Guide to United States Department of Education Growth Model Pilot Program 2005-2008

A paper commissioned by the
CCSSO Accountability Systems and
Reporting State Collaborative

COUNCIL OF CHIEF STATE SCHOOL OFFICERS

The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

STATE COLLABORATIVE ON ASSESSMENT AND STUDENT STANDARDS
The State Collaborative on Assessment and Student Standards (SCASS) Project was initiated in 1991 to encourage and assist States in working collaboratively on assessment design and development of high quality assessments in relation to standards for student learning. State education agencies voluntarily support collaborative work across a variety of topics and subject areas.

The Accountability Systems and Reporting (ASR) state collaborative project has been assisting States to develop and improve statewide systems since 2000. Currently 28 state education agencies are participating members, together with researchers, consultants, and measurement specialists. Priority activities of the ASR Collaborative are: state planning and development to meet requirements for state systems under NCLB, sharing strategies and research-based evidence to improve the validity of accountability systems, developing reporting formats that improve information access for constituents, and establishing standards for the essential components of accountability systems. Funding support for this paper was entirely from member States of the Accountability Systems and Reporting State Collaborative (ASR-SCASS). For further information about ASR SCASS and resources on accountability across the 50 States, see the CCSSO website: http://www.ccsso.org/projects/Accountabilty_Systems/.

Council of Chief State School Officers
Attn: Publications
One Massachusetts Ave., NW, Suite 700
Washington, DC 20001
202-336-7016
Fax: 202-408-8072
www.ccsso.org

Guide to United States Department of Education Growth Model Pilot Program 2005-2008

ASR Working Group
Kimberly O'Malley, Pearson
William Auty, Consultant
Paul Bielawski, MI
Robert Bernstein, CA
Tom Boatman, MN
Tom Deeter, IA
Pete Goldschmidt, CRESST
Glenn Hirata, HI
Christy Hovanetz, MN
George Michna, CT
Shula Nedley, PA
Becky Odneal, MO
Joanne Reihm, DE
Lauren Stillman, CCSSO
Rolf Blank, CCSSO

A paper commissioned by the CCSSO Accountability Systems and Reporting State Collaborative
March 2009

Information on state growth models was collected from state accountability directors, state websites, and proposals and reports submitted to US ED.

Council of Chief State School Officers, Washington, DC.

Table of Contents

Guide to United States Department of Education (US ED) Growth Model Pilot Program 2005-2008

Background

Interest in using student progress or growth information for evaluating schools/districts and teachers is growing. In November 2005, The United States Department of Education (US ED) requested state proposals for accountability models that incorporate measures of student growth. States were encouraged to submit proposals to the Department for using growth models to demonstrate accountability under the federal No Child Left Behind (NCLB) Act. States submitting proposals were required to show how their growth-based accountability models satisfied NCLB alignment and foundational elements explained in the November 2005 letter. As part of the pilot, US ED announced in 2005-2006 that it would approve no more than 10 high-quality growth models. After more than 20 states submitted proposals and 9 states were approved for the pilot study, Secretary Spellings announced in December 2007 that US ED was opening the growth pilot study to all eligible states.

Purpose

The purpose of this document is to describe the US ED growth model pilot program from its inception to the point at which the program was no longer a pilot and was opened up to all states. In addition, the paper is designed to help states in the process of planning to submit a growth model proposal. The information in this document is mostly public information from a variety of sources. By pulling the facts and figures together and organizing them as they are in this document, states should find the paper valuable, as it will provide a rich source of information about what types of growth models have been approved, features of growth models that have been important in the program, and the impact the growth models have had on state's AYP calculations.

This document summarizes activities related to the growth pilot study from the start of the pilot in 2005 to the end of 2008. Though US ED continues to encourage states to submit growth-based accountability proposals, this document includes only events and approved models through December 2008. The document summarizes which states submitted proposals and which were approved. The paper describes the principles around the growth pilot study and the extent to which approved models meet those principles. Furthermore, it lists growth model features and the extent to which those features vary across approved models. Finally, it summarizes how including growth into AYP calculations has impacted the states with approved growth models and highlights next steps for growth-based accountability.

This document is organized around seven sections:

1. US ED pilot study events
2. Requirements and guidance for growth-based accountability models
3. Three basic growth model types approved by US ED
4. General descriptions of submitted models
5. Features of 11 approved growth models

6. Summary of approved growth models and how the growth models are used in adequate yearly progress (AYP) calculations
7. Impact of growth models on schools and districts
8. Conclusions and next steps in growth-based accountability models

In compiling this information, the Accountability Systems and Reporting (ASR) SCASS met two challenges. First, states' growth models have often evolved through the proposal and application processes. For example, several states were given feedback from the peer reviewers that resulted in changes to the growth models, and some states made changes to their growth models based on their experience applying the models. Second, documentation of the proposals, model features, and communications between states and US ED has not been consistent state to state. For some states, all documents, from initial proposal to final proposal have been easily accessible on the US ED website. For other states, only the initial proposal is available. Given these challenges, the authors of this paper have, to the extent possible, contacted state representatives to obtain information and reviews to promote accuracy in this document. However, the authors note that each state proposal is complex and in many instances went through several iterations to gain approval; moreover, states continuously update early growth model results resulting in variations in reported effects. It is also important to note that different sources report slightly nuanced results; e.g. results relating to schools "making AYP as a result of growth" vs. schools "making AYP based on growth." The former intends to demonstrate the marginal impact of adding a growth model to the existing accountability model, whereas the latter intends to demonstrate how many schools would make AYP based only on a growth model. For recommended changes or corrections to this document, please contact the authors or CCSSO.

US ED Pilot Study Events

From 2005-2008, the Department received three rounds of proposals and several resubmissions within each round. In the first round in February 2006, the Department received 20 proposals. Seven (Hawaii, Maryland, Nevada, New Hampshire, Ohio, Pennsylvania, and South Dakota) of the twenty proposals applied to evaluate growth in the subsequent year (i.e., the 2006-2007 school year), so they were not evaluated in the first round of reviews. Of the 20 first round submissions, eight (Alaska, Arizona, Arkansas, Delaware, Florida, North Carolina, Oregon, and Tennessee) were sent to be evaluated by a peer review group. The five states that submitted for the 2005-2006 school year but were not sent by US ED for peer review included Colorado, Indiana, Iowa, South Carolina, and Utah. Most of these states were not likely to have state assessment systems that would have been approved through the NCLB peer review process for the 2005-06 school year.

In May, 2006, the US ED approved the growth pilot proposals from North Carolina and Tennessee. These two states evaluated how including growth affected their federal accountability calculations for the 2005-2006 school year. In general, the addition of growth to the accountability system was not expected to and did not result in larger numbers of schools meeting AYP. The expectation was that for North Carolina, 40 of the 932 schools that missed AYP under the status model were expected to meet AYP due to the growth model. For Tennessee, 47 of the 353 schools that missed AYP under the status model were expected to meet AYP under the growth model. A September 2006 ED WEEK article reported that only 8 schools

in Tennessee met AYP due to the growth model. For North Carolina, the exact number of schools that met AYP due to growth was not reported, though the numbers were likely small or zero based on Lou Fabrizio's comment, "it's not helping much at all."

Of the six states reviewed by the peer reviewers in round one, five submitted revised proposals in September 2006. These included Alaska, Arkansas, Delaware, Florida, and Oregon. Arizona submitted a revised proposal in November. In November 2006, US ED announced that Arkansas, Delaware, and Florida received approval or conditional approval for growth-based accountability pilot projects for use in the 2006-2007 school year.

