Proposal for a Growth Model to Evaluate Adequate Yearly Progress for Schools and Districts

Tom Horne
Superintendent of Public Instruction
Arizona Department of Education

Submitted by the Arizona Department of Education, February 17, 2006.
Revised March 17, 2006

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

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.

Notes on the Revised Proposal

In response to comments by U.S. Department of Education staff, the Arizona Department of Education (ADE) has significantly altered the growth model it originally proposed on February 17. The changes are:

- **New time horizon to attain proficiency.** The original model expected all students to achieve proficiency by the time they reached eighth grade. The amended model presented here expects students to achieve proficiency within three years or by eighth grade, whichever comes first.

- **Growth calculation for <u>all</u> students.** The original model employed growth as an additional safe harbor for groups that failed to meet the status goals. The amended model presented here is used for all students.

Additional changes, explanations, and examples have been added to the proposal in response to the March 10 letter from Henry L. Johnson to the ADE. These changes are cross-referenced and summarized in an appendix.

Summary

The Arizona Department of Education (ADE) is submitting this proposal requesting the opportunity to incorporate a model of individual student growth into the adequate yearly progress (AYP) evaluations of districts and schools. If approved, the model can be implemented in the AYP evaluations for the 2005-06 school year.

The purpose of including a growth model in AYP evaluations is to recognize the success of schools in improving student achievement, even though that achievement may fall short of proficiency. With recognition comes the incentive for schools to concentrate further effort toward assisting students who fall short of proficiency. Increased attention toward the lowest performing students will keep them from falling further behind their peers. Educators in successive grades will find more of their students within striking distance of proficiency allowing the achievement gap to be closed more quickly.

Arizona has publicly reported its measure of student growth for schools, the Measure of Academic Progress (MAP), for seven years. It is included on school report cards and used by third party information sources like GreatSchools.net. It is also used as a performance pay indicator for teachers. Parents, the public, school administrators, and teachers are all familiar with the specific measure and the concept of growth in general. With increasing emphasis on state standards, the old MAP based on the norm-referenced test was starting to be deemphasized. With the introduction of a new growth measure based on state standards, we expect additional use of the growth measure to develop state and local policy and treat the instructional needs of individual students. Including growth as part of the AYP determination will bring the various accountability measures in the state into greater congruence, increasing the effectiveness and validity of the state's overall school accountability system.

Components in Place
Arizona has in place all the necessary elements for a student-level growth model:

- The state's Student Accountability Information System (SAIS) has been in place since 2002. Each student of a public (including charter) school in the state is assigned a unique SAIS ID that follows her through her career from kindergarten to high school graduation. The SAIS system also contains demographic and program membership information for all students.

- All students are tested in both reading and math in grades three through eight. The state's assessment system has been submitted for NCLB peer review. The AIMS test is vertically scaled for grades three through eight. Arizona has vertically aligned, articulated standards for all grades.

- Arizona tested in all grades operationally in the 2004-05 school year, so field test data will not be needed to implement the model.

The growth model submitted in this proposal will be calculated for all subgroups of students and used as a parallel AYP evaluation for both schools and districts. Schools and districts that make AYP in the growth model will be deemed to have made AYP and not be designated for school improvement.

The model looks at progress individual students make toward proficiency from one year to the next. The goal is proficiency within three years or by the eighth grade, whichever comes first. Annual growth targets are set that measure each student's progress toward that goal. Students are deemed to have made sufficient progress if they meet their annual growth target. Scores for individual students are aggregated by the relevant subgroups. If the percentage of students in a subgroup that meets the target for growth is equal to or greater than the AMO, then the subgroup is considered to have met AYP.

Details of Model
The proposed model calculates the annual progress made by each student toward proficiency in state standards. Progress is measured against the goal of proficiency within three years or by eighth grade, whichever comes first. This represents a horizon of no more than three years, and less for students in higher grades. The steps for calculating it are:

1. For each student an **annual growth target** is set for each subject: reading and math. The growth target is how much improvement measured by scale score points a student would have to make over her previous year's score in equal intervals in order to achieve proficiency by the eighth grade. The growth target is set by subtracting the student's current year scale score from the scale score for proficiency three grades higher and dividing by the number of remaining grades. Since Arizona's vertically scaled series of tests ends at eighth grade, the scale score for eighth grade will be used for grades five through seven. The targets are

rounded to the nearest whole number. Demographic factors are not used to set the target.

Example: A student scores 402 on the 3rd grade math test in 2005. The passing score on the 6th grade math test is 496. The student's math score must improve 31 points each year—(496 – 402)/(6-3) = 94/3 = 31—for him to reach proficiency by 6th grade.

Example. A student scores 469 on the 6th grade reading test in 2005. The passing score on the 8th grade reading test is 499. The student's reading score must improve 15 points each year—(499-469)/(8-6) = 30/2 = 15—for her to reach proficiency by 8th grade.

2. For each student **actual growth** is measured for each subject.

Example: A student scores 402 on the 3rd grade math test in 2005 and 442 on the 4th grade math test in 2006. The student's actual growth is 40 points (442 – 402).

3. For each student **actual growth** is compared to the **annual growth target**. If actual growth is greater than or equal to the annual growth target, the student is deemed to have made adequate growth.

Example. A student's growth target is 31. The actual growth shown is 40. The student is considered to have made adequate growth.

Students who pass in the current year are automatically considered to have met their growth target regardless of how much actual growth they make.

Example 1. A student scores 457 on the third grade math test. Thus, the student's growth target is 13. The student scores a 460 on the fourth grade test—a passing score. Even though the student only grew by 3 points, he is proficient in the current year and thus is considered to have met his growth target.

Example 2. A student scores 457 on the third grade math test. Thus, the student's growth target is 13. The student scores a 450 on the fourth grade test—a passing score. Even though the student's score *decreased* by 7 points, he is proficient in the current year and thus is considered to have met his growth target.

This requirement makes the aggregate growth measure a status-plus measure of performance. The growth measure for a subgroup is all students who are proficient plus the number who met their growth targets divided by the total number of students included in the analysis.

The percentage of students meeting their annual growth target for a group is compared to the annual measurable objective (AMO) for that subject and grade. If the percentage is above the AMO, then the group is deemed to have made AYP. Because the AMOs

increase over time, the growth targets increase over time as well. In 2014, all students will either have to be proficient or be on the road to achieving proficiency within three years or less. This is consistent with the NCLB goal.

Example. This example shows how the growth measure can be calculated for a subgroup. This is a subgroup of 4th graders, and we are examining if it met the AMO for 4th grade math. The passing score for 4th grade is 448.

Student ID	Actual Score 3rd Grade	Passing Score 6th Grade	Growth Target	Actual Score 4th Grade	Actual Growth	Met Growth Target
1	440	496	19	450	10	Y
2	400	496	32	440	40	Y
3	450	496	15	440	-10	N
4	390	496	35	430	40	Y
5	445	496	17	440	-5	N
6	420	496	25	440	20	N
7	460	496	12	465	5	Y
8	435	496	20	454	19	Y
9	425	496	24	440	15	N
10	415	496	27	445	30	Y

The AMO for fourth grade math is 54 percent of students passing. The percentage of students meeting their growth target is 60 percent, above 54 percent, so the group will be deemed to have made AYP. A 99 percent confidence interval will be used for this determination. If the upper bound of the confidence interval is greater than 54 percent, then the subgroup will be deemed to have made sufficient growth.

Conclusion

Arizona is well qualified to incorporate a growth model into its AYP evaluations for schools and districts.

1. Arizona has in place all the necessary elements for a growth model. The SAIS database assigns each student a unique identifier that tracks students across grades, schools, and districts. It also holds student demographic information, allowing growth to be measured for all the subgroups required by NCLB. The AIMS test has been administered operationally in all grades three through eight, is vertically scaled, and will remain stable for the next four years.

2. Arizona has experience with growth measures. Student growth has been a part of Arizona's school accountability system since the previous century. Growth measures have been reported publicly and used by administrators. We will continue to do so regardless of the status of this proposal.

3. Arizona's growth model provides a more valid measure of school effectiveness than the current status-only evaluation. It recognizes and gives schools the incentive to improve student performance at all levels.

4. Arizona sets ambitious growth targets for all students in all schools. Expected growth is not dependent on students' demographic characteristics. It sets the ultimate goal of all students being proficient or on track to proficiency by 2014.

If approved, the Arizona growth model would lessen the possibility that students would fall behind. It will increase the number of students who move on to higher grades that are ready to learn. Consequently, adding the growth model to the AYP determination will help close the achievement gap more rapidly.

