed the schools on

file and the data available were correct and complete as listed and produced an AYP determination for these schools based on that information.

Specifically, administrators were expected to:

1. Provide updated contact information.
2. Confirm that entities were in fact schools that were in open.
3. Confirm the grade configurations of the schools.
4. Confirm school type: e.g. Traditional, alternative, or K-2.
5. Notify ADE of any significant discrepancies in the data: e.g. the test scores of entire classes or grades miscoded to another school.
6. Reconcile the demographic data obtained through the AIMS test sheets to the demographic data obtained through the Student Accountability Information System (SAIS).

The web site application allowed schools to submit contact information, and information regarding school type and grade configurations. If large-scale corrections were required, schools were asked to download their test score data from the common logon web site and submit the corrected data to the Research and Policy section via its Achieve inbox (achieve@ade.az.gov).

IMPORTANT NOTE: The criteria used to select AIMS scores for evaluation of AYP differ from the criteria used to select scores for Arizona LEARNS. Indeed, the criteria differ among the separate components of the AYP evaluation. The criteria also differ from the scores provided to schools by the testing contractor, the scores publicly reported by ADE, and the scores available through the ADE's AIMS Wizard located at www.ade.az.gov/profile/publicview.

Complete verification of the data was a condition for being permitted to appeal a school's evaluation. That is, schools that did not verify their data were not allowed to appeal their AYP designation.

Timeline for School AYP

The following was the timeline for the school AYP evaluation process in 2003.

- August 14, 2003. Opening of data verification process.
- August 27, 2003. Closing of data verification process.

- August 29, 2003. Preliminary release of AYP determinations for Title 1 schools already in school improvement (priority Title 1). Opening of window for appeals submission for priority Title 1 schools.
- September 3, 2003. Closing of appeals window for priority Title 1 schools.
- September 15, 2003. Preliminary release of AYP determinations for all other schools. Opening of window for appeals submission for other schools.
- September 18, 2003. Closing of appeals window.
- September 29, 2003. Public release of AYP determinations for priority Title 1 schools.
- October 15, 2003. Public release of AYP determinations for all other schools.

Timeline for District AYP

The following was the timeline for the district AYP evaluation process in 2003/2004.

- December 10, 2003. Opening of data verification process.
- December 22, 2003. Closing of data verification process.
- January 5, 2004. Preliminary release of district AYP determinations. Opening of window for appeals submission for districts.
- January 16, 2004. Closing of appeals window for districts.
- January 19 – 28, 2004. Reviewing of submitted appeals.
- January 30, 2004. Public release of district AYP determinations.

Meeting the Annual Measurable Objectives for Proficiency

This section describes the calculation used to determine if schools met the annual measurable objectives (AMOs) for student proficiency in math and reading/language arts. NCLB requires that schools meet the AMOs set by the state in order to make AYP. A description of how the AMOs were set is given in section 2. Schools must meet the AMOs for each subjects/grade combination and all the applicable subgroups.

Calculation of Annual Measurable Objectives (AMOs)

The formula used to calculate the percentage of students passing is:

$$\text{Percent Pass} = \frac{\text{Number of students meeting/ex ceeding the standard on AIMS}}{\text{Number of students tested}}$$

This fraction was rounded to two digits, e.g.: .941=.94; .946=.95.

Confidence Intervals

To ensure that the decision regarding whether a school met the AMOs was statistically reliable and not overly influenced by random factors, the determination for meeting the AMOs was made employing a 99 percent (one-tailed) confidence interval. The confidence interval methodology is designed to ensure that 99 out of 100 times the confidence interval will contain a school's true performance level. If the AMO in question was below the upper bound of the confidence interval calculated for the school, the school was deemed to have met the standard.

Example. Twenty-nine percent of a school's third graders passed the AIMS mathematics test. The upper bound of the 99 percent confidence interval for this subject/grade combination for this school is calculated to be 35 percent. Since this is greater than the intermediate goal of 32 percent, the school is considered to have met the standard.

**Calculation of the
Upper Bound of the
Confidence Interval.**

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 99 percent confidence interval (UB99) is:

$$UB99 = p + 2.33\sqrt{p(1-p)/n}.$$

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 4.1.

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

Table 4.1. Rules for Determining UB99 for Small n and p.