In November, 2006, nine states submitted proposals for round two of the pilot study and proposed to use a growth model in the 2006-2007 school year. At this time, only five of the ten potential pilot slots remained. The nine states with submissions included Arizona, California, Iowa, Hawaii, Ohio, Nevada, New Hampshire, Pennsylvania, and Utah. US ED sent all of these states' proposals for peer review in March 2007 except Utah's proposal. In May 2007, the five states that submitted additional information requested by reviewers included Alaska, Arizona, Iowa, New Hampshire, and Ohio. Then, US ED announced in May 2007 that Iowa and Ohio were approved (Ohio was conditionally approved), and US ED announced in July that Alaska and Arizona were approved. At this point, nine of the ten planned approvals had been granted. One spot remained.

In December 2007, before the final spot was filled, US ED opened the growth pilot project to all eligible states. In the third round, six proposals were submitted including those from the District of Columbia, Michigan, Minnesota, Missouri, New Mexico, and Pennsylvania. After the peer review process, US ED announced in May 2008 that Michigan and Missouri were approved for the growth pilot program.

Table 1. State Proposal Submissions and Those Leading to Approvals

Date	States
February 2006 (Round 1 Submission)	Alaska, Arizona, Arkansas, Colorado**, Delaware, Florida, Hawaii*, Indiana**, Iowa**, Maryland*, Nevada*, New Hampshire*, **North Carolina,** Ohio*, Oregon**, Pennsylvania*, South Carolina**, South Dakota*, **Tennessee**, Utah
September, 2006 (Round 1 Resubmission)	Alaska, **Arkansas**, **Delaware**, **Florida**, Oregon
November, 2006 (Round 2 Submission)	Alaska, Arizona, Hawaii, Iowa, Ohio, Nevada, New Hampshire, Pennsylvania, Utah**
May, 2007 (Round 2 Resubmission)	**Alaska**, **Arizona**, **Iowa**, New Hampshire, **Ohio**
February-April, 2008 (Round 3 Submission)	District of Columbia, **Michigan**, Minnesota, **Missouri**, New Mexico, Pennsylvania

Note. States listed in bold font indicate state submissions that led to approvals. *These states submitted proposals for a future year, so they were not evaluated in round in which they were submitted. **State proposals that were not sent for peer review.

Requirements and Guidance for Growth-Based Accountability Models

Since the 2005 pilot study announcement, the interpretation of the core principles guiding growth-based accountability models have evolved. For some principles, the trend has been that the interpretation of the principles has resulted in more flexibility whereas in other principles, it has resulted in less flexibility. For example, US ED initially was encouraging to states that did not count students in the numerator of AYP calculations if the students met proficiency but did not meet growth expectations. Tennessee, for example, has three options for meeting AYP, one using the status model, another using safe harbor, and the third that counts students in the numerator of AYP calculations for schools if students are projected to be proficient in a future year. If a student in Tennessee is proficient but not projected to meet proficiency, that student is not counted in the AYP numerator for that third AYP option. Later, states were granted approval by US ED, even though students in those states were counted in the AYP numerator if they met proficiency but were not projected to be proficient in a future year. The increased flexibility in the interpretation of some principles has likely contributed to more schools meeting AYP due to growth in states for which the growth models have introduced the additional flexibility. As mentioned, the first two states approved for the growth pilot, Tennessee and North Carolina, did not have many additional schools identified as meeting goals set by NCLB with the addition of a growth model. In contrast, Iowa which had a growth model approved a year later than Tennessee and North Carolina reported that more than 100 schools made AYP due to the Iowa growth model.

For other principles, just the opposite has happened. The interpretation of the principles has been applied with decreased flexibility over time. For example, US ED initially required that states not use large confidence intervals in their growth-based accountability systems. The confidence intervals were not limited to those around growth estimates. The confidence intervals that US ED examined were any that were used in AYP calculations. The initial requirement indicated that confidence intervals were allowed, though their use was limited. However, in the August 18, 2008 letter, US ED noted, “For example, based on the peers’ comments and concerns, we have concluded that wide confidence intervals are inappropriate in measuring individual student growth. In fact, we have not approved the use of any confidence intervals in the growth model pilot.” The changes in US ED application of this principle over time were to make the application more stringent.

The principles and guidance for including a growth measure in states’ federal accountability system come from several sources. The next sections summarize the principles and guidance from three sources: the alignment and foundational elements announced by Secretary Spellings in 2005 and repeated in US ED letters to states about growth, the cross-cutting themes reported by the first set of growth model pilot peer reviewers, and guidance from US ED reviews of state proposals in noted publications. None of the approved states have growth-based accountability systems that satisfy all of the principles and all of the guidance from the peer/expert group. In the next sections, each principle is listed and explained. In addition, information about flexibility in the application of the principle is provided.

NCLB Alignment and Foundational Elements

1. The accountability model must ensure that all students are proficient by 2013-14 and set annual goals to ensure that the achievement gap is closing for all groups of students.

Explanation. According to this principle, the growth-based accountability model must be designed so that all students are proficient by 2013-2014. The NCLB legislation specifically requires that, "each State shall establish a timeline for adequate yearly progress. The timeline shall ensure that not later than 12 years after the end of the 2001-2002 school year, all students in each group described in subparagraph (all students, economically disadvantaged students, students from major racial and ethnic groups, students with disabilities, and students with limited English proficiency) will meet or exceed the State's proficient level of academic achievement on the State assessments". Though this goal of NCLB has been a critical element of the NCLB legislation since the beginning, recent presentations by US ED staff about this aspect of the NCLB legislation have indicated that the requirement is now that "all students are learning at grade level by 2013-2014." Based on these recent presentations, the interpretation of this requirement has clearly shifted.

The shift in the interpretation of this element is also seen in regards to US ED communications around the growth models. In US ED decisions for approving models, this requirement has been applied with some flexibility. Though many of the 11 approved models have growth-based accountability systems that meet this requirement as originally phrased, many others require that students either meet proficiency by 2013-2014 or are found to be "on track" to meeting proficiency. For example, Missouri's recently approved proposal notes that the state, "proposes to incorporate a growth model calculation into its accountability system at grades 3-8, establishing unique growth trajectories that will ensure that, by 2014, all students will either be proficient or 'on-track to be proficient' by the end of grade 8, or within four years of the baseline score, whichever is reached first". Furthermore, in the September 23, 2008 NCLB growth model proposal peer recommendations, the guidance for this element is written such that growth model proposals should "facilitate adequate student progress over time towards the goal of 100% proficiency by the year 2013-2014." The phrase "towards the goal" serves as evidence of the evolving nature of the interpretation of this element.

2. The accountability model must not set expectations for annual achievement based upon student background and school characteristics.

Explanation. According to this principle, states should not include background or school variables in the growth models. The intent behind this principle was thought to be to promote the use of variables in growth models that schools, districts, and the state have the ability to change. Schools, districts, and the state have the ability to change student performance and school performance on state mandated tests, but they do not have the ability to change student-level background variables, like the socio-economic status (SES) of students. All of the approved growth-based accountability systems meet this requirement, as no states have proposed models that set growth expectations that differ based on students'

background characteristics (e.g., gender or ethnicity) or school characteristics (e.g., percent free and reduced lunch).