Table of Contents

1 Overview 8
1.1 A Growth Model Will Improve School Accountability in Arizona 8
1.2 Arizona Meets the New Equation for Flexibility 9
1.3 Core Elements in Place for Growth Model 10
1.4 Development and Implementation 10
2 The Proposed Model 12
2.1 The Current System for AYP Determination in Arizona 12
2.2 Incorporating the Growth Model into the AYP Evaluation 13
2.3 Calculation of Growth 14
2.4 Details of the Calculation 16
2.5 Alternate Tests 20
2.6 Inclusion Rules and Missing Data 20
2.7 K-2 Schools, High Schools, and Third Grade 21
3 Arizona's Growth Model Meets the Seven Core Principles 22
3.1 100 percent proficiency by 2014 and Incorporating Decisions about Student Growth into School Accountability 22
3.2 Establishing Appropriate Growth Targets at the Student Level 24
3.3 Accountability for Reading/Language Arts and Mathematics Separately 25
3.4 Inclusion of All Students 25
3.5 State Assessment System and Methodology 26
3.6 Tracking Student Progress 29
3.7 Participation Rates and Additional Academic Indicators 32
4 Conclusion 33

1 Overview

The Arizona Department of Education (ADE) is submitting this proposal requesting the opportunity to incorporate a model of individual student growth into the adequate yearly progress evaluations (AYP) of districts and schools.

Arizona has the components in place—a statewide student data system, vertically scaled standards and tests—to implement a growth model. Furthermore, Arizona meets the criteria set out in the New Equation for providing flexibility in implementing NCLB.

Arizona's first model of school accountability, introduced in 1998-99, was a growth model: the Measure of Academic Progress (MAP). The MAP was based on growth measured by stanine to stanine movement on the state's norm-referenced test. At the time it was the only test administered statewide, and until the 2004-05 school year, the only test administered in all grades three through eight. The ADE has reported MAP results by school since 1999. Now that the state's criterion-referenced test, the AIMS, is given at every grade three through eight, the ADE intends to introduce a new growth measure based on growth measured against state standards.

If approved, the model can be implemented in the AYP evaluations for the 2005-06 school year.

1.1 A Growth Model Will Improve School Accountability in Arizona

The purpose of including a growth model in AYP evaluations is to recognize the success of schools in improving student achievement, even though that achievement may fall short of proficiency. With recognition comes the incentive for schools to concentrate further effort toward assisting students who fall short of proficiency. Increased attention toward the lowest performing students will keep them from falling further behind their peers. Educators in successive grades will find more of their students within striking distance of proficiency. This will allow the achievement gap to be closed more quickly.

Including a growth measure in the state's AYP determinations will raise expectations by making explicit the goal that ***all*** students in Arizona are to show improvement. It is widely recognized that a simple status model of accountability that only looks at percent proficient gives schools the incentive to focus on so-called "bubble" students—those who fall just short of proficient. Arizona addressed this problem when designing its state accountability system AZ LEARNS. AZ LEARNS credits schools for improvement across all levels of student achievement. The Measure of Academic Progress has also been included in AZ LEARNS. By expecting improvement from students at all levels, schools will not just focus on the bubble students and allow the lowest performing to languish.

A growth model will make the state's AYP determinations a more valid measure of school effectiveness. Arizona currently bases its AYP determinations on percent proficient with the safe harbor provided for in legislation. Because status models do not account for improvement shown by all students, they can only imperfectly distinguish

schools that are truly effective. Schools that are highly successful in raising student achievement can be falsely identified as ineffective and in need of improvement if a proficiency target is not met. Using growth as a component in the AYP determination will help ADE in identifying those schools where student achievement is truly stagnant. This will allow ADE to focus resources and assistance in helping those schools.

Arizona has publicly reported its measure of student growth for schools, the MAP, for seven years. It is included on school report cards and used by third party information sources like GreatSchools.net. It is also used as performance pay indicator for teachers. Parents, the public, school administrators, and teachers are all familiar with the specific measure and the concept of growth in general. With increasing emphasis on state standards, the old MAP based on the norm-referenced test was starting to be deemphasized. With the introduction of a new growth measure based on state standards, we expect additional use of the growth measure to develop state and local policy and treat the instructional needs of individual students. Including growth as part of the AYP determination will bring the various accountability measures in the state into greater congruence, increasing the effectiveness and validity of the state's overall school accountability system.

1.2 Arizona Meets the New Equation for Flexibility

Arizona has a history of commitment to educational reform and accountability. Initiatives that have been in place since the 1990s include school report cards, ratings of school effectiveness, statewide assessment of proficiency in state standards, proficiency as a graduation requirement, and school choice through charter schools and open enrollment.

Our state accountability system, AZ LEARNS, allows schools only three years to move out of the lowest performance category before they face state intervention. The state has already put into place turnaround principals in several of its failing schools as well as ATLAS teachers who are teacher leaders/mentors that model best practices. Before this happens, however, each year a school identified as underperforming must still notify the public of its underperforming status and file a school improvement plan. Meanwhile the state provides assistance through solutions teams.

Arizona is meeting the key NCLB principles for school accountability:

- It has in place an accountability system for all schools and districts that tests in all grades three through eight, and high school. The system holds entities accountable for subgroup performance in reading and mathematics.

- School performance measures are made available to parents both in paper and electronically through school, district, and state report cards. The report cards include student proficiency by subgroup, teacher experience; the number of classes taught by highly qualified teachers; and attendance, promotion, dropout, and graduation rates.

- Choice and supplemental services are available for students in underperforming schools, as well as the options for school choice available for all students in the state. Schools in school improvement receive visits from solutions teams made up of top educators. ADE also assigns a coach to monitor the implementation of their school improvement plan. Coaches also evaluate compliance with federal and state accountability. A school in improvement uses Arizona's Standards and Rubrics for School Improvement to identify the strengths and limitations of the school's overall program. The Standards and Rubrics are anchored in the scientifically research-based principles and indicators that consistently distinguish top-performing schools. ADE also provides a web resource guide to support the Standards and Rubrics.

- Arizona has issued standards for qualified teachers and reports to parents the qualifications of teachers at their children's schools. Arizona has in place a career ladder program that encourages performance-based compensation plans to provide incentives to teachers to make career advancements without leaving the classroom or the profession. Arizona requires a proficiency assessment for most certifications, and will require a performance assessment for future teacher certifications. Arizona has also established a classroom site fund in which teachers are compensated for student performance and based on school improvement goals identified in the school's improvement plan.

1.3 Core Elements in Place for Growth Model

Arizona has in place all the necessary elements for a student-level growth model:

- The state's Student Accountability Information System (SAIS) has been in place since 2002. Each student of a public (including charter) school in the state is assigned a unique SAIS ID that follows her through her career from kindergarten to high school graduation. The SAIS system also contains demographic and program membership information for all students.

- All students are tested in both reading and math in grades three through eight. The state's assessment system has been submitted for NCLB peer review. The AIMS test is vertically scaled for grades three through eight. Arizona has vertically aligned, articulated standards for all grades.

- Arizona tested in all grades operationally in the 2004-05 school year, so field test data will not be needed to implement the model.

1.4 Development and Implementation

The Arizona Department of Education developed the growth model presented here with comments and suggestions from:

- The State Assessment and Accountability Advisory Committee (SAAAC), a group of 40 members representing districts, charter schools, and other stakeholders statewide.

- The National Assessment and Accountability Advisory Committee (NAAAC), a group of six university scholars with expertise in assessment and accountability.

- Arizona's Title I Committee of Practitioners (COP)

- The Accountability Systems and Reporting working group of CCSSO.

In general it was the recommendation of the groups that the growth model used by ADE be fair, simple, and consistent across accountability models.

The model also holds to the criteria set down by the U.S. ED, especially:

1. Demographic characteristics of students are not used to set different growth targets.

2. Growth is measured toward the goal of proficiency.

3. Growth by students above proficient does not compensate for the growth of less than proficient students.

Aggregate results and the results for subgroups will be reported publicly and placed on school, district, and state report cards. The proposed measure was chosen to be straightforward, intuitive, and easily understood by principals, teachers, and parents. As always, ADE is ready to offer supplementary material and training to further understanding.

It is the intent of the ADE to continue to monitor the validity of the growth measure. The growth measure proposed here will be used as a safe harbor appended to the current status model. This will enable an easy comparison of the results produced with the growth model to those that would have been provided under the current model. The ADE will provide the necessary information for the evaluation of the program by U.S. ED.

2 The Proposed Model

The growth model submitted in this proposal will be calculated for all subgroups for both schools and districts, parallel to the traditional AYP evaluation. The model looks at progress individual students make toward proficiency from one year to the next. The goal is proficiency within three years for grades three and four, or by the eighth grade for grades five through seven. Annual growth targets are set that measure each student's progress toward that goal. Students are deemed to have made sufficient progress if they meet the annual growth target. Scores for individual students are aggregated by the relevant subgroups. If the percentage of students in a subgroup that meets the target for growth is equal to or greater than the annual measurable objective (AMO), then the subgroup is considered to have met AYP.