<u>If $n \geq 30$ and $n < 35$, and</u>	<u>If $n \geq 40$ and $n < 45$, and</u>
$p \geq 0$ and $p < .05$, UB99=.16	$p \geq 0$ and $p < .05$, UB99=.13
$p \geq .05$ and $p < .10$, UB99=.25	$p \geq .05$ and $p < .10$, UB99=.22
$p \geq .10$ and $p < .15$, UB99=.33	$p \geq .10$ and $p < .15$, UB99=.28
$p \geq .15$ and $p < .20$, UB99=.38	$p \geq .15$ and $p < .20$, UB99=.35
$p \geq .20$ and $p < .25$, UB99=.45	
$p \geq .25$ and $p < .30$, UB99=.51	<u>If $n \geq 45$ and $n < 50$, and</u>
	$p \geq 0$ and $p < .05$, UB99=.12
<u>If $n \geq 35$ and $n < 40$, and</u>	$p \geq .05$ and $p < .10$, UB99=.21
$p \geq 0$ and $p < .05$, UB99=.15	$p \geq .10$ and $p < .15$, UB99=.27
$p \geq .05$ and $p < .10$, UB99=.24	$n \geq 50$ and $n < 55$, and
$p \geq .10$ and $p < .15$, UB99=.30	$p \geq 0$ and $p < .05$, UB99=.11
$p \geq .15$ and $p < .20$, UB99=.36	$p \geq .05$ and $p < .10$, UB99=.20
$p \geq .20$ and $p < .25$, UB99=.43	
	<u>If $n \geq 55$ and $n < 60$, and</u>
	$p = 0$, UB99=.10
	<u>If $n \geq 60$ and $n < 100$ and</u>
	$p = 0$, UB99=.09
	<u>If $n \geq 100$ and $n < 200$</u>
	<u>and $p = 0$, UB99=.06</u>
	<u>If $n \geq 200$ and $p = 0$,</u>
	UB99=0

Safe Harbor

Even if after calculating the confidence interval the percent of students proficient in a subgroup still fell short of the AMO, the group could still make AYP if its achievement indicators met certain safe harbor provisions. To make safe harbor a subgroup had to meet the following two-part test:

- Make a 10 percent improvement in the fraction of students meeting the standard (i.e. passing AIMS) from the previous year, and
- Have a 94 percent attendance rate for that group, or make a one-percentage point improvement in the group's attendance rate over the previous year. Since graduation rate data was not available for the applicable subgroups, attendance rate was used as the additional

safe-harbor indicator for high schools as well. The 2002 attendance rate data for the ethnic, low-income, special education, and English language learner subgroups was not available. Consequently, the 1 percent improvement criterion was not applied. That is, if a subgroup failed to meet the 94 percent standard, it did not make safe harbor.

Examples

1. In 2003, 20 percent of fifth graders in Gila Monster Elementary passed the AIMS reading test. The upper bound of the confidence interval was 25 percent, still below the annual measurable objective of 32 percent. However, in 2002, 18 percent of fifth graders passed the AIMS reading test, thus Gila Monster El. saw an improvement of over 10 percent in performance $[(20-18)/18 = 11 \text{ percent}]$. Furthermore, the attendance rate for Gila Monster's fifth grade was 96 percent, greater than the standard of 94 percent. So, Gila Monster's fifth graders make AYP in reading.
2. In 2003, 20 percent of eighth graders in Javalina Middle School passed the AIMS reading test. The upper bound of the confidence interval was 27 percent, still below the annual measurable objective of 31 percent. In 2002 19 percent of fifth graders passed the AIMS reading test, thus Javalina M.S. saw an improvement of only 5 percent in performance $[(20-19)/19 = 5 \text{ percent}]$. Even though the attendance rate for Javalina's eighth grade was 96 percent, greater than the standard of 94 percent, it fails to make the safe harbor provisions, and so does not make AYP in eighth grade reading.
3. In 2003, 30 percent of third graders in Gila Monster Elementary passed the AIMS reading test. The upper bound of the confidence interval was 40 percent, still below the annual measurable objective of 44 percent. However, in 2002 25 percent of third graders passed the AIMS reading test, thus Gila Monster El. saw an improvement of over 10 percent in performance $[(30-25)/25 = 20 \text{ percent}]$. However, the attendance rate for Gila Monster's third grade was 90 percent, less than the standard of 94 percent and identical to last year's attendance rate, so Gila Monster's third graders fail to make AYP in reading.

4. In 2003, 20 percent of third graders in Saguaro Elementary passed the AIMS reading test. The upper bound of the confidence interval was 25 percent, still below the annual measurable objective of 32 percent. However, in 2002 15 percent of fifth graders passed the AIMS reading test, thus Saguaro El. saw an improvement of over 10 percent in performance $[(20-15)/15 = 33 \text{ percent}]$. The attendance rate for Saguaro's third grade was 90 percent, less than the standard of 94 percent. However, in 2002 the attendance rate for Saguaro's third grade was 89 percent. Since Saguaro saw a 10 percent improvement in the fraction of third graders meeting the standard in math *and* a 1 percent improvement in the attendance rate for third graders, it meets the safe harbor provision for third grade math, and thus makes AYP.