3. The accountability model must hold schools accountable for student achievement in reading/language arts and mathematics.

Explanation. All approved growth models have held schools accountable for student achievement in reading/language arts and mathematics. Some models like Ohio and Tennessee make student projections based on content area scores in subjects other than reading/language arts and mathematics, such as science, social studies, and writing. However, these other content-area scores in the Tennessee and Ohio models are used as predictors to enhance the accuracy of the student projections. In addition, Tennessee combines student performance in the two subjects in AYP calculations. Schools meet AYP under the projection model option in Tennessee if the percentages of students projected to meet proficiency are at or above annual AYP proficiency targets in ***both*** reading/language arts and mathematics.

4. The accountability model must ensure that all students in the tested grades are included in the assessment and accountability system. Schools and districts must be held accountable for the performance of student subgroups. The accountability model includes all schools and districts.

Explanation. All states with approved growth models have demonstrated their efforts to include as many students as possible in their assessment and accountability systems. However, due to the difficulty of tracking all students in a state, no state can demonstrate 100% participation in the assessment and accountability systems. US ED guidance for the growth proposal review process has shown increased emphasis on match rates and the inclusion of as many students as possible in the assessment and accountability systems. Specifically, in the September 23, 2008 document entitled, "Growth Model Proposal Peer Recommendations for the NCLB Growth Model Pilot Applications," seven recommendations are made about presenting match rates. These recommendations include suggestions about the types of evidence and data tables states should include in proposals. The added guidance reflects the increased emphasis US ED is placing on this element. One concern stems from the fact that even with very high participation rates in a single year, match rates tend to decrease over time, and often differentially for specific subgroups. A 90% match rate from one year to the next could result in a third of the students not matched after three years in a growth system.

US ED has also focused increased attention on the inclusion of students who take alternate assessments in the growth models. States initially approved, such as Tennessee and North Carolina, did not include students taking alternate assessments in the growth models. These two states included students taking alternate assessments in the accountability system based on the status models. More recently, in part due to US ED's emphasis on including students taking alternate assessments in the growth model, states are including students taking alternate assessments in their growth models. Evidence that US ED has been strongly encouraging states to include students taking alternate assessments in the growth models is found in the November 9, 2006 decision letter to Arkansas in which US ED wrote, "Arkansas' growth model does not currently include students

taking alternate assessments based on alternate achievement standards. The Department is very interested in improving the quality of these assessments and ensuring appropriate accountability for students taking them. Therefore, we would like Arkansas to work with the Department and its technical assistance provider on how to incorporate results from these assessments into the growth model for the 2007-08 school year." Similar encouragement was provided in the US ED decision letter to North Carolina.

In Arizona, for example, among students who take the alternate assessment, those taking AIMS-A are included in the growth model. Students who "move up" a performance level are considered to have met their growth target. Although the AIMS-A is not grade specific (i.e., the same tests are given to students regardless of grade), the scores required to meet proficiency increase by grade. In this respect, Arizona offers that movement from one performance level to another does in fact represent an improvement in student performance.

5. This foundational element was originally written in the 2005 announcement of the growth model pilot program to say, "The State's assessment system, the basis for the accountability model, must receive approval through the NCLB peer review process for the 2005-06 school year. In addition, the full NCLB assessment system in each of grades 3-8 and in high school in reading/language arts and math must have been in place for two testing cycles". The element as reworded in the August 18, 2008 US ED letter to chief state school officers was, "Include assessments that produce comparable results from grade to grade and year to year in grades three through eight and high school in both reading/language arts and mathematics, that have been operational for more than one year, and that have received* Full Approval *or* Full Approval with Recommendations *before the State determines AYP based on 2008-09 assessment results."

Explanation. All states needed to have approval for their assessment system before including a growth measure in the federal accountability system. However, some states obtained approval for their assessment system, implemented a growth-based accountability system, and then introduced new assessments. For example, Florida was approved to implement a growth model in their 2006-2007 AYP calculations. Florida received approval of their assessment system through the peer review process on June 27, 2007 with a full approval rating. However, in a June 11, 2008 letter from USED, Florida's new alternative assessment was not approved in the peer review process. US ED allowed Florida to continue to use their growth model for AYP results based on the 2007-2008 school year assessments even though their alternate assessment was not fully approved.

6. The accountability model and related State data system must track student progress.

Explanation. In order to implement a growth model, states needed to demonstrate that their data systems were robust enough to track students across grades. US ED emphasis on this element has increased since the inception of the growth model pilot program. The US ED guidance document from September 23, 2008 noted, "Effective growth models—that is, models that accurately target and capture changes in individual student achievement—are only possible if States are employing reliable and valid assessment systems that successfully track students

over time and across schools, such as by assigning unique identification numbers beginning with their entry into the State system." All states with approved growth models have systems in place for effectively tracking students and student data. As an example, Arizona has the ability to track students in grades K-12 through it's Student Accountability Information System (SAIS). Unique student identifiers have been in place since the 2001-02 school year, and a student's unique identifier will remain with the student across grades, schools, and districts until that student graduates from high school. The SAIS also links key demographic information (gender and race) and cohort membership (English language learners, students with disabilities, economically disadvantaged). In 2003-04, SAIS student identifiers were connected with Arizona's testing data. In Arizona's growth model pilot proposal, reference was made to a 90% match rate using 2004 and 2005 test data which did not differ by cohort membership or proficiency level.

7. The accountability model must include student participation rates in the state assessment system and student achievement on an additional academic indicator.

Explanation. All states with approved growth models have demonstrated the inclusion of student participation rates and student achievement on an additional academic indicator. Michigan, for example, proposed to continue to require schools and districts to meet the participation requirements related to all students in the tested grades. Michigan proposed to continue to use the other academic indicators of attendance rates for elementary and middle schools and graduation rates for high schools as required elements of AYP.

Summary of Alignment and Foundational Elements in Approved State Proposals

Table 2 summarizes the extent to which the 11 approved state proposals meet the alignment and foundational elements.

Table 2. Summary of Compliance with Alignment and Foundational Elements of 11 Approved State Models

	AK	AZ	AR	DE	FL	IA	MI	MO	NC	OH	TN
All proficient by 2013-14	1	1	1	✓	1	1	1	1	1	1	1
No expectations based on background characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Accountable for reading/ELA and mathematics separately	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
All students included in accountability system	2,6	2,4	2,4	2	5	2	2,3	2,4	2,3	2,3	2,3
Assessment system approved and in place for two years	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Student progress tracked	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Accountability model includes student participation and additional academic indicator	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes:
1 Includes on track to proficiency in 2013-14
2 Students without growth information (e.g., new to school) are included based on status.
3 All grade 3 and H.S. students are included based on status.
4 All grade 3, 8 and H.S. students are included based on status.
5 Grade 3 students (except those previously retained in grade 3) are included based on status.
6 All grade 3 students are included based on status.

Cross Cutting Themes

Additional guidance for the US ED growth proposal pilot program was provided on May 17, 2006, after the first set of growth peer reviewers completed their review of the round one submissions. The guidance came in the form of a document entitled, "U.S. Department of Education Summary by the Peer Review Team of April 2006 Review of Growth Model Proposals." The document described seven cross-cutting issues that the peer reviewers identified as important in their review of the first eight state proposals.