2.1 The Current System for AYP Determination in Arizona

The present method for AYP evaluation in Arizona examines student performance in reading and math for each grade three through eight, and high school (10th grade). There are separate annual measurable objectives (AMOs) for each grade. For the past three years, elementary and middle schools have only been evaluated for grades three, five, and eight. The approved AMOs for these grades are given in table 1. The 2004-05 school year was the first year the state's criterion referenced test, the AIMS, was administered operationally in grades four, six, and seven. However, the results for these grades will not be incorporated into the AYP evaluation until the 2005-06 school year. (The proposed AMOs for these grades are given in table 2.)

Table 1. Approved AMOs

Grade 3	Reading AMO	Math AMO
2005-07	53.3 %	43.3 %
2008-10	62.6 %	54.6 %
2011	71.9 %	65.9 %
2012	81.2 %	77.2 %
2013	90.5 %	88.5 %
2014	100 %	100 %
Grade 5		
2005-07	43.3 %	33.3 %
2008-10	54.6 %	46.6 %
2011	65.9 %	59.9 %
2012	77.2 %	73.2 %
2013	88.5 %	86.5 %
2014	100 %	100 %
Grade 8		
2005-07	42.5 %	22.5 %
2008-10	54 %	38 %
2011	65.5 %	53.5 %
2012	77 %	69 %
2013	88.5 %	84.5 %
2014	100 %	100 %

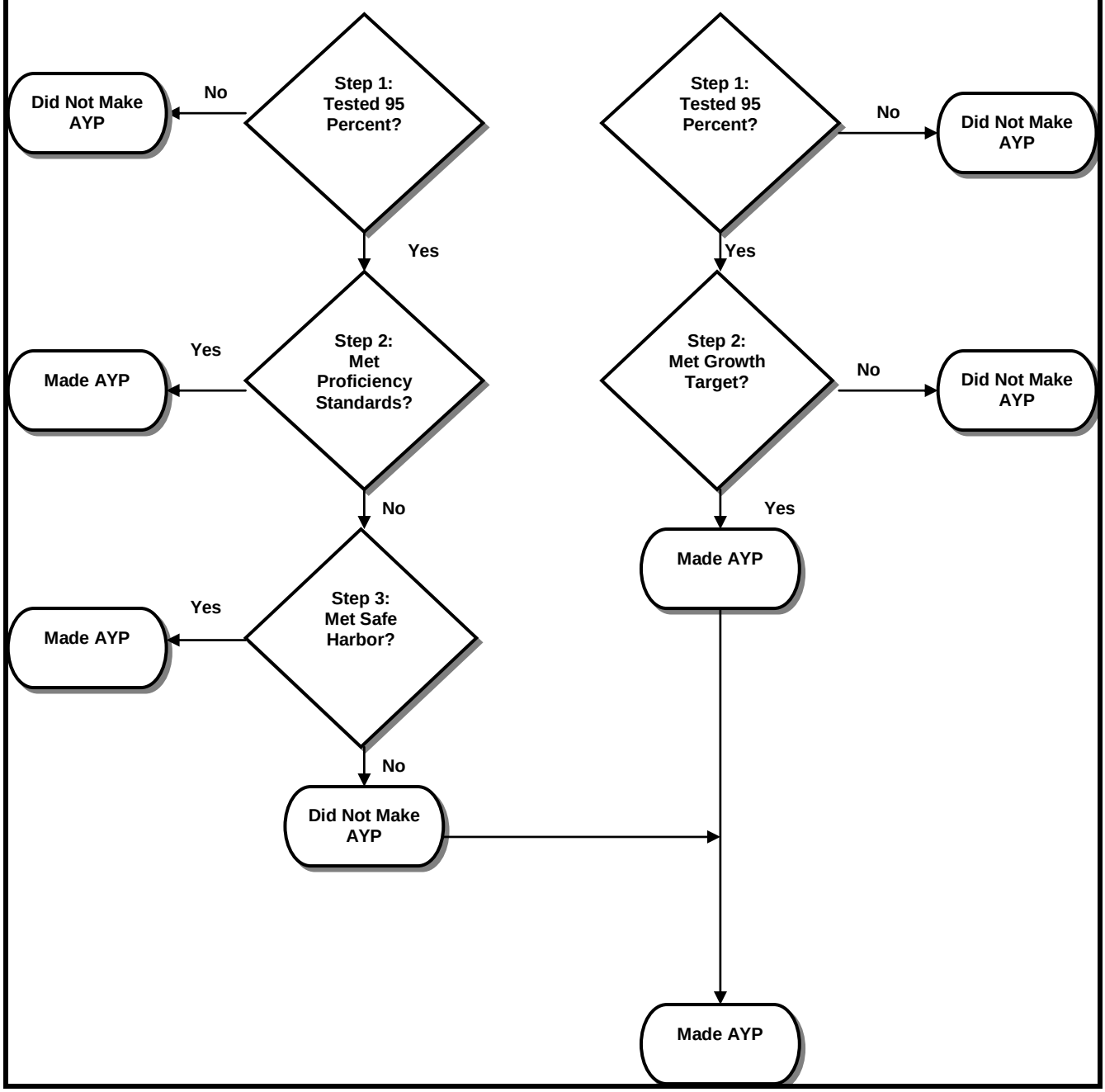
Table 2. Proposed AMOs		
Grade 4	**Reading AMO**	**Math AMO**
2005-07	45%	54%
2008-10	56 %	63.2 %
2011	67 %	72.4%
2012	78 %	81.6 %
2013	89 %	90.8 %
2014	100 %	100 %
Grade 6		
2005-07	45%	43%
2008-10	56 %	54.4 %
2011	67 %	65.8 %
2012	78 %	77.2 %
2013	89 %	88.6 %
2014	100 %	100 %
Grade 7		
2005-07	49%	48%
2008-10	59.2 %	58.4 %
2011	69.4 %	68.8 %
2012	79.6 %	79.2 %
2013	89.8 %	89.6 %
2014	100 %	100 %

The AYP evaluation does not average across grades or years when determining if a subgroup met the AMO. The minimum group size is 40 and a 99 percent confidence interval is used. Safe harbor is evaluated per statute without a confidence interval.

2.2 Incorporating the Growth Model into the AYP Evaluation

The growth model will be incorporated into the AYP evaluation for each subgroup as shown in the flow chart on the next page. The growth model will be calculated for all subgroups. If the subgroup meets the growth target, then it is considered to have made AYP. Meeting the growth target does not overrule the requirement that 95 percent of the subgroup be tested. If 95 percent of the subgroup is not tested, then it fails to make AYP regardless of percent proficient, safe harbor, or growth. Schools or districts that make AYP via the growth model will be considered to have made AYP regardless of the traditional AYP determination.

Figure 1. AMO Determination with Growth



2.3 Calculation of Growth

The proposed model calculates the annual progress made by each student toward proficiency in state standards. Progress is measured against the goal of proficiency within three years. Because Arizona's vertical scale ends at eighth grade, a goal of

proficiency by eighth grade will be used for grades five through eight. The growth target represents a horizon ranging from one to three years. The steps for calculating it are:

1. For each student an **annual growth target** is set for each subject: reading and math. The growth target is how much improvement measured by scale score points a student would have to make over her previous year's score in equal intervals in order to achieve proficiency within three years or by the eighth grade, whichever comes first. The growth target is set by subtracting the student's current year scale score from the scale score for proficiency in the eighth grade and dividing by the remaining number of grades. The targets are rounded to the nearest whole number. Demographic factors are not used to set the target.

   Example: A student scores 402 on the 3rd grade math test in 2005. The passing score on the 6th grade math test is 496. The student's math score must improve 31 points each year—(496 – 402)/(6-3) = 94/3 = 31—for him to reach proficiency by 6th grade.

   Example. A student scores 469 on the 6th grade reading test in 2005. The passing score on the 8th grade reading test is 499. The student's reading score must improve 15 points each year—(499-469)/(8-6) = 30/2 = 15—for her to reach proficiency by 8th grade.

2. For each student **actual growth** is measured for each subject.

Example: A student scores 402 on the 3rd grade math test in 2005 and 442 on the 4th grade math test in 2006. The student's actual growth is 40 points (442 – 402).

3. For each student **actual growth** is compared to the **annual growth target**. If actual growth is greater than or equal to the annual growth target, the student is deemed to have made adequate growth.

Example. A student's growth target is 31. The actual growth shown is 40. The student is considered to have made adequate growth.

Students who pass in the current year are automatically considered to have met their growth target regardless of how much actual growth they make.