Special rules

1. If a group had less than 30 students tested in 2002, it was automatically considered to have met the safe harbor provision. If a group had an ADM of less than 30 in 2003, then it was automatically considered to have met the attendance rate criterion for safe harbor.
2. Because ADE attendance data does not track the Title 1 status of students, attendance rate for low-income students was not included in the safe harbor determination. The safe harbor determination was based solely on the year-to-year improvement in percentage of students meeting the standards.

Data used

Students included in the calculation

Students were included in the calculation if they met the following criteria:

- Took either the AIMS or AIMS-A and received a score of FFB or above,
- Began the year in the same school. (Answered yes or left blank question number 3 on the AIMS demographic questionnaire; the field STARTYR in the AIMS data set = Y or blank.)

Applicable subgroups

Students in each of the following subgroups in every subject/grade combination were required to meet the annual measurable objective.

- ***Ethnicity.*** The fraction of students meeting the standard was calculated for each of the five ethnic groups—White, Black, Asian/Pacific Islander, American Indian, and Hispanic. The ethnicity of students tested was taken from the answer to the ethnic information question on the AIMS test sheet. Students with ethnicity of Other/blank were not included in the calculation.
- ***English Language Learners (ELL).*** ELL status was determined using the answer to question number 8 on the AIMS test sheet (ELLPROF in the ADE AIMS database.) Students with ELLPROF = 1 were considered English language learners. Students with ELLRPOF = 2 or blank were considered English proficient.
- ***Special Education Students.*** A student was identified as special education if she tested out of level (grade enrolled greater than test level in the AIMS file), took the AIMS-A, or was specified as a member of a special education program (question number 7 on the AIMS test sheet).
- ***Low income.*** A student was identified as being from a low-income family if the AIMS demographic information indicated she was a Title 1 student on AIMS testing documents. All students attending schools that had a school-wide Title 1 designation were identified as Title 1 students.

Special rule

Minimum group size. A group or subgroup was not evaluated if it had less than 30 test scores that met the selection criteria. Thirty is the sample size conventionally considered large enough to provide statistically meaningful results.

Meeting the Standard for Number of Students Tested

This section describes the calculation used to determine if a school had assessed 95 percent of its students. To make AYP, schools must test 95 percent of their students in reading and mathematics in all grades in which AIMS is administered, and must test 95 percent of their students in each applicable subgroup.

Calculation

The formula used to calculate the percentage of students tested is:

$$\text{Percent Tested} = \frac{\text{Number of students tested}}{\text{Number of students enrolled}}$$

The fraction of percent tested was rounded to two digits, e.g.: .941=.94; .946=.95.

Data used

Number of students tested. All students who took either the AIMS or AIMS-A and received a score of FFB or above. Students who received a score of Did Not Attempt (DNA) were excluded from the calculation.

Number of students enrolled. The denominator for the percent tested calculation was an unduplicated enrollment count at the school level. For grades 3, 5, or 8, enrollment used for all subjects was the first day of the week of testing: April 28, 2003—as reported to the Student Accountability Information System (SAIS). For grade 10, enrollment used was for the day the test was administered: April 28, 2003 for mathematics, and April 29, 2003 for reading.

The rules used to select students for enrollment were:

- Students were not selected if they were concurrently enrolled in more than one school on the relevant day.
- Students were not selected if they had more than one attribute on the relevant day. For example, a student could not have been assigned to more than one grade or more than one ethnicity on that particular day.

Applicable subgroups

- ***Ethnicity.*** Schools were required to have tested 95 percent of their students in the five ethnic groups—White, Black, Asian/Pacific Islander, American Indian, and Hispanic. The ethnicity of students tested was taken from the answer to the ethnic information question on the AIMS test document. The number of students enrolled for each ethnicity was taken from SAIS enrollment data for the selected days described in chapter 2.

Demographic information in the AIMS data files allowed students to be identified by the ethnicity of “Other” or have no ethnicity indicated, while SAIS permits students to be identified only by the five ethnic groups listed above. This raised the possibility of schools failing to meet the 95 percent threshold because of students identified as “Other” or blank. To resolve this problem, ADE used an algorithm to allocate other and unidentified students in the AIMS test files to one of the five ethnic groups. The rules followed by the algorithm were:

1. Students in the category of “Other” or blank were allocated to the five ethnic groups on a pro rata basis according to enrollment.
2. Students were not allocated to an ethnic category if that category already met the 95 percent threshold.
3. Students were not allocated to an ethnic category if the number of students in that category was already below the thirty-student threshold (as measured by enrollment) for minimum group size to be evaluated.
4. If the AIMS data had more students tested in an ethnic category than SAIS showed as enrolled, the extra students were allocated to other ethnic categories.
5. Fractions of students were allowed.