1. States should incorporate available years of existing achievement data, instead of relying on only two years of data.

Explanation. Though the growth peer reviewers encouraged the use of all available student data in the growth models, the states that have been approved by US ED to incorporate a growth measure in their accountability systems have typically not followed this guidance. Only two states have included all available student data in their projection/growth models. These states implement a multi-level regression model and include all available years of testing data as predictors in the model. States that implement a growth to proficiency type model tend to use only data from one year, the baseline year, in setting growth targets. These states use data from subsequent years to evaluate whether students meet growth targets or not. However, these states do not use data from multiple years to set growth targets. States that implement a value table or a transition table approach tend to focus on

student performance across two years, so these states rely typically on only two years of data.

As an example, the Arizona growth model calculates how much a student must grow per year to reach proficiency in three years (or by 8th grade, which ever comes first) but uses only the previous year's score to estimate a predicted score to determine whether the student truly meets the growth target. The rational Arizona provided in support of this approach is that this approach is designed to (1) include as many students in the growth calculation as possible where a continuous enrollment school and/or cohort is not necessary, and (2) to ease planning and understanding for teachers, administrators, and parents since it sets the same growth target for students scoring at the same level. Arizona offers that by setting different growth targets based on a stating point several years in the past would be confusing and perceived as unfair.

2. States should consider the impact on student growth trajectories of varying school configurations and of student movement between schools and districts.

Explanation. This aspect of growth models is one that states have applied in different ways. Most states with approved growth models, such as Alaska, Delaware, Florida, North Carolina, and Tennessee, set growth targets in a way that do not line up with typical school configurations. Alaska, for example, sets growth targets so that students with a starting point in grades 3-6 have four years to become proficient. For students with a starting point in grade 7 or higher, Alaska sets targets so the student is proficient in grade 10. Given the way Alaska's model works, student growth trajectories span typical school configurations. Nonetheless, schools are held accountable for student progress in the year in which the student is in the school. A few states have set growth targets in ways that more directly account for typical school configurations. Ohio, for example, uses a projection model such that students are counted as proficiency for AYP if they are on a path to reach proficiency by the end of the first year at their next school or within four years, whichever comes first.

3. States should make growth projections for all students, not just those below proficient.

Explanation. All states with approved growth models evaluate growth for all students and not just for students below proficient. Arizona's growth model, for example, holds schools and districts accountable for the academic growth of all students. This is different in comparison to the current status-only AYP method where schools have the incentive to only focus on the students who are near proficiency. As another example, Arkansas calculates growth thresholds for students above proficient so that these students participate in the growth model.

4. States should hold schools accountable for same subgroups as they did under status model.

Explanation. All states with approved growth models have continued to hold schools accountable for the same subgroups as they did under the status model.

5. States should not use wide confidence intervals (US ED has not approved the use of confidence intervals in any pilot proposal).

Explanation. As noted in the introduction to this section, this is one of the growth model principles that has received stricter interpretation over time. Initially, guidance from US ED suggested that a proposal could be accepted with some application of confidence intervals. However, more recent communications from US ED indicate that growth-based accountability proposals should not incorporate the use of confidence intervals. Several states received conditional approval for their pilot proposals, where one condition was that the state eliminate the use of confidence intervals. For example, Delaware initially proposed to incorporate the use of a confidence interval, but US ED would only approve the system without the confidence interval. In the November 9, 2006 letter from US ED informing Delaware that their proposal had been approved, Secretary Spellings wrote, "Delaware's proposal contemplates the use of a confidence interval in the growth calculation, although the proposal does not specify any details. As recommended by the peer reviewers, we are approving Delaware's growth model without the use of a confidence interval." US ED noted in their August 18, 2008 letter that, "In fact, we have not approved the use of any confidence intervals in the growth model pilot."

The Arizona model uses the lower bound of a confidence interval around a student's predicted score as a means of assuring the student's predicted true score is sufficient to meet the annual target. Despite Arizona's approved model including a confidence interval, the way in which it is applied is unique among growth models and is not used to accept the notion that a predicted score is statistically the same as the target. Instead, it is applied so that the predicted score can be considered significantly above the required target. This method intends to account for regression to the mean and measurement error. In addition, Missouri applies a confidence interval to status calculations, but Missouri does not apply a confidence interval to growth calculations.

6. States should not reset growth targets each year.

Explanation. This recommendation is one that has also been applied differently by states with approved growth models. Sates applying growth to proficiency type models tend to meet this requirement, as these states identify a base year for students, set targets for three or four years in the future, and hold students to those targets each year. Students are unable to have targets reset. For states applying a value table or transition table approach, growth targets are reset each year to some extent, given only two years of scores are evaluated each year. For example, suppose a student starts below proficiency in grade 3 and transitions to a higher sublevel in grade 4. That student would add a positive value to a value table or meet growth targets in a transition table. Then suppose that student drops to a lower level in grade five. That student in grade 5 would not add a positive value to the value table or meet growth targets in a transition table. It is possible, however, that the student again demonstrates improved scores from grade 5 to grade 6. That student would again receive credit for growth even though the student has not

made progress overall across the years. For states applying the regression-based projection models, growth targets are reset each year. Though the initial peer reviewers recommended that growth targets not be reset each year, they were the peer review group that approved Tennessee's model which does reset students' growth targets each year. In Tennessee, students are evaluated for growth by projecting students' performance three years in the future. A student receives a new projection each year.

7. States should not average scores between proficient and non-proficient students.

Explanation. This feature was intended to prevent growth from one student to compensate for lack of growth from another student. The way in which this feature has been applied by states has varied. States implementing growth to proficiency type models, transition type models, or projection models have typically met this requirement, as each student receives an individual growth decision that is not averaged with other student's. States implementing value tables combine values of proficient and non-proficient students, as the score for a campus is the average of students' scores.

Other Guidance

1. States should carefully justify the implementation of a growth model in addition to an index system.

Explanation. There has been little formal guidance from US ED on the extent to which states can implement a growth model in addition to an index system. In Secretary Spellings' November 18, 2005 initial letter to chief state school officers announcing the growth model pilot program, US ED encouraged states to consider implementation of an index. The letter suggested that an index could be implemented during the time when states developed the necessary elements for a growth model. In one of the attachments to the letter, it states, "An index model provides a way for States that cannot track individual student progress to implement a form of accountability that captures subgroup growth, thus providing an alternative to the status model of AYP decision-making that only gives a school credit for the percentage of students at the proficient or above level."

A few states have been successful in obtaining approvals to implement both a growth model and an index system. Michigan, for example, was approved to implement a transition table growth model in addition to a Proficiency Index for AYP calculations. The Proficiency Index that Michigan implements is used to combine grade-level proficiency information and targets to make AYP decisions across grades. In addition, Pennsylvania was approved to implement a projection growth model in addition to using the Pennsylvania Performance Index in 2009 AYP calculations. Other states have not been approved to use a growth measure in addition to a performance index. The June 10, 2008 US ED letter to Minnesota announcing that Minnesota's proposal was not approved states, "The heart of the peer's concerns relates to the interaction of Minnesota's existing performance index with the value table. The Department has concerns about the appropriateness of allowing a state to include both a performance index and a growth model in its accountability system." The letter goes on to state that US ED would bring this

issue to the National TAC and provide guidance to states in summer 2008. The issue is that peers have been concerned with the potential lack of integration of an index system and a growth model, and the (often) result of doubling the number of ways in which schools might make AYP.