Example 1. A student scores 457 on the third grade math test. Thus, the student's growth target is 16. The student scores a 460 on the fourth grade test—a passing score. Even though the student only grew by 3 points, he is proficient in the current year and thus is considered to have met his growth target.

Example 2. A student scores 457 on the third grade math test. Thus, the student's growth target is 16. The student scores a 450 on the fourth grade test—a passing score. Even though the student's score *decreased* by 7 points, he is proficient in the current year and thus is considered to have met his growth target.

This requirement makes the aggregate growth measure a status-plus measure of performance. The growth measure for a subgroup is all students who are proficient plus the number who met their growth targets divided by the total number of students included in the analysis.

The percentage of students meeting their annual growth target can be calculated by subject and grade to provide scores for subgroups, grades, schools, and districts.

The percentage of students meeting their annual growth target for a group is compared to the annual measurable objective (AMO) for that subject and grade. If the percentage is above the AMO, then the group is deemed to have made AYP. Because the AMOs increase over time, the growth targets increase over time as well. In 2014, all students will either have to be proficient or be on the road to achieving proficiency withing three years or less. This is consistent with the NCLB goal.

Example. This example shows how the growth measure can be calculated for a subgroup. This is a subgroup of 4th graders, and we are examining if it met the AMO for 4th grade math. The passing score for 4th grade is 448.

Student ID	Actual Score 3rd Grade	Passing Score 6th Grade	Growth Target	Actual Score 4th Grade	Actual Growth	Met Growth Target
1	440	496	19	450	10	Y
2	400	496	32	440	40	Y
3	450	496	15	440	-10	N
4	390	496	35	430	40	Y
5	445	496	17	440	-5	N
6	420	496	25	440	20	N
7	460	496	12	465	5	Y
8	435	496	20	454	19	Y
9	425	496	24	440	15	N
10	415	496	27	445	30	Y

The AMO for fourth grade math is 54 percent of students passing. The percentage of students meeting their growth target is 60 percent, above 54 percent, so the group will be deemed to have made AYP. A 99 percent confidence interval will be used for this determination. If the upper bound of the confidence interval is greater than 54 percent, then the subgroup will be deemed to have made sufficient growth.

2.4 Details of the Calculation

This section gives more details of the calculations, and provides additional reasoning behind the model.

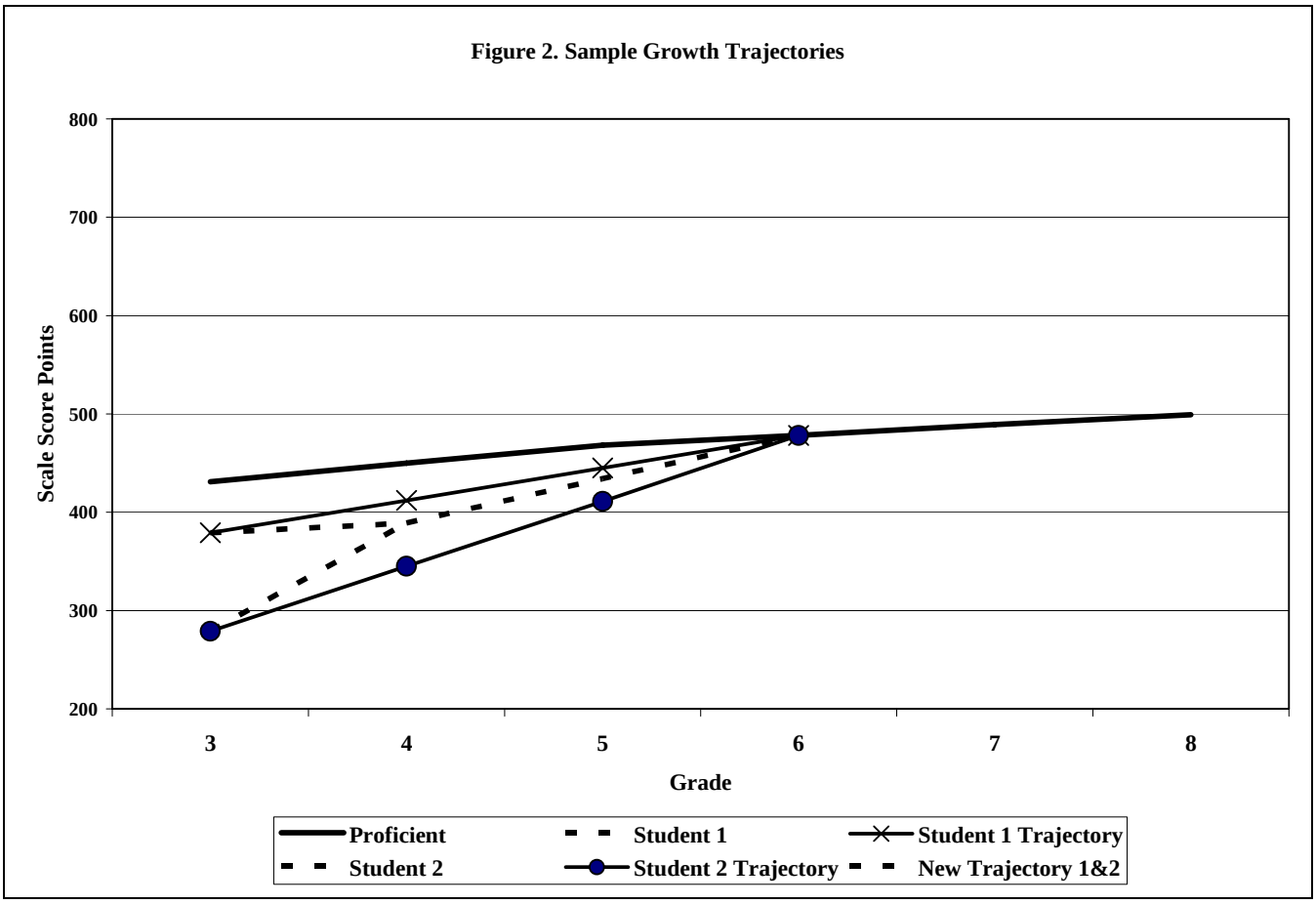
- **The model sets ambitious growth targets.** The growth model sets a target of proficiency for students within one to five years, depending on the grade. This is a much more demanding expectation of non-proficient students than under the

current AYP determination. Under the current status-only model, students who are not proficient may reach the 8th grade without making any progress at all.

- **New growth targets are set for each student, each year.** The model proposed here will set new individual growth targets for each student each year. Growth targets will be based on length of time enrolled in Arizona schools and grade level. The growth target will expect a student to become proficient within three years of first enrolling in a grade between three and eight, or by eighth grade—whichever comes first. The clock will start over for students with discontinuous enrollment histories.

  This rule was reached to address the situation shown in figure 2. The dark solid line shows the scale score considered proficient for each grade three through eight. Two students score below proficient in third grade. The lines labeled 'Student "X" Trajectory' show the amount of growth each student would need each year to reach proficiency within three years. Both score the same on the 4th grade test. Student 2 exceeded her growth target but student 1 did not. The question is: What growth target should the school set for each student for the fifth grade? Under the model proposed here, the growth target would be a new, identical trajectory for both students starting from the fourth grade score. <u>Note that the original goal of proficiency by sixth grade is maintained.</u>

  This rule has two advantages. First, it attempts to include as many students in the growth calculation as possible. Growth targets are set according to enrollment not test history, so a continuous test history is not required for a student to be included in the model. Second, it greatly eases planning and understanding for teachers, principals, and parents since it sets the same growth target for students scoring at the same level. Setting different growth targets based on a starting point that could be several years in the past would be confusing and perceived as unfair. Given the state's variation in grade configurations, student mobility, and the prevalence of charter schools, starting points could be in different schools and even unavailable to the school currently responsible for the student.



- **Average or expected growth.** The growth model proposed here sets growth targets that expect students to achieve proficiency in three years or less. It does not make use of the concepts of average, expected, or one-year's growth as they have been typically used in growth models in the past. Growth expectations in this model are not set by determining if growth is greater than a past average for the state or school. Growth expectations are not set by determining if a student maintained a given performance level from year to year.

  Given that students below proficient may have up to three years to traverse the two performance levels below proficient, it necessarily follows that a student could meet the growth target and not change his performance level. This is even more likely given the relative scale score span of the lowest performance level (falls far below) compared to the next highest (approaches the standard).