Table 5.1. Allocation of Students with “Other/Blank” Ethnicity

Ethnicity	Number tested (AIMS)	Number enrolled (SAIS)	Number tested after reallocation	Percent tested
White	42	40	40	100%
Black	30	40	38	95
American Indian	45	50	47.5	95
Asian	18	20	18	90
Hispanic	48	50	48	96
Other/blank	9		0.5	

The above table shows how the ADE allocated students in the AIMS data who had ethnicities of “Other” or blank. The table shows the number of test scores and enrollment for 5th grade math for Typical Elementary. In this case Typical tested 95 percent of its students for this subject/grade ($192 \div 200 = 96\%$) but could fail to meet the threshold for testing 95 percent of its African American and Native American students. Although the number of Asian students tested is below the 95 percent threshold, the enrollment for Asian students is below the 30-student threshold for a group to be evaluated. There are two more White students tested than enrolled. The algorithm takes the “extra” White students along with the nine Other/blank students and reallocates them to ethnic categories that fall below the 95 percent requirement. Eight students are allocated to the African American ethnic category and 2.5 students are allocated to the Native American ethnic category. The Asian category receives none since it is below the 30-student minimum group size.

English Language Learners (ELL). Schools were required to test 95 percent of their English language learners in order to make adequate yearly progress (AYP). The number of ELL students tested was the sum of all students identified as ELL students in the AIMS test file. ELL status was determined using the answer to question 8 on the AIMS test sheet (ELLPROF in the ADE AIMS database.) Students with ELLPROF = 1 were considered English language learners. Students with ELLRPOF = 2 or blank were considered English proficient. The number of ELL students enrolled was taken from SAIS. Students must have been enrolled as an English Language Learner on the relevant day in order to be included in the enrollment counts for this subgroup.

Special Education Students. Schools were required to test 95 percent of their special education students in order to make adequate yearly progress (AYP). The number of special education students tested was the sum of all students identified as special education in the AIMS test file. A student was identified as special

education if she tested out of level (grade enrolled greater than test level in the AIMS file), took the AIMS-A, or was specified as a member of a special education program (question 7 on the AIMS test sheet). The number of special education students enrolled was taken from the SAIS system. Students must have been enrolled in a special education program on the relevant day in order to be included in the enrollment counts for this subgroup.

Special rules

100 percent tested. If a school tested 100 percent of its students overall in a subject/grade category, it was assumed that it tested 100 percent for each subgroup in that subject/grade category, regardless of how students were labeled or mislabeled on testing documents.

Minimum group size. A group or subgroup was not evaluated if it had less than 30 students enrolled on the relevant day. Thirty is the sample size conventionally considered large enough to provide statistically meaningful results.

Other Indicators of School Performance

This section describes the calculation used to determine if a school met the other performance indicators for AYP. NCLB requires that schools be evaluated on a third performance indicator as well as percentage of students assessed and percentage of students proficient in the standard. The law requires that graduation rate be used for the third indicator for high schools, and gives states the discretion to choose the third indicator for elementary schools. Arizona has chosen the schoolwide attendance rate as the third indicator for elementary schools. To make AYP a high school must have a graduation rate of 71(70.5) percent; an elementary school must have an attendance rate of 94(93.5) percent.

Attendance Rate

Calculation. The formula used to calculate the attendance rate is:

$$\text{Schoolwide Attendance Rate} = \frac{\text{Average Daily Attendance}}{\text{Average Daily Membership}}$$

The attendance rate was rounded to two digits, e.g.:
.941=.94; .946=.95.

Data used. The average daily attendance (ADA) and average daily membership (ADM) for the 100-day counts for all grades offered by a school, except for pre-school, were used in the calculation.

Safe Harbor. If a school demonstrated a one-percentage point improvement in its attendance rate from the previous year, it was deemed to have met the performance standard. The growth rate was rounded to the nearest hundredth of a point, e.g. .009 = .01, .004=.00.

Example. Gila Monster Elementary had an attendance rate in 2003 of 92 percent, less than the standard of 94 percent. However, its 2002 attendance rate was 90 percent. Gila Monster El. demonstrated an improvement of 2 percentage points over the previous year, and so is deemed to have met the requirements for attendance rate.