Since indexes and growth measures both allow states flexibility in AYP calculations, states will likely have to provide evidence showing that the implementation of both of these measures does not violate NCLB core principles. In addition, evidence suggests that US ED has been focusing and gathering input from experts on this issue. The October 2008 Department of Education Federal Register publication states, "…the appropriate use of confidence intervals and performance indexes in determining AYP are issues that would benefit from immediate consideration by the National TAC." More formal guidance on this issue might be published by US ED, once US ED obtains recommendations from the National TAC.

2. States must determine a uniform minimum group size for reporting AYP for student groups.

Explanation. The minimum group size was not a feature that was emphasized in the original growth model pilot announcement. The 2005 growth model pilot program announcement required that states report their minimum group size and explain how the minimum size would continue in a state AYP definition that incorporates growth. In the 2006 cross-cutting themes document, the peer reviewers stressed the importance of match rates and group sizes. They stressed that groups of students on whom AYP determinations are based need to be representative of the particular subgroup of students at the school. The September 23, 2008 US ED peer recommendations did not emphasize the importance of minimum group sizes more than in the 2006 document. Though the minimum n-size issue has not been stressed in guidance documents, this feature has appeared to be an important one in approvals for states. For example, the US ED June 10, 2008 decision letter to Missouri stated that Missouri's approval to include a growth measure into AYP calculations was conditional on "…Missouri's adopting a uniform minimum group size for all students in the State, including students with disabilities and limited English proficient students." A similar condition was noted in the approval letter for Ohio.

3. States should set the number of years a student has to reach proficiency at a reasonable amount.

Explanation. US ED has not limited the number of years a student can score below proficiency but count for AYP due to growth to a specific number of years. In the November 18, 2005 US ED letter announcing the growth pilot program, guidance on the first core principle included a statement about the Department allowing models that set a point in time as the goal for growth targets. The examples the letter noted that growth targets might be set at the end of a grade in a particular school or within four years. Though there is not a set number of years for growth targets, most states set growth targets such that the maximum number of years for a growth target is at three or four years. For some examples, the maximum number of years a student can count for AYP without meeting proficiency is two years for Ohio. For Alaska, Arizona, Florida, and Michigan, it is three years. For North Carolina, Iowa, and Missouri, it is four years. For Arkansas and Tennessee, the

maximum number of years is five and six, respectively. For Delaware, there is not a set number of years for students, given that student's changes in performance levels allow students to earn points each year.

Three Basic Growth Model Types

The eleven approved growth models can be categorized into three basic types of growth models, though models in each type vary in the specific ways in which they are implemented. The three types include:

1. Growth to Proficiency Models
2. Value Tables/Transition Models
3. Projection Models

Growth to proficiency models, also called growth to standards models and trajectory models, are growth models designed to inform about whether students are on track to meeting the proficiency standard in some specified point in the future. The key element is that these models work backwards from a proficient score in some future grade and then divide required student gains into annual pieces. The number of years states specify varies, but is typically 3 or 4 years. When these models are included in AYP calculations, schools are given credit for students who have not yet passed but are making score gains such that the students will pass by the third or fourth year if score gains continue. Of all approved growth models, this type has been most popular. Six states with growth to proficiency models have been approved for the pilot program including Alaska, Arizona[1], Arkansas, Florida, Missouri, and North Carolina.

Value tables and transition models are growth models that evaluate student transitions across performance levels or subdivisions of performance levels. These models give credit to schools for moving students into higher levels or sublevels during the school year. The focus of these models tends to be on student changes typically over 2 years. When using these models, states will subdivide performance levels and will expect students to progress across performance levels in such a way that students reach proficiency in a set number of years (typically 3 or 4). States implementing these models tend to report student growth at the campus, district, and state level. They do not always report growth at the student level. States approved for this type of model include Delaware, Iowa, and Michigan.

Projection models predict or project student performance into the future. Credit with these growth models is given to schools for students who have not met proficiency yet are predicted or projected to reach proficiency in the future (typically no more than 3 or 4 years). The key element is that these models project student performance based on past performance and the performance of prior cohorts in the target grades and compare this projection to the proficiency standard for the target grade.[2] These types of models tend to be more statistically complex than the other types of models approved by US ED for the pilot program, given

[1] Arizona uses both a regression model to generate a predicted score and a fixed annual target in applying it growth model.

[2] The difference between a Growth to Proficiency Model and a Projection Model is that the former uses only a student's current status and the future grade level proficiency score to determine growth, whereas the latter uses a student's current status and the past typical average growth of a previous cohort that already reached the target grade to determine growth.

they typically involve multi-level regression equations. States with approved growth models that are of this type include Ohio and Tennessee.

General Descriptions of Submitted Models

The table below provides a general description of the growth models from states that submitted proposals for the growth model pilot program. In addition, for all states that submitted proposals, either the approved status is listed or reasons are provided for why the proposed models were not approved.

Table 3. Summary of Submitted Growth Proposals and US ED Decision

STATE	GENERAL GROWTH MODEL DESCRIPTION	APPROVED/REASON(S) REJECTED
Alaska	To calculate the growth determination, Alaska determines which students are on track to be proficient within four additional years for those in grade 4-6, three additional years in grade 7, two additional years in grade 8, and one additional year in grade 9. Students in grade 10 will be evaluated based on status as are students in grade 3. Students in grades 4-6 who are in the LEA or state for the first year, who made a 25% gain from the previous year, and who are on track to be proficient within three additional years will be considered proficient. Once a student has been in the system for a third year they will have to demonstrate a one-third gain, one half gain the next year, and, finally, they will be required to be proficient by the fourth year, but no later than grade 10.	Approved
Arizona	In Arizona, the growth model includes growth targets and predicted scores. The growth target part of the model requires that students reach proficiency within three years or by the eighth grade, whichever comes first. Arizona subtracts a student's current year scale score from the scale score for proficiency in the target grade and divides by the number of remaining grades. Regression analyses are used to create the prediction equation using current and previous year scale scores. The predicted score is compared with the target score in the predicted year.	Approved
Arkansas	Arkansas implements nonlinear growth trajectories for students in grades 4-8 with the expectation that students will reach proficiency by eighth grade. Growth increments required to reach proficiency vary across the years.	Approved
Colorado	Colorado implements a quantile regression growth model with the expectation that all students will reach proficiency within 3 years or by 10th grade. The quantile regression growth model is combined with all available prior test	Colorado proposal was not sent to peer reviewers when initially submitted in February 2006.

	scores to determine whether each student is on track to be proficient or on track to maintain proficient within 3 years or by 10th grade, whichever comes first.	
Delaware	Delaware implements a value table approach to growth-based accountability. Each student in a subgroup earns points for moving across proficiency levels. The number of points increases as students move from levels below proficiency to levels at proficiency. The points in the value tables were set by committee.	Approved
District of Columbia	The District of Columbia implements a vertical scale. The growth model trajectory targets are set for each grade level based on the grade that a student enters a school, the target grade, and the proficiency cut score for that target grade. The expected annual growth targets are calculated by dividing the difference between a student's baseline scaled score and the ending scaled-score target (cut score) by the number of steps. For most students, the number of steps is three. For students with a baseline year in grade 6, the number of steps is four, and students in grade 8 are expected to reach proficiency by grade 10 (in one step). A student is "on trajectory" if the student's observed score is at or above the target. Growth targets are reset when students successfully complete all grades in one school and move to another, such as matriculating from elementary school to middle school.	➢ Maturity of assessment and tracking system ➢ Extent to which scale and standard-setting support proposed growth model ➢ Resetting of growth for transfer students (move to middle school) was not found to be appropriate
Florida	Florida calculates students' average annual projected growth rate by taking the difference between students' current scale scores and students' first scale score and dividing the difference by the number of years the student has been in school over that time.	Approved
Hawaii	Hawaii proposes to use a linear mixed model to allow probabilistic estimates based on a three-year projected growth curve for each student in each content area using that student's achievement scores. Hawaii also uses piecewise linear regression to compare student growth rates that occurred under two distinct educational periods: elementary school and middle school.	➢ Assessment system was not fully approved; concern centered on major revisions to be required in alternate assessment ➢ Longitudinal cohort match rates deemed insufficient
Indiana	Indiana's proposed growth model focuses on cohorts of students over two years. Under the model, groups of students would meet AYP if the reduction in students scoring below proficient from the previous school year to the current school year meets the safe harbor target.	
Iowa	Iowa divides the scale score range below proficient into three categories and has established category boundaries on the scale	Approved