Example 1. A student scores 200 on the third grade reading test. The scale score for proficiency in sixth grade is 478. The student's annual growth target is (478-200)/3 = 93. If the student met the growth target each year, her performance levels would be:

Grade	Scale Score	Performance Level
3	200	Falls far below
4	293	Falls far below
5	386	Falls far below
6	479	Meets the standard (proficient)

Example 2. A student scores 386 on the third grade math test. The scale score for proficiency in sixth grade is 496. The student's annual growth target is (496-386)/3 = 37. If the student met the growth target each year, her performance levels would be:

Grade	Scale Score	Performance Level
3	386	Approaches
4	423	Approaches
5	460	Approaches
6	496	Meets the standard (proficient)

- **Statistical significance of growth.** In order to ensure that the model produces valid and reliable determinations of school effectiveness, a 99 percent confidence interval will be used to determine if a subgroup met the growth goal (AMO). If the upper bound of the confidence interval surrounding the subgroup's growth measure is greater than the AMO, then the subgroup is deemed to have made AYP.

  The purpose of the confidence interval is to guard against error caused by the variation in student performance from year to year. That is, aggregate student performance may vary from year to year due to random factors outside school control. The confidence interval prevents these factors from erroneously labeling a school as in need of improvement. Because the growth model is still a conjunctive model (all subgroups must make AYP for a school to make AYP) ADE believes that a 99 percent confidence interval applied at the subgroup level produces an acceptable level of error at the school or district level.

  The model proposed here does not use the standard error of measurement (SEM) around test scores to determine if annual growth targets or actual growth at the individual level are significant.

2.5 Alternate Tests

Arizona only administers one alternate test, the AIMS-A. The AIMS-A is designed for only the most significantly cognitively disabled students. Students who take the AIMS-A will be included in the growth model. Since the AIMS-A does not have a scale, the percent of expected growth results will be determined by examining the change in performance level. Like the AIMS, the AIMS-A has four performance levels: falls far below (FFB), approaches (A), meets (M), and exceeds (E). The last two levels are considered proficient. As above, students who pass in the current year, regardless of performance level, will be considered as having met their annual growth target. Students not proficient in the current year who fall a performance level will be considered as having not met their growth target. Students not proficient in the current year but moving up one performance level will be considered as having met their growth target.

Table 3. Met Annual Growth Target for Alternate Assessment

	Performance Level Current Year			
Performance Level Previous Year	**FFB**	**A**	**M**	**E**
FFB	N	Y	Y	Y
A	N	N	Y	Y
M	N	N	Y	Y
E	N	N	Y	Y

The above table would also be used in the unlikely event that a student would take the AIMS-A one year and the AIMS the next.

2.6 Inclusion Rules and Missing Data

To be included in the growth measure a student must have been enrolled in a school for the full academic year. However, to be included a student does not have to be enrolled in the same school or district for two consecutive years. Furthermore, to be included a student does not have to have been enrolled for a full academic year the previous year. ADE considers a student to have been enrolled for a full academic year if the student was continuously enrolled in a school from within the first ten school days through the testing date.

Example 1. In 2005 student A was enrolled in Gila Monster Elementary for the full academic year. In 2006, she enrolled in Roadrunner Elementary at the beginning of the year and remained enrolled up to the time she took the test. She will be included in the growth model.

Example 2. In 2005 student B moved in from out of state in the middle of the year and enrolled in Gila Monster Elementary. In 2006, he continued at Gila Monster and remained enrolled the entire year. He will be included in the growth model.

Example 3. In 2005 student C was enrolled in Gila Monster Elementary for the full academic year. She stayed with relatives out of state and re-enrolled in Gila Monster two

months after the beginning of the 2006 school year. She will not be included in the growth model.

Because the proposed model sets its goals based on length of time enrolled missing test scores are not as significant an issue as they would be for multi-year models. ADE will evaluate and report the growth for subgroups with ten students or more. However, the growth measure will be used in the determination of AYP for a subgroup only if the number of matched students is greater than or equal to 90 percent of the total number of students used in the status determination or 40 students, whichever is greater. The conventional status measure will be used to evaluate subgroups with too few matched students. Subgroups with less than 40 students will not be evaluated.

Example 1. A subgroup has 50 students, of which 48 have been successfully matched. Since 48 is 96 percent of 50, the subgroup will be evaluated using the growth measure.

Example 2. A subgroup has 50 students, of which 44 have been successfully matched. Since 44 is only 88 percent of 50, the subgroup will be evaluated using the status measure.

Example 3. A subgroup has 40 students of which 39 have been successfully matched. Since 39 is less than 40, the subgroup will be evaluated using the status measure.

Example 4. A subgroup has 35 students. Since 35 is less than 40 the group will not be evaluated regardless of the number of students successfully matched.

2.7 K-2 Schools, High Schools, and Third Grade

A growth model will not be used in the evaluation of K-2 or high schools. Those schools will continue to be evaluated using the current method. Third grade will continue to be evaluated using the current status/safe harbor method as well.

3 Arizona's Growth Model Meets the Seven Core Principles

3.1 100 percent proficiency by 2014 and Incorporating Decisions about Student Growth into School Accountability

3.1.1 How does the state accountability model hold schools accountable for universal proficiency by 2013-14?

Arizona's growth model holds schools and districts accountable for the academic growth of all students. This makes it an improvement over the current status-only AYP evaluation, where schools have the incentive to only focus on the students who are near proficiency.

Growth is measured as the change in scale score points in consecutive years. A student's individual growth target is the growth necessary to achieve proficiency by the time the student reaches eighth grade. The growth measure for a subgroup is the percentage of students in the subgroup meeting their growth targets. The percentage required for all subgroups are the AMOs for proficiency (shown in tables 1 and 2) for that grade, subject, and year. As a result, the number of students required to meet their growth target grows along with the AMOs. Thus, by 2013-14 the target is that all students will be on track to attain proficiency.

As the flowchart in figure 1 shows, all students will be included in the growth model. It will be a parallel evaluation of schools and districts. Consequently, status and safe harbor will continue to be calculated and used for evaluation as before. This will facilitate the monitoring and evaluation of the effectiveness and validity of the growth model.

The growth model will be used for grades four through eight. Grade three, high schools, and K-2 schools will be evaluated using the current approved method.

3.1.2 Has the state proposed technically and educationally sound criteria for growth targets for schools and subgroups?

The Arizona growth model preserves the annual expectation that the goal of all students in all grades and schools is to reach proficiency. The model sets growth targets so that students will close the achievement gap each year so that they will reach proficiency in three years or less. (Refer to section 2.3 for the details of the calculation.) Thus the growth target is the same for all subgroups in the relevant grade.

The AMOs provide challenging and continuously improving expectations for student growth. They are goals familiar to schools that have already been the focus of planning for the past three years. The AMOs have an empirical basis in their starting point.

3.1.3 Has the state proposed a technically and educationally sound method of making annual judgments about school performance using growth?

The growth measure will be applied in the same manner to all subgroups for both subjects—reading and math—for all grades four through eight. Its use will ensure that schools receive credit for, and hence have the incentive to, improve student achievement over all performance levels. This will help to see that as students advance up through the grades some will not be left further behind. It will also help ensure that there are no false negatives; that schools that are successful in raising the performance of low-achieving students will not be mistakenly identified as needing improvement.

The growth measure will include all students who have attended a school for the full academic year that can be successfully matched from the previous year—including students who take the alternate assessment. If the match rate for any subgroup is less than 90 percent, the growth measure will be reported but not used in the AYP evaluation. This is to ensure the validity of the data and to remove the incentive to selectively thwart students from being matched. A confidence interval of 99 percent will be applied.

The growth measure will not affect the other indicators of school performance: the 95 percent assessment requirement, the graduation rate, or the attendance rate.

The growth measure will be released to the public and placed on the school, district, and state report cards. It is intuitive and transparent and will be familiar to most parents, educators, and policymakers in Arizona. Its use in AYP evaluations minimally disrupts the formula.

The growth goals for subgroups will increase along with the AMOs so by the year 2013-14 100 percent of students must either be proficient or on track to be proficient within three years. If we return to our example from section 2.3 we see that this subgroup would NOT make AYP in 2014 because only 60 percent of the students are proficient or on track to proficiency.

Student ID	Actual Score 3rd Grade	Passing Score 6th Grade	Growth Target	Actual Score 4th Grade	Actual Growth	Met Growth Target
1	440	496	19	450	10	Y
2	400	496	32	440	40	Y
3	450	496	15	440	-10	N
4	390	496	35	430	40	Y
5	445	496	17	440	-5	N
6	420	496	25	440	20	N
7	460	496	12	465	5	Y
8	435	496	20	454	19	Y
9	425	496	24	440	15	N
10	415	496	27	445	30	Y

The subgroup in the table below, however, would be deemed to have made AYP. Even though not all students are proficient, all students are on track to be proficient within three years.