Special rules. A school's attendance rate was not evaluated if it had an ADM of less than 30.

Graduation Rate

The graduation rate is an important complement to the AYP determination for high schools. Graduation rates indicate the success of students in meeting course requirements and achieving passing grades in subject areas not covered by the AIMS test. Graduation rates are used solely in the calculation of high school AYP. High school status was granted to any school that reported data in grade ten for each of the relevant school years (2000-2001, 2001-2002 and 2002-2003).

The Graduation Rate is a four-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 four years of beginning high school.

Calculation. The formula used to calculate the graduation rate is:

Graduation Rate	=	Number of Cohort members who graduated within four years				X 100
		Original Cohort Membership	+ Transfers In	- Transfers Out	- Deceased	

The graduation rate was rounded to three digits, e.g.:
.7045=.705; .7044=.704.

Data used. Federal requirements mandated that Arizona use the four-year graduation rate rather than the five-year rate used for Arizona LEARNS. The threshold graduation rate was for the cohort class of 2002, which represents the most recent graduation rate statistics. The graduation rate for the cohort class of 2001 was used for the determination of safe harbor.

Safe Harbor. If a school demonstrated a one-half-percentage point improvement in its graduation rate from the previous year, it was deemed to have met the performance standard. The growth rate was rounded to the nearest thousandth of a point, e.g. .0045 = .005, .0044=.004.

Example. Gila Monster High had a graduation rate in 2002 of 69.0 percent, less than the standard of 70.5 percent. However, its 2001 graduation rate was 67.0 percent. Gila Monster High demonstrated an improvement of 2 percentage points over the previous year, and so is deemed to have met the requirements for graduation rate.

Special rules. A school's graduation rate was not evaluated if it had an ADM of less than 30.

Calculation of Adequate Yearly Progress for K-2 Schools

The No Child Left Behind Act requires that a state evaluate *all* schools. Consequently, an alternative methodology for determining adequate yearly progress (AYP) had to be developed for schools that did not offer any of the grades in which AIMS is administered. In Arizona, this group consisted of schools that offered grades 2 and below.

The methodology for calculating AYP for K-2 schools was analogous to that used for other schools. K-2 schools were evaluated based on two criteria: whether they met the annual measurable objectives and attendance rate. Because AIMS is not administered in these schools, the AMO evaluation used student performance on the second grade Stanford 9 (SAT-9) tests for reading and mathematics. The percentage of students assessed was not used in determining AYP. As for other schools, the conjunctive model was used. A K-2 school had to meet both the AMO and the performance standard for attendance rate to make AYP.

Meeting the Annual Measurable Objectives for Proficiency for K-2 Schools

Setting of Baselines and Annual Measurable Objectives

The Arizona Department of Education established the starting points for K-2 schools in a manner analogous to that used for other schools. To determine the baselines for each subject, all K-2 schools in Arizona were ranked in descending order according to the percentage of students scoring in the 50th percentile or above in the Stanford 9. Then, cumulative enrollment was calculated adding upward from the bottom of the list of schools. The baseline was then set to be equal to the fraction of students for that subject in the school where the cumulative enrollment was equal to 20 percent of enrollment for K-2 schools. The data used for this calculation were the Stanford 9 results for the spring of 2002. The baselines calculated were:

- Math—20 percent.
- Reading—19 percent.

As with other schools, the baselines also served as the AMOS for the first year of AYP determinations.

Calculation

The calculation for determining if K-2 schools met the AMOs was equivalent to that used for other schools. The formula used to calculate the percentage of students who rank in at least the 50th percentile is:

$$\text{Percent} \geq 50\text{th percentile} = \frac{\text{Number of students} \geq 50\text{th percentile}}{\text{Number of students tested}}$$

This fraction was rounded to two digits, e.g.: .941=.94; .946=.95.

Confidence Intervals. The confidence interval methodology for AMO determination as described in chapter 4 for other schools was applied to AMO determination for K-2 schools.

Safe Harbor. The safe harbor methodology for AMO determination as described in chapter 5, section 1.3 for other schools was applied to AMO determination for K-2 schools.

Data used. Students were included in the calculation if they were in the second grade, took the Stanford 9, and began the year in the same school. (The field STARTYR in the SAT-9 data set = Y or blank.)

Applicable subgroups. The AMO determination for K-2 schools was done only for the whole of grade 2 for reading and mathematics. There were no separate AMO determinations for subgroups—ethnicity, ELL status, etc.—as was done for other grades.

Minimum group size. A subject group was not evaluated if it had less than 30 test scores that met the selection criteria.