	score system for non-proficient students across grades. A student's growth trajectory must cross a category boundary in order to count for Adequate Yearly Growth.	
Maryland	Maryland submitted a proposal explaining that the state planned to develop a growth model similar to models that would be approved for the 2005-2006 school year. The model would require all students to meet proficiency in no more than three years.	➢ Data systems for tracking students year to year not fully developed until a later year
Michigan	The Michigan growth model divides each of the assessment performance levels (not proficient, partially proficient, proficient, and advanced) into three sub-levels (low, middle, and high), and tracks students transitions from one year to the next (e.g. from the middle of the not proficient category in grade 3 to the top of the partially proficient category the next year in grade 4).	Approved
Minnesota	Minnesota proposed to implement a value table growth model, where the achievement levels below proficiency are subdivided into two groups (i.e., "low" and "high"). Each student in a subgroup earns points for moving from a lower achievement level or sublevel to a higher achievement level or sublevel during one year. The points increase as the discrepancy between the two levels increases and the highest point in this table is 100. If the student makes two consecutive years of growth, then the students is eligible for additional points, called compounding points, which is one-half the difference in the points for the next highest performance range. At a campus level, students' average points are used in the accountability system.	➢ Concerns about the compounding growth ➢ Insufficient rationale for the values ➢ Concerns about the interaction between Minnesota's index system and the proposed growth model
Missouri	Missouri calculates the student's annual growth target by taking the difference between the student's base year scale score and the student's proficiency goal scale score (in four years or by grade 8). The difference is divided by the number of years the student has to become proficient. This average is added to the student's baseline score each year to determine the student's annual growth target.	Approved
Nevada	Nevada proposed a value table approach to growth-based accountability. Each student in a subgroup earns points for moving across the six levels: the lowest of the four proficiency levels will be split into two levels. The number of points will increase as students move from the lowest level (Emergent Low) to the highest level (Exceeds). The points in the value tables are planned to be set by committee.	➢ Values awarded to student growth were questioned ➢ Student performance above proficiency could compensate for student performance below proficiency ➢ Rationale for excluding students with severe cognitive disabilities needed ➢ Concerns that data systems would not accurately track students over time

New Hampshire	New Hampshire sets growth targets for students based on the number of base-year standard deviation units a student is below proficient. Students are expected to reach proficiency in 1-3 years.	➢ Concerns about the technical structure of the proposal ➢ Inconsistencies in proposal text and formulas
New Mexico	New Mexico implements the Individual Student Academic Change (ISAC) growth model. This model tracks students who are non-proficient in the first year and compares student gains from the first to the second year with gains that a student would need to make in order to meet proficiency in 2013-2014.	➢ Lack of detail around development of growth intervals ➢ Lack of reporting to parents and the public ➢ No subgroup disaggregation ➢ Concerns that some students' performance might compensate for the performance of other students ➢ Insufficient data systems
North Carolina	North Carolina uses a time-locked modified z-scale to evaluate whether students are on track to become proficient in four years (or less). At the onset of the trajectory growth calculations, students in grade 3 who are not proficient are expected to decrease by 25% each year the difference between the grade 3 pre-test results (on the change scale) and the expected proficiency at the end of grade 6 four years later. Other grade levels follow a similar pattern, however the number of steps to proficiency decrease gradually and the "percent difference closed per step" increases, as the student progresses academically to grade 8.	Approved
Ohio	Ohio uses a projection model such that students are counted as proficient for AYP if they are on a path to reach proficiency within two years (with the exception of 7th grade, where students must reach proficiency by 8th grade).	Approved
Oregon	Oregon uses students' highest test score in each year to fit a three level, hierarchical linear growth model. The model calculates a linear growth trajectory for students in grades 3-8 and 10. Students are considered on track for proficiency if their growth trajectories project that they will be proficiency three years in the future.	➢ Concerns that planned changes to the assessment system would not support the proposed growth model
Pennsylvania	Pennsylvania implements a projection model that uses all available math and reading achievement data to project a student's math or reading performance in 1-3 years in the future. Students are included if they have at least 3 data points, i.e.,each test scores is a data point.	➢ Model was acceptable but proposal was denied due to Pennsylvania's insistence that they retain their previously approved Pennsylvania Performance Index. ➢ The model would have been approved if Pennsylvania had agreed to discontinue use of their index.
South Carolina	The proposed South Carolina model implemented a two-part safe harbor	➢ Concerns that the state had a higher minimum group size for

	component. Part 1 of safe harbor tracked student scores from one year to the next to determine whether a sufficient number of students in the cohort scored proficient in the second year for there to be a ten percent reduction in students not scoring proficient from the prior year. Part 2 of the growth safe harbor used longitudinally-matched student data and tracked students' performance by scale scores for two years. The students' scale scores were converted to point weights drawn from conversion tables, and the weights were averaged across all students and compared to the previous year. The difference between the two years was compared to "rating criteria."	the students with disabilities and LEP subgroup ➢ The coherence of the accountability system was a concern when a growth model was added to the existing structure ➢ The model did not appear to require 100 percent proficiency by 2014
South Dakota	South Dakota's proposed model was not intended by the state to be implemented and used in AYP determinations in the 2005-06 school year, given state was making significant changes to its assessments. The proposed growth measure looked at the average "gain" toward proficiency for students from one year to the next, where gain for students not proficient is defined as a decrease in the number of scale score points from the proficiency line year after year.	➢ State did not have two years of assessment data for grades 3-8 and 10 ➢ It was unclear how growth would be applied to students meeting proficiency
Tennessee	Tennessee implements a projection model that uses all available past scores to project if a student will score proficient or advanced on the statewide assessment three years in the future. Growth is projected for students in grades 3 - 8 who have at least one previous regular test score in the subject. Other students (students in grade 3 and high school, students with only the current year's score, and students who take the alternate assessment) do not get growth projections and are included based on their proficiency in the current year. The projection method is based on linear regression with the assumption that students will receive the average Tennessee schooling experience in the future. Individual student projections are based on all available test scores for the student in each subject and assume the student will receive future instruction of average effectiveness.	Approved
Utah	Utah implements a progress method that results in a progress category of high, medium, or low based on a weighted composite of several indicators: language arts performance, mathematics performance, science performance, attendance, and graduation rate (for high school only). Student performance on language arts, mathematics, and science, defined as their year-to-year performance category change on the state test, is mapped to a value from a value table.	➢ The proposed system did not require all students to be proficient by 2013-2014 ➢ The proposed system was compensatory, as performance across all AYP subjects are combined and not in separate categories ➢ The proposed method for holding subgroups accountable was not approved by US ED

Features of 11 Approved Growth Models

In this section, practical, psychometric, and accountability features of the approved growth model proposals are summarized. The features described below are those that are not directly related to the guidance for the growth pilot program, so these features are not already listed in the paper. Explanations of the features are provided initially, followed by summary tables with features on specific state proposals that were approved. Only the 11 state proposals approved in 2005 to 2008 are included here. The states in the summary tables are organized according to when they received US ED approval. For example, the first two columns compare and contrast features of the North Carolina and Tennessee proposals, as these were the first two states approved by US ED.