Student ID	Actual Score 3rd Grade	Passing Score 6th Grade	Growth Target	Actual Score 4th Grade	Actual Growth	Met Growth Target
1	440	496	19	450	10	Y
2	400	496	32	440	40	Y
3	450	496	15	465	15	Y
4	390	496	35	430	40	Y
5	445	496	17	462	17	Y
6	420	496	25	445	25	Y
7	460	496	12	465	5	Y
8	435	496	20	454	19	Y
9	425	496	24	449	15	Y
10	415	496	27	445	30	Y

Note that the model implies that all students enrolled in an Arizona school in 2014 will be proficient by 2017. After 2017, only students who move in from out of state after 2014 would not be required to be proficient for a school to make AYP.

3.1.4 Does the state-proposed growth model include a relationship between consequences and rate of student growth consistent with Section 1116 of ESEA?

The growth model will be a parallel evaluation to the traditional AYP determination. Regardless of the determination by the traditional model, schools found to have made AYP by the growth model will be considered to have made AYP and hence not in school improvement. We anticipate the growth model being used by the state's school improvement team and educators for diagnostic and planning purposes. However, a school or district's performance measured by the growth model will not mitigate, alter, or enhance the consequences of being identified for improvement. Schools and districts identified for improvement by the growth model will have to meet the consequences of improvement—parental notification, choice, supplemental services—as outlined in section 1116.

3.2 Establishing Appropriate Growth Targets at the Student Level

A detailed description of the student level growth measure is given in section 2.3 beginning on page seven. The annual growth target is the difference between the student's previous year scale score and the proficient score for three grades higher, divided by the number of intervening grades. Students who are proficient in the current year are considered to have met their growth target. This individual growth measure provides an easily interpretable and transparent measure of student progress. The simple calculation facilitates its use in planning for students, classrooms, and schools.

The growth targets, the annual measurable objectives, and the rationale for their use are described in section 3.1.2 above. We underscore that we are setting expectations for growth that are not based on student demographic characteristics.

The tests administered in 2006 will represent the second year that Arizona has tested at all grades three through eight. It will also be the second administration of the test with a

vertical scale. As the data set for the new test deepens, ADE intends to continually evaluate the growth goals set for students. The growth goals proposed here represent ambitious one-year targets for students. Additional years of data will shed light on what growth trajectories are both necessary and feasible for a student to reach proficiency.

3.3 Accountability for Reading/Language Arts and Mathematics Separately

As described in section 2.3 and section 3.1 the growth calculation mirrors the AMO determination: it is done separately for reading and math for each subgroup. It will also be reported by subject. Arizona also assesses writing but does not include it in the AYP evaluation at this time.

The growth model submitted ensures the validity and reliability of its determinations by:

- Placing a minimum match requirement of 90 percent for the growth measure to be considered in the AYP determination for a subgroup.

- Using a confidence interval of 99 percent to minimize the impact of measurement error.

- Using a dichotomous yes/no determination as to whether a student met her growth target. This will mitigate the impact of outliers and fluke results, and prevent growth above proficient from compensating for growth that will not cause a student to reach proficiency.

ADE evaluates AYP for small schools by averaging over three years' of data. Since the new vertically scaled test has only been in place for two years, the growth model could not be incorporated into the evaluation of small schools until the evaluations released in the fall of 2007.

3.4 Inclusion of All Students

The growth model will be applied to all students in all schools that serve grades four through eight, subject to the rules regarding minimum group sizes as outlined in section 2.6. Third grade, high schools, K-2 schools, and small schools will be evaluated using the current, approved method (see section 2.7). Students taking the alternate assessment will be included according to the rules outlined in section 2.5. Arizona law requires that limited English proficient students take the English language version of the AIMS. The students in this subgroup will be evaluated using the same method for all students described above. English language learners will not have extended time horizons to meet proficiency.

Students who change program membership will be counted in the subgroup to which they belong in the current year. Growth for promoted or retained students will be measured by the change in scores for consecutive tests, regardless of the grade level of the tests.

3.5 State Assessment System and Methodology

3.5.1 Has the state designed and implemented a statewide assessment system that measures all students annually in grades 3-8 and one high school grade in reading/language arts and mathematics in accordance with NCLB requirements for 2005-06, and have the annual assessments been in place since the 2004-05 school year?

Arizona's Instrument to Measure Standards Dual Purpose Assessment (AIMS DPA) is used in grades three through eight, Arizona's Instrument to Measure Standards High School (AIMS HS) is used in high school, and Arizona's Instrument to Measure Standards-Alternate (AIMS-A) is used to assess the competency and growth of students with significant cognitive disabilities in grades three through eight and high school. AIMS DPA and AIMS HS are selected response tests for all grade levels and AIMS-A uses a performance evaluation, a parent interview and activity-based evidence. Students are currently assessed in reading, mathematics, and writing. A science test is under development.

Arizona tested all grades three through eight and high school operationally in 2004-05. In 2005 the test underwent a standard setting and a vertical scale was developed for grades three through eight.

AIMS DPA is a statewide assessment that is both standards-based, measuring students knowledge against the Arizona Academic Standards, and norm referenced, comparing student knowledge against students nationwide. TerraNova, norm-referenced test items included in the test, compare a student's performance to students nationwide. AIMS test items measure a child's knowledge without comparison to other students based on whether the child is proficient in accordance with the Arizona Academic Standards.

The AIMS DPA provides a norm-referenced test score (NRT) and an AIMS score. The results of the tests are used by classroom teachers to guide instruction and improve student learning. The NRT compares a child's performance on certain test items to the performance of students nationwide based on percentile ranking. The AIMS score assesses a child's knowledge of the Arizona State Standards. Only the AIMS score is used for accountability purposes. Each child will receive one of the following scores for each of the three areas tested - reading, writing, and mathematics:

Exceeds the Standard – denotes superior academic performance on challenging subject matter reflected by the content standards
Meets the Standard – denotes solid academic performance and understanding of the state content standards
Approaches the Standard – denotes partial understanding of the skills and knowledge necessary for proficient work at grade level
Falls Far Below the Standard – denotes insufficient understanding of the prerequisite skills. Students who achieve at this level have serious gaps in knowledge and skills and may require remediation.

Performance levels of meets and exceeds are considered proficient.

The state has submitted State wide assessment system evidence for NCLB Peer Review for the November 2005 review. We are awaiting a response from the U.S. ED.

Arizona Academic Standards

The Arizona Academic Standards are established academic expectations for teaching and learning. They are statements of knowledge and skills that every child is expected to learn. The standards were developed by the state's educational community and adopted by the State Board of Education to meet federal and state guidelines. Currently, academic standards for grades three through eight have been developed in the following content areas: reading, writing, mathematics, science, social studies, foreign language, technology, arts, comprehensive health/PE, and workplace skills.

Arizona's Instrument to Measure the Standards (AIMS) is aligned to the grade-level Arizona Academic Standards in reading, writing and mathematics. The test blueprints for AIMS DPA, AIMS HS and AIMS-A were followed during item selection as test forms were developed. A committee met in the spring of 2003 and developed assessment blueprints for reading and mathematics that covered the content standards. The emphasis that they placed on the number of items for future assessments reflected both the content standards and classroom instruction across the state. The blueprints were then reviewed by both the state (SAAAC) and national (NAAAC) technical advisory committees. An external alignment study confirming the degree of alignment of the assessments to the standards was conducted July 25-27, 2005.

3.5.2 How will the state report individual student growth to parents?

The Arizona Department of Education, in collaboration with CTB/McGraw-Hill, developed paper reports that can be appropriately interpreted for students, grade level groups, schools, districts, counties and the state. Parent reports contain the students' scores from previous years so that parents may track student growth. Reports for AIMS HS and AIMS DPA are explained in interpretive guides that accompany reports when they are distributed to schools. These guides are also posted on the Arizona Department of Education website for school use. ADE has posted on the website and distributes at public forums parent brochures to help parents understand AIMS DPA and AIMS HS. Student reports reflect student performance in academic content domains (reading, writing, and mathematics) and indicate achievement levels (falls far below, approaches, meets, or exceeds standards) to which the student's scores correspond. On the back of each student report form the performance level descriptors are provided to enhance understanding. The report visually displays how the student is performing to assist families who do not read English or Spanish.

3.5.3 Does the statewide assessment system produce comparable information on each student as he/she moves from one grade level to the next?

The criterion-referenced portions of the AIMS tests in grades three through eight were placed on a vertical scale using TerraNova NRT items as external anchors. Parameters for these items were obtained during national standardization. The AIMS scales for

grades three through eight ranges from 200–800. The AIMS high school assessments were not placed on the grades three through eight scale.

All items that made up the AIMS criterion referenced test (CRT) for each subject and grade were calibrated in Pardux using the 3 PLM and placed on the TerraNova scale using the TerraNova NRT standardization item parameters as anchor estimates. Equating was done using the Stockard and Lord procedure that matches the test characteristic curves of an anchor test and the focal test. The mean and standard deviation of the Arizona students' scores on the CRT were then calculated on the TerraNova scale for each grade. The smoothed means of AIMS CRT on the TerraNova scale fitted to a second degree polynomial was used to articulate the vertical scale.