Attendance Criteria for K-2 Schools

To make AYP a school had to meet the schoolwide attendance goal of 94 percent. The calculation of the attendance rate for K-2 schools is the same as that for other schools as described in chapter 7.

Summary of AYP Determination for K-2 Schools

Table 7.1 below summarizes the AYP determination process for K-2 schools. The conjunctive model requires that all cells in the table have a “yes” for a school to make AYP.

Table 7.1. AYP Determination for K-2 Schools

Met AMO for SAT-9 Reading?	Met AMO for SAT-9 Math?	Met Goal for Attendance Rate?
Yes/No	Yes/No	Yes/No

Determining Adequate Yearly Progress for School Districts and Charter Holders

The No Child Left Behind Act requires that local education agencies (LEAs), districts and charter holders, be evaluated for Adequate Yearly Progress. The method for determining AYP (AYP) for districts² is analogous to that used for schools with data being aggregated to the district level as if a district were one large school. The details of the AYP calculation for districts are nearly identical to that for schools.

- Districts are evaluated for percentage of students passing AIMS, percentage of students assessed, and a third indicator.
- Annual Measurable Objectives (AMOs) and the performance goals for percentage of students assessed, attendance rate, and graduation rate are the same for districts as they are for schools.
- The applicable subgroups for AYP evaluation are the same for districts as they are for schools.
- Confidence intervals, safe harbor provisions, and minimum group size requirements are applied to district AYP using the same methodology and parameters as for school AYP.
- District AYP uses a conjunctive model. To make AYP, a district must meet all the performance standards for all subjects, grades, and subgroups that are applicable.

Differences between District and School AYP Evaluation Methods

There are four differences between the AYP evaluation method used for districts and that used for schools.

1. ***Measure of student mobility.*** NCLB requires that students mobile with respect to an entity are not included in the AMO part of the AYP evaluation. For a school, this meant

² All statements in this section apply to both districts and charter holders. For the sake of brevity, we use “district” to refer to both types of entities/LEAs.

excluding students who did not start the year at that school. District level mobility was determined by the answer to the question “Number of Years in District” (question 4 on the AIMS document.) Students who answered “Less than 1 year” were excluded from the sample for AMO determination for districts.

2. ***Limit on the number of students with alternative assessment who count toward meeting the proficiency standard.*** NCLB mandates that the number of students who take an alternative assessment who count as being proficient may not be greater than 1 percent of the total number enrolled. This limit only applies per subject for an entire grade level but not for any subgroups. For the 2003 AYP determination, only students who took the AIMS-A are considered to have taken an alternate assessment.

Example. Gila Monster Elementary has 100 students who have taken AIMS for third grade reading; 72 students, including two who have taken the AIMS-A, passed the test. To comply with the required limit, only one AIMS-A student may count as proficient (1 is 1 percent of 100). Therefore, only 71 students count as proficient when determining if Gila Monster El. met the AMO for third grade reading. One of the students who took the AIMS-A is white, and the other is Hispanic. Both students count as proficient when determining if the white and Hispanic subgroups at Gila Monster El. met the AMO for third grade reading.

3. ***Graduation/Attendance Rates.*** Graduation rate was used as the third indicator required by NCLB for unified and high school districts. Attendance rate was used for elementary districts.
4. ***Consistency Check for Number Tested.*** Subgroups for a subject/grade were deemed to have met the 95 percent goal for percentage of students tested if the data for the entire subject/grade implied that missing the goal was mathematically impossible.

Example. Data for Gila Monster Elementary show 98 students have taken AIMS for fifth grade reading and 100 fifth grade students enrolled. However, data submitted to ADE show 50 students enrolled in special education programs but only 45 students with special education status indicated on their AIMS demographic information. Since only two students did not take the test, it is mathematically impossible for Gila Monster to have tested less than 48 of its special education students—there must be special education students in the data who are not labeled as such. Since $48/50 = .96$

> .95, Gila Monster meets the 95 percent assessed threshold for its special education students.

The Adequate Yearly Progress Appeals Process

The Adequate Yearly Progress (AYP) Appeals Process developed by the Arizona Department of Education (ADE) provided schools and districts the opportunity to appeal 2002-2003 AYP determinations. In accordance with Title I, Section 1116 of the No Child Left Behind Act of 2001 (NCLB), the ADE allowed schools and districts to appeal their respective AYP determinations for statistical and/or substantive reasons.

Step One: Data Verification

The first step in completing the AYP Appeals Process required *all* schools and districts 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 included two distinct types of data, both of which had to have been verified by administrators: school/district information and data information.