Submission leading to approval—This row indicates the pilot proposal submission month and year that led to US ED approval.

Growth measure name—For states that have named their measure of growth, the name is listed.

First year all AYP grades tested in reading and math—The years listed in this field indicate the first year that the state assessed students in the NCLB grades in reading and mathematics.

Grades for which growth calculations made for students—This row lists the grades for which growth calculations are made for any students or student progress is used for growth at the campus level. The grades listed are only those for which growth or projections are made. A grade would not be listed if scores from that grade are used in the growth calculations, but for which students are not evaluated for growth or progress. As an example, if grade 3 scores are used to evaluate growth in grade 4 but students are not evaluated for growth in grade 3, grade 3 would not be listed. However, if student projections are made for grade 3 students in a state, grade 3 would be listed for that state. In addition, for models in which growth is not reported at the student level (e.g., Delaware), these grades indicate the grades in which student progress is used in evaluating growth for the campus.

Maximum number of years any student can meet growth when below proficiency—In this row, the number of years (not necessarily consecutive) in which a student can meet growth expectations even when that student is below proficiency is indicated. For example, if a state implements a projection model in which students are given a projection each year, and students can be below proficiency yet still meet growth expectations (projected to be proficient at set point in future) for all grades 3-8, that state would have 6 years noted in the table. As another example, a state would have 4 listed in this table if that state allows students to count for growth each time the students cross a performance division and students must cross 4 performance divisions in 6 years to be proficient.

Growth only for below proficient students—This field indicates if state growth models are only applied to students who are below proficient. For states that track growth for students at and above proficiency, this field will have a "No."

Growth possible for grade 3 students—States that calculate growth for grade 3 students are noted in this field. As an example, North Carolina calculates growth for grade 3 students using a pretest given at the beginning of grade 3.

Growth for students w/o prior year score—States that measure growth for students who are missing the prior year test score are noted in this field. As examples, states may calculate growth for students without the prior year score by using scores from more than one year prior, the mean of the grade-level scores the prior year, a baseline score from two or more years before, or a pretest score.

Growth for SWD taking alternate assessment—For states that have an alternate assessment for students with disabilities, this field indicates whether the proposed growth model or a different growth model can be applied to scores on that alternate assessment. In other words, if students taking an alternate assessment will have growth calculated, this field will indicate "yes." However, if students taking an alternate assessment are only evaluated for status, this field will indicate "no."

Growth for ELL Taking Alternate Assessment—For states that have an alternate assessment for English language learners, this field indicates whether the proposed growth model or a different growth model can be applied and used in AYP calculations to scores on that alternate assessment. In other words, if students taking an ELL assessment will have growth calculated and that growth can be included in AYP calculations, this field will indicate "yes." However, if students taking an alternate assessment are only evaluated for status, this field will indicate "no."

Growth reported at student level —This field indicates whether growth at the student level is reported. For some states, student-level growth is calculated but not reported, as the student-level growth is used in calculating campus-level AYP. For these states, this field will indicate "no."

Growth reports limited to grades 3-8—This field indicates if states have proposed using student/campus growth for AYP only in grades 3-8. For example, a state that uses a projection model in grades 3-8 and only reports student projections in grades 3-8, even if that model projects grade 11 from grade 8, will have a "yes" in this field.

Growth calculations use all prior-year data— This field indicates whether a state uses all prior-year data in the growth calculations. A state that uses all prior-year data in a given subject will have a "yes" indicated. For example, a state that uses all of a student's reading scores in grades 3-5 to evaluate that student's reading growth in grade 5 would have a "yes" in this field. If a state uses only scores from one or two years, that state would have a "no" indicated.

Vertically aligned standards—This field indicates whether the state vertically aligned its academic achievement standards. Though most state proposals have indicated that their academic achievement standards have been vertically aligned, the ways in which this was accomplished and the extent to which the standards have been aligned varies. This table does not indicate details about how the standards were aligned, only if they were aligned.

Vertical scale in any grades for which growth calculated—This field indicates whether the state has implemented a vertical scale in any grades in which growth is calculated. If a state has a vertical scale in a grade for which growth is calculated, but the vertical scale is not used in the growth calculations, a footnote is provided.

Confidence interval for growth—This field indicates if the state will use a confidence interval in growth calculations. If states propose using a confidence interval around the percent of students reaching proficiency by having met growth targets for AYP purposes, these states will have a "yes" in this field.

Growth targets reset—This field indicates if states ever reset students' growth targets. For states in which students have a baseline year and a set number of years to meet proficiency, this field would indicate "no." For states in which projections are made every year, this field will indicate "yes."

Target timeline Aligns with Grade Configuration—This field indicates if the growth target or projection endpoint matches with the typical grade configurations in the state. For example, if students below proficiency are expected to reach proficiency in three years, does the three years end at grades that a typical school unit serves, such as grade 8 in states in which middle schools serve grades 6-8.

Students above proficiency with no/negative growth counted in AYP—This field indicates if students who are proficient in a current year but who are not meeting growth targets will count for AYP. For example, if states count students in the numerator of AYP calculations if they meet status OR if they meet growth, this field would indicate "yes." In other words, a state with a "yes" would allow a student who met proficiency standards but did not meet growth expectations to count in the numerator of the AYP proficiency ratio. A state with a "no" would only allow students to count in the numerator of the AYP proficiency ratio if students meet status and growth criteria.

Averaging of calculations for AYP—This field identifies states that average calculations over more than one year or over students within subgroups for AYP.

Minimum n same as for AYP status model—This field identifies states that apply a minimum sample size rule to the growth model. For example, many states designate a minimum n (or sample size), such that the state does not include a subgroup in AYP calculations or growth model calculations if that subgroup has fewer students than the minimum sample size.

Growth in addition to status and safe harbor provisions for AYP—This field indicates states that use growth as an additional method for meeting AYP after applying the status model and the safe harbor provisions.