Transformation constants were calculated to convert the mean and standard deviation for the third grade to a mean of 450 with a standard deviation appropriate to the content area. The same constants were then applied to the means and standard deviations for all other grades to place the AIMS CRT on a scale of 200-800.

After transformation to the new scale, means, standard deviations, and distributions were examined to ensure acceptable articulation. Minor adjustments were made to eliminate overlap. Standard deviations were adjusted slightly and the eighth grade observed mean was used instead of the smoothed mean.

The lowest and highest obtainable scale scores were set for each grade and subject using the following criteria:

The highest obtainable scale score (HOSS) had to be greater than the scale score associated with one less than the maximum raw score; increase monotonically; be high enough to prevent a ceiling effect; low enough so that the Standard Error of Measurement (SEM) does not exceed ten times the minimum SEM; and the scale score gap between the HOSS and the scale score associated with one minus the raw score maximum had to be similar to the scale score gap between the scale scores associated with the raw score maximum minus one and the raw score maximum minus two.

The lowest obtainable scale score (LOSS) also had to increase monotonically; be low enough to prevent floor effects; and high enough so that the SEM associated with the LOSS is less than ten times the minimum SEM.

Grades four, six, and seven were tested for the first time in the academic year 2004 – 2005. ADE will provide evidence for equating test scores between grades three through eight using the mean equating method after the 2005-06 administration. A set of anchor items from the 2005 test has been reserved to control for drift in the scale.

The cut scores for each subject and grade are given in the tables below.

Table 4. Mathematics Cut Scores for Grades 3 Through 8						
	Grade					
	3	4	5	6	7	8
Approaches	386	414	442	463	484	505
Meets	420	448	476	496	517	537
Exceeds	492	521	550	574	599	623

Table 5. Reading Cut Scores for Grades 3 Through 8						
	Grade					
	3	4	5	6	7	8
Approaches	379	402	424	433	443	452
Meets	431	450	468	478	489	499
Exceeds	516	536	556	571	587	602

3.5.4 Is the statewide assessment system stable in its design?

There will be no changes to the grades assessed, content, instruments, scoring, and cut-scores for the next four years.

3.6 Tracking Student Progress

ADE has the ability to track students through their career kindergarten through graduation in Arizona's public schools via its Student Accountability Information System (SAIS). Since the 2001-2002 school year, students enrolling in Arizona schools have been assigned a unique identifier, the SAIS id that remains with them across grades, schools, and districts until they graduate from high school. SAIS also contains student demographic information—gender, ethnicity, etc.—and program membership, whether the student is an English language learner, special education, or eligible for a free/reduced lunch. Student testing data is maintained in a separate database. Starting with the 2003-04 school year, SAIS ids were attached to the state's testing data.

As mentioned before, Arizona has published a measure of student growth, the MAP, since 1999. Matching for the MAP was based on student names, birthdates, and other characteristics. ADE achieved a highly successful match rate via these means. ADE intends to match students on their SAIS ids to determine growth. The table below shows the results from matching 2005 to 2004 test data using student SAIS ids. The method achieves nearly a 90 percent match rate that does not differ by program membership or by proficiency level. As can be expected, the match rate is lower for mobile students, those not enrolled in a school for the full academic year.

Table 6. Match Rates	
Subgroup	Match Rate
Total	89%
English language learners	85%
Free/Reduced lunch	89%
Special Education	89%
Full academic year	93%
Non-full academic year	63%
Proficient	94%
Not-proficient	91%

The key to having a high match rate is to make every effort to ensure every test record has a valid student identifier attached. ADE takes several steps, outlined in figure 3, to see that this is the case.

1. ADE distributes pre-printed test labels for all students in the state containing students' unique identifiers. This minimizes errors due to students or administrators bubbling in student information on test answer documents.

2. After the test is scored and ADE receives the test score results from the testing company, ADE matches the test scores back to its student database, SAIS, to update the student demographic information associated with the test scores.

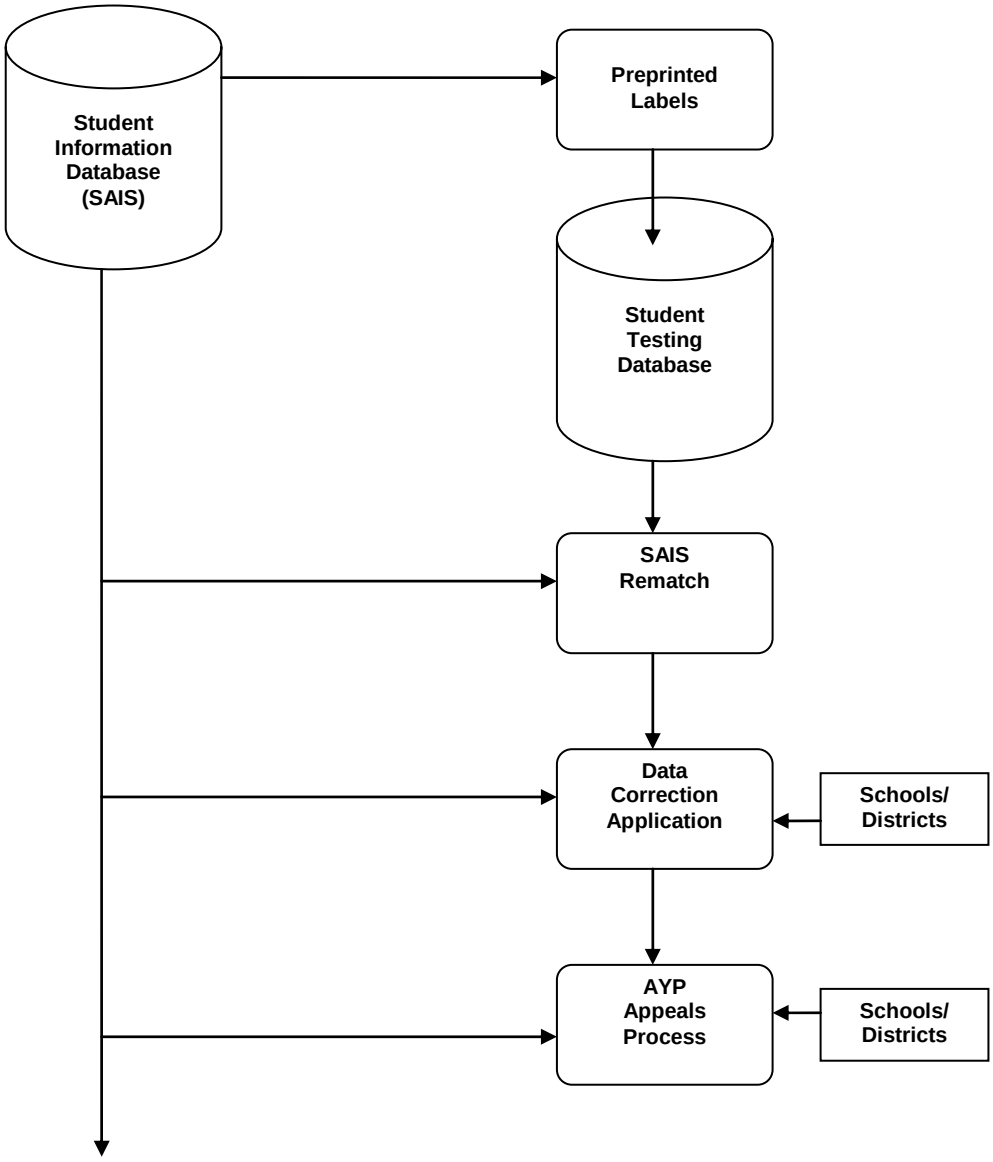
3. ADE then makes the test score data available to schools and districts. ADE has two online applications that allow schools to download a copy of their data as it exists in the ADE test database, make corrections, and upload corrected files to the database. At the same time schools may correct the student demographic and program information in SAIS. The only corrections schools may make directly to the test score database are test accommodations, school information, and the SAIS identification number associated with the test. All corrections to student demographic information must be made through SAIS. Demographic information associated with test scores is then corrected by repeated re-matching of the test score file to the student information file.

4. ADE provides additional information to schools to help them identify test records with faulty student identifiers. First, it provides on-line reports that allow schools to view records in the test database with faulty ids. Second, it has added a field to the test score database that identifies records with invalid SAIS ids. Schools can then see those records when they download their data.

5. ADE conducts an AYP appeals process during which schools are given a final opportunity to correct their data.

ADE also conducts trainings for administrators on both accountability and use of the SAIS database, and has in place standard procedures to maintain the integrity of the SAIS database.

Figure 3. ADE Data Verification Process



Section 2.6 describes the policies for missing data.