Step Two: Appeal Application

School/district administrators choosing to appeal an AYP determination had to complete the AYP Appeal Application, which was accessible via the Common Logon during the specified appeal timeframe in order to indicate the exact issue(s) of appeal. Appeals were only accepted through the website application. Appeals sent to ADE via email, fax, or mail/delivery were not accepted.

Schools/districts were able to appeal AYP determinations in two categories: data (statistical) and non-data (substantive) reasons - schools/districts were not limited to one category and were able to appeal in both if necessary.

Statistical appeals

Schools/districts that appealed based on statistical arguments could have argued one or more of the following:

1. Calculation of 95% tested (This argument was valid only for high schools which had no opportunity to conduct make-up examinations. Elementary schools had two a week test window in which students could make up exams and therefore were unable to appeal the 95% tested objective).
2. Calculation of AMO due to the inclusion of invalid scores. This appeal could have included appealing the inclusion of English Language Learners not yet considered English proficient and/or the inclusion of Special Education students with modifications.
3. Calculation of additional indicators: attendance (elementary schools) and graduation rate (high schools).

Substantive appeals.

Schools/districts that appealed based on substantive arguments could have argued that mitigating circumstances, outside of the school/district's control, negatively impacted the quantity or quality of test data. This included circumstances that affected test conditions, test scores, percent tested, and/or additional indicators (attendance – elementary schools; graduation rate – high schools).

Important notes for the appeals process. Administrators that chose to appeal an AYP determination had to clearly articulate the issue(s) they believed merited an appeal through the AYP Appeal Application. School/district administrators had to submit evidence that the issue(s) they believed merited an appeal directly resulted in a *significant* decrease in student academic achievement as demonstrated on AIMS and/or a decrease in student participation during the administration of AIMS. The evidence had to have been submitted to ADE at the time the appeal was submitted. Failure to provide this evidence resulted in the appeal not being granted. Evidence that was submitted after the appeal deadline closed was not considered. Once appeals were submitted through the Common Logon, the school/district/charter received an email verifying that the appeal had been received.

The ADE, if necessary, requested that a school/district administrator provide additional information/evidence to assist in the appeals process. Only those requests for additional information that were provided during the specific timeframe were 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 was not accepted.

Step Three: Appeal Resolution

After all appeals were submitted and the appeal window closed, the ADE began to process the appeals. Appeals were addressed categorically, not necessarily in the order received so the fact that a school/district submitted its appeal during the first day of the appeal window did not mean it would necessarily receive a decision first during the resolution process. The appeal process was implemented in three stages.

Stage 1 – Statistical appeals processed

All appeals of a statistical nature based on data discrepancies were reviewed.

Appeals that challenged the calculation of 95% tested and/or attendance/graduation were processed by verifying that the information taken from the Student Accountability Information System (SAIS) and the numbers used in the calculations were true and accurate. **Note: It was the responsibility of the school district/charter to ensure that the information reported to SAIS was accurate and the district's numbers matched those reported to ADE.**

Appeals that challenged the calculation of AMO due to the inclusion of invalid scores were evaluated using three modified data sets. The first was a data set that excluded the English Language Learners (ELL), not yet English Proficient, from the calculations to determine if the school would in fact have met the AMO objective had it not been for the inclusion of invalid test scores. The second modified data set excluded Special Education students who received non-standard accommodations from the calculations to determine if AYP was met without the use of invalid test scores. The third modified data set excluded both invalid scores of the ELL and Special Education students from the calculations.

All statistical appeals needed to be supported with compelling evidence. While some statistical arguments seemed compelling and may even have seemed obvious to the school/district submitting the appeal, evidence needed to be provided to support all appeals. For example, if the percent of students tested objective was not met in the ELL subgroup because of the miscoding of ELL students on the test, evidence of that miscoding needed to be provided. Simply stating, “ELL numbers at ADE don’t match the district’s” was not compelling; ADE needed to know *why* the numbers were different; meaning that particular students needed to be identified as miscoded. **Note: Some schools/districts, when providing information in the appeals mentioned specific details about students such as names, id#, ethnicity, and specific attendance/student record information which violated**

guidelines set forth by FERPA. Schools/districts were strongly encouraged to follow the FERPA guidelines in the future. When referring to students in appeals, identifying student information such as name, id#, etc. is not to be submitted with the appeal. Instead, students are to be referred to as student #1, student #2, etc.

Stage 2—Substantive appeals processed.

Substantive appeals were resolved through a committee process. All committee members were nominated and represented a diverse background to ensure that appeals were considered from multiple perspectives. Among those perspectives were those of principals, teachers, school administrators, department administrators, researchers, and Title I representatives.