3.7 Participation Rates and Additional Academic Indicators

The use of the growth model will not affect how participation rates or the additional academic indicators are used in AYP evaluations.

4 Conclusion

Arizona is well qualified to incorporate a growth model into its AYP evaluations for schools and districts.

1. Arizona has in place all the necessary elements for a growth model. The SAIS database assigns each student a unique identifier that tracks students across grades, schools, and districts. It also holds student demographic information. The AIMS test has been administered operationally in all grades three through eight, is vertically scaled, and will remain stable for the next four years.

2. Arizona has experience with growth measures. Student growth has been a part of Arizona's school accountability system since the previous century. Growth measures have been reported publicly and are used by administrators. We will continue to do so regardless of the status of this proposal.

3. Arizona's growth model provides a more valid measure of school effectiveness than the current status-only evaluation. It recognizes and gives schools the incentive to improve student performance at all levels.

4. Arizona sets ambitious growth targets for all students in all schools. Expected growth is not dependent on students' demographic characteristics. It sets the ultimate goal of all students being proficient or on track to proficiency by 2014.

If approved, the Arizona growth model would lessen the possibility that students would fall behind. It will increase the number of students who move on to higher grades that are ready to learn. Consequently, adding the growth model to the AYP determination will help close the achievement gap more rapidly.

Appendix: Responses to March 10 Letter from Henry L. Johnson

1. **Please describe the expected or average growth each year for a student.** The model sets individual growth targets for students that will put them on the path to achieve proficiency within three years or by the time they reach eighth grade—whichever comes first. Each year the student must show sufficient growth so that within three years she is proficient.

   The model does not make use of the concepts of average, expected, or one-year's growth as they have been typically used in growth models in the past. Growth expectations in this model are not set by determining if growth is greater than a past average for the state or school. Growth expectations are not set by determining if a student maintained a given performance level from year to year.

   For more details, please see sections 2.3 and 2.4 above

2. **Please clarify if it is possible that a student's growth target could be less than their "expected" growth to maintain their current level of performance.** No, because students below proficient have to grow enough to increase their performance level, growth must be greater than the growth of the performance level cut scores. However, given that students below proficient may have up to three years to traverse the two performance levels below proficient, it necessarily follows that a student could meet the growth target and not change his performance level. This is even more likely given the relative scale score span of the lowest performance level (falls far below) compared to the next highest (approaches the standard). Please see the examples in section 2.4.

3. **Please provide a rationale for the use of a confidence interval in the growth model.** The purpose of the confidence interval is to ensure valid and reliable evaluations by guarding against error caused by the variation in student performance from year to year. That is, aggregate student performance may vary from year to year due to random factors outside school control. The confidence interval prevents these factors from erroneously labeling a school as in need of improvement. Because the growth model is still a conjunctive model (all subgroups must make AYP for a school to make AYP) ADE believes that a 99 percent confidence interval applied at the subgroup level produces an acceptable level of error at the school or district level. Please see section 2.4 for additional discussion.

4. **Provide an example of how the model would be implemented for a typical school in 2013-14.** In order for a school to make AYP in 2013-14 100 percent of the students in all of its subgroups must either be proficient or have shown sufficient growth to attain proficiency within three years. The subgroup in the table below would be deemed to have made AYP even though not all students are

proficient. (Proficient students are **in bold**.) The model implies that all students enrolled in an Arizona school in 2014 will be proficient by 2017.

Student ID	Actual Score 3rd Grade	Passing Score 6th Grade	Growth Target	Actual Score 4th Grade	Actual Growth	Met Growth Target
1	**440**	**496**	**19**	**450**	**10**	**Y**
2	400	496	32	440	40	Y
3	**450**	**496**	**15**	**465**	**15**	**Y**
4	390	496	35	430	40	Y
5	**445**	**496**	**17**	**462**	**17**	**Y**
6	420	496	25	445	25	Y
7	**460**	**496**	**12**	**465**	**5**	**Y**
8	**435**	**496**	**20**	**454**	**19**	**Y**
9	**425**	**496**	**24**	**449**	**15**	**Y**
10	415	496	27	445	30	Y

5. **Please clarify that the proposed model applies consequences consistent with Section 1116.** The consequences of not making AYP are not mitigated, enhanced, put out of sequence, or modified in any other way by the model. Schools identified as not making AYP will face all the consequences of school improvement as specified in Section 1116.

6. **Please clarify whether the growth model will be applied to all students in every school in the State.** The growth model will be applied to all students in every subgroup in grades four through eight. Students in K-2 schools, high schools, small schools, and third grade will be evaluated using the current, approved system. Please see sections 2.2, 2.7, and 3.4 for further discussion.

7. **Please clarify how the growth model applies the full academic year requirement.** To be included in the growth measure a student must have been enrolled in a school for the full academic year. However, to be included a student does not have to be enrolled in the same school or district for two consecutive years. Furthermore, to be included a student does not have to have been enrolled for a full academic year the previous year. ADE considers a student to have been enrolled for a full academic year if the student was continuously enrolled in a school from within the first ten school days through the testing date.

   Example 1. In 2005 student A was enrolled in Gila Monster Elementary for the full academic year. In 2006, she enrolled in Roadrunner Elementary at the beginning of the year and remained enrolled up to the time she took the test. She will be included in the growth model.

   Example 2. In 2005 student B moved in from out of state in the middle of the year and enrolled in Gila Monster Elementary. In 2006, he continued at Gila Monster and remained enrolled the entire year. He will be included in the growth model.

Example 3. In 2005 student C was enrolled in Gila Monster Elementary for the full academic year. She stayed with relatives out of state and re-enrolled in Gila Monster two months after the beginning of the 2006 school year. She will not be included in the growth model.

8. **Provide an example of how the State will implement the minimum 'n' in the growth model.** Here are four examples:

   Example 1. A subgroup has 50 students, of which 48 have been successfully matched. Since 48 is 96 percent of 50, the subgroup will be evaluated using the growth measure.

   Example 2. A subgroup has 50 students, of which 44 have been successfully matched. Since 44 is only 88 percent of 50, the subgroup will not be evaluated using the growth measure. It will be evaluated using the status measure.

   Example 3. A subgroup has 40 students of which 39 have been successfully matched. Since 39 is less than 40, the subgroup will be evaluated using the status measure.

   Example 4. A subgroup has 35 students. Since 35 is less than 40 the group will not be evaluated regardless of the number of students successfully matched.

9. **Provide a description of how limited English proficient students will be included in the growth model.** Arizona assesses its students only in English. LEP students will be included in the model using the exact same calculation and growth targets as for other students. LEP students will not be exempt for a certain number of years nor have longer horizons to attain proficiency.

10. **Please clarify how Arizona proposes to report student growth to parents.** Aggregate measures of growth for schools and districts will be reported on school and district report cards. Parents are provided individual reports that show their child's test history, and interpretive guides to help them understand his or her progress. Given that these reports are printed and distributed in May, and the lead time involved with our testing contractor, it is too late to include the AYP growth measure on the reports distributed this year. ADE will explore the feasibility of including the measure on future reports, and look for other avenues to provide parents student-level information.

11. **Provide additional information on match rates for above proficient and below proficient students.** The match rate for proficient students is 94 percent. The match rate for below proficient students is 91 percent.

12. **Provide additional information regarding the quality assurance procedures used to maintain the accuracy of the student matching system.** The key to having a high match rate is to make every effort to ensure every test record has a

valid student identifier attached. ADE takes several steps, outlined in figure 3, to see that this is the case.

1. ADE distributes pre-printed test labels for all students in the state containing students' unique identifiers. This minimizes errors due to students or administrators bubbling in student information on test answer documents.

2. After the test is scored and ADE receives the test score results from the testing company, ADE matches the test scores back to its student database, SAIS, to update the student demographic information associated with the test scores.

3. ADE then makes the test score data available to schools and districts. ADE has two online applications that allow schools to download a copy of their data as it exists in the ADE test data base, make corrections, and upload corrected files to the database. At the same time schools may correct the student demographic and program information in SAIS. The only corrections schools may make directly to the test score database are test accommodations, school information, and the SAIS identification number associated with the test. All corrections to student demographic information must be made through SAIS. Demographic information associated with test scores is then corrected by repeated re-matching of the test score file to the student information file.

4. ADE provides additional information to schools to help them identify test records with faulty student identifiers. First, it provides on-line reports that allow schools to view records in the test database with faulty ids. Second, it has added a field to the test score database that identifies records with invalid SAIS ids. Schools can then see the records when they download their data.

5. ADE conducts an AYP appeals process during which schools are given a final opportunity to correct their data.

ADE also conducts trainings for administrators on both accountability and use of the SAIS database, and has in place standard procedures to maintain the integrity of the SAIS database.