Once the committee was assembled, the appeals were evaluated utilizing an appeals rubric that evaluated the significance of the argument presented and how the circumstances presented in the argument affected the school's performance. The committee based their decisions on the following criteria:

1. ***Was the circumstance that affected the school outside the school's control?*** If the school was negligent in its test administration and/or data collection, the appeal was not deemed relevant and the appeal was not considered. For example, if the school forgot to test a certain class in a certain grade and remembered after the test window closed, that circumstance was not outside of the school's control and therefore not a valid argument for appeal. Conversely, if the school did test everyone and some of the tests were lost by the testing contractor, then that situation would be outside of the school's control.
2. ***Did the special circumstance actually have an impact on the school's performance?*** Not all circumstances at a school impact test data. For example, if the school had a long-lasting construction project on campus, did the actual test environment suffer during the test week? How? Or if a teacher left mid-year, did the learning environment suffer? How? If the answers to these questions did not show adequate impact on the test environment, then the event most likely did not affect the actual performance at the school. Conversely, if it could be demonstrated that the event did influence the performance of the students then that was a valid argument.
3. ***Is this problem one that was recurring and likely to happen in the future?*** Appeals made based on

policy(s) at the school that impact test collection/data results, which contradicted ADE/NCLB policy(s), were not accepted. For example, if the school's enrollment numbers were inflated and they failed to make the 95% test objective because the school did not withdraw students after the 10th consecutive missed day, opting instead to withdraw students after the 20th day, they could not appeal the enrollment figures because they were not following ADE policy.

4. ***Is the problem eligible for appeal?*** Arguments that target NCLB regulations and ADE policy were not valid. For example, schools could not argue that the 95 % tested threshold be lowered for their school or that certain subgroups be excluded from the requirements. **Note: As mentioned in the statistical section, certain data related to specific populations could be statistically excluded from the analysis of AMO (such as ELL and Special Education students) after the initial calculations were computed, but they could not be exempted from the 95% tested and/or attendance requirements. All students and all subgroups had to be tested and attend school.**
5. ***Did the school provide compelling evidence of the circumstance?*** While some situations may have seemed compelling and may even have seemed obvious to the school submitting the appeal, compelling evidence of impact needed to be provided to support all substantive appeals. For example, if percent of students tested objective was not met, specific details to support the claim needed to be provided with the appeal at the time it was submitted. Simply stating "Students were absent and unable to make up the test" is not compelling; the committee needed to know *why* the students were unable to make up the test such as being extremely ill, suspended, incarcerated, or dealing with a family emergency for the entire test window. **Note: Some schools/districts, when providing information in the appeals mentioned specific details about students such as names, id#, ethnicity, and specific attendance/student record information which violated guidelines set forth by FERPA. Schools/districts were strongly encouraged to follow the FERPA guidelines in the future. When referring to students in appeals, identifying student information such as name, id#, etc. is not to be submitted with the appeal. Instead, students are to be referred to as student #1, student #2, etc.**

Appeal resolution notes. If the school submitted both a statistical and substantive appeal, the statistical appeal was evaluated first. Only after the statistical arguments had been exhausted was the appeal sent to the substantive committee for evaluation.

The appeals should have addressed the appropriate category. Some appeals addressed issues that were not relevant to the school's AYP designation. For example, some schools presented arguments that invalid assessment results for ELL and Special Education students should be taken out of the analysis (an appeal for the AMO category) yet their school failed to meet the 95% tested objective. Therefore, the appeal was not relevant and denied. Since all subgroups had to be tested at the 95% level, the school could not argue that invalid assessment results for ELL and SPED students be excluded from this requirement, thus affecting the testing requirement. Schools/districts in these situations should have addressed in their appeal rationales as to *why* the school did not test 95% of the students.

Schools/districts needed to be certain that if they failed in multiple categories that the appeal addressed all deficiencies. Some appeals submitted addressed only one deficient area. While those arguments were compelling in that criterion, the overall AYP designation for the school could not change because additional deficiencies related to the AYP designation were not addressed in the appeal.

Schools/districts needed to be certain to provide all relevant information/support when submitting the appeal; late information to support the appeal was not accepted (unless ADE specifically asked for additional information as noted above).

Stage 3 – Notification of Results Sent to Schools/districts

Once all appeals were resolved, notifications went to the schools/districts that filed appeals. The contact person of record for the school received an email from Achieve with directions as to how to access appeal information via the Common Logon when the appeal had been processed. Schools/districts were notified before the final public release of the AYP determinations as to the outcome of the appeal process. All appeals were final.