Table 4a. Summary of Practical, Psychometric, and Accountability Features of Approved Growth-Based Accountability Models

	State	North Carolina	Tennessee	Arkansas	Delaware	Florida
PRACTICAL	Submission leading to approval	Feb. 2006	Feb. 2006	Sept. 2006	Sept. 2006	Sept. 2006
	Growth measure name	ABCs	N/A	N/A	N/A	N/A
	First year all AYP grades tested in reading and math	1992-93	2001-2002	2005-2006	2001-2002	2000-2001
	Grades for which growth calculations made for students[1]	3 - 8	3 - 8	4-8	3-10	4-10
	Maximum number years any student can meet growth when below proficiency[2]	4	6	5	N/A	3
	Growth only for below proficient students	Yes	No	No	No	No
	Growth possible for grade 3 students	Yes, pretest	No	No	Yes	No
	Growth for students w/o prior year score	Yes[3]	No	No	No	No
	Growth for SWD taking 2% alternate assessment	No	No	Yes	Yes	Yes
	Growth for SWD taking 1% alternate assessment	No	No	No	No	Yes
	Growth for ELL taking alternate assessment	No	No	No	Yes	Yes
	Growth reported at student level	No	Yes	Yes	No	Yes
	Growth reports limited to grades 3-8	Yes	Yes	4-8	3-10	4-10
PSYCHOMETRIC	Growth calculations use all prior-year data	No	Yes	No	No	No
	Vertically aligned standards	Yes	Yes	Yes	Yes	Yes
	Vertical scale in any grades for which growth calculated	Yes[4]	Yes[4]	Yes, Grd 3-8	No	Yes
	Confidence interval for growth	No	95%	No	No	No
	Growth targets reset	No	Yes	No	Yes	No
	Target timeline aligns w/ grade configuration	No	Yes	No	Yes	No
	Growth model type	Growth to Prof.	Projection Model	Growth to Prof	Value Table	Growth to Prof.
ACCOUNTABILITY	Students above proficiency with no/negative growth counted in AYP	Yes	No	Yes	It is possible	Yes
	Averaging of calculations for AYP	Yes	No	No	Yes	No
	Minimum n same as for AYP status model	Yes	Yes	Yes	Yes	Yes
	Growth in addition to status and safe harbor provisions for AYP	Yes	Yes	Yes	Yes	Yes

Notes. [1]Only grades for which growth is calculated, not grades for which data used, [2]Year not counted if growth target is proficiency that year [3] pretest scores used, [4]The growth calculations do not use the vertical scale

Table 4b. Summary of Practical, Psychometric, and Accountability Features of Approved Growth-Based Accountability Models

	State	Alaska	Arizona	Iowa	Ohio	Michigan	Missouri
PRACTICAL	Final Approval granted	May 2007	May 2007	May 2007	August 2007	April 2008	April 2008
	Growth measure name	Alaska's AYP Growth Model 2004-2005 for grades 3-9; 2005-2006 for grade 10	N/A	N/A	N/A	N/A	N/A
	First year all AYP grades tested in reading and math		2004-05	2005-2006	2005-2006	Yes	2005-2006
	Grades for which growth calculations made for students[1]	4-9	4-7	3-8	4-8	Yes	3-8
	Maximum number of years any student can meet growth when below proficiency[2]	3	3; end of gr. 8	4	2	3	4; end of gr. 8
	Growth only for below proficient students	Yes	No	No	Yes	No	No
	Growth possible for grade 3 students	No	No	No	No	No	No
	Growth for students w/o prior year score	No	Yes	No	Yes	Yes	No
	Growth for SWD taking 2% alternate assessment	No	No	No	N/A	N/A	Yes
	Growth for SWD taking 1% alternate assessment	No	Yes	No	No	No	N/A
	Growth for ELL taking alternate assessment	No	Yes	No	N/A	Yes	Yes
	Growth reported at student level	No	Yes	No	No	Yes	Yes
	Growth reports limited to grades 3-8	No	No	Yes	Limited to grades 4-8	Yes	Yes
PSYCHOMETRIC	Growth calculations use all prior-year data	No	No	No	Yes	No	No
	Vertically aligned standards	Yes	Yes	Yes	Yes	Yes	Yes
	Vertical scale in any grades for which growth calculated	No	Yes	Yes	No	Yes	Yes
	Confidence interval for growth	No	Yes	No	No	No	No
	Growth targets reset	No	Yes	No	Yes	Yes	Yes
	Target timeline aligns w/ grade configuration	Yes	Yes	No	Yes	Yes	No
	Growth model type	Growth to Prof.	Growth to Prof.	Transition Model	Projection Model	Transition Model	Growth to Prof.

ACCOUNTABILITY	Students above proficiency with no/negative growth counted in AYP	Yes	Yes	Yes	Yes	Yes	Yes
	Averaging of calculations for AYP	No	No	2 or 3 years	Yes	Yes	Yes
	Minimum n same as for AYP status model	Yes	Yes	Yes	Yes	Yes	Yes
	Growth in addition to status and safe harbor provisions for AYP	Yes	Yes	Yes	Yes	Yes	Yes

Notes. [1]Only grades for which growth is calculated, not grades for which data used, [2]Year not counted if growth target is proficiency that year [3] pretest scores used, [4]The growth calculations do not use the vertical scale

Approved Growth Models

The following pages summarize the growth models approved for the eleven states and ways in which growth is used in these states' accountability models.

Alaska

Growth Model Description: Alaska administers annual assessments in grades 3-10 in the content areas of reading, writing, and mathematics. Reading and Writing scores are combined to yield a Language Arts score. Alaska also administers annual assessments in grades 4, 8, and 10 in science. In Alaska, a score of 300 or above on the statewide assessment defines proficiency for each content area. Scores on these state assessments are not on vertical scales. For students with scores below 300 in mathematics, growth targets are set. For students with scores below 300 in language arts, growth targets are set. The targets depend on how far a student's score is below the proficiency cut score, the student's grade, and the results from the student's base testing year. The first year that a student is tested is considered the student's "base year," and the student's scaled score on that test is the student's "base score."

Using the student's base score, a student will be assigned a "target score" to be achieved each of the subsequent years the student has to become proficient. If the student's observed scaled score on the Standards Bases Assessment (SBA) is equal to or higher than the target score, and equal to or higher than the score from the previous grade level, the student will be considered to be "on track to becoming proficient" for that school year. If the observed scaled score on the SBA is less than the target or less than the score from the previous grade level, the student will be considered to not be on track and therefore will not count positively for his/her school. The target score will be calculated by first estimating the student's true score (using classical measurement theory) for the base year. Making those calculations requires the grand mean for the state and the reliability of the SBA taken in that base year.

A student's estimated true score (ETS) is calculated as follows:
ETS = Grand Mean + Reliability * (Observed Score – Grand Mean)

If the student's base year is grades 3-6, the student is given four years to become proficient. If the base year is grade 7 or higher, then the student is given the difference between the base year and 10; for example, if the student's base year is grade 7, the student is given 3 years to become proficient, and if the student's base year is grade 9, the student is given one year to become proficient. Students must be proficient by the end of grade 10 to count positively for their school.

As an example, a student who starts in an LEA in grade 3 and scores 200, which is below the grade 3 proficiency score of 300, will have AYP determined by the growth model. Given the state mean for Math is 355, the student's score of 200 is converted to an estimated true score: ETS = 355 + .91 * (200 – 355), or 214.

The increments for the above student are calculated as follows: proficiency score (300) minus the ETS (214) divided by the number of years to proficient (4) = 21.5.

For this student, the following growth targets apply:

- Grade 4: 235
- Grade 5: 257
- Grade 6: 278
- Grade 7: the student's scale score is expected to be at proficiency, or 300.

Student Growth in AYP Calculations: The Alaska AYP growth model first evaluates how many subgroups in a campus and district meet minimum n requirements. For those subgroups and schools, participation requirements are then evaluated. Then, subgroups and schools are evaluated on the performance requirement, meaning the percentages of students meeting proficiency or on track to meeting proficiency under the growth model separately in language arts and mathematics are compared with the annual measurable objectives. The comparison includes use of a confidence interval. Then, schools are evaluated based on the other academic indicator. If schools and subgroups do not meet the performance requirement, Alaska implements an improvement or safe harbor provision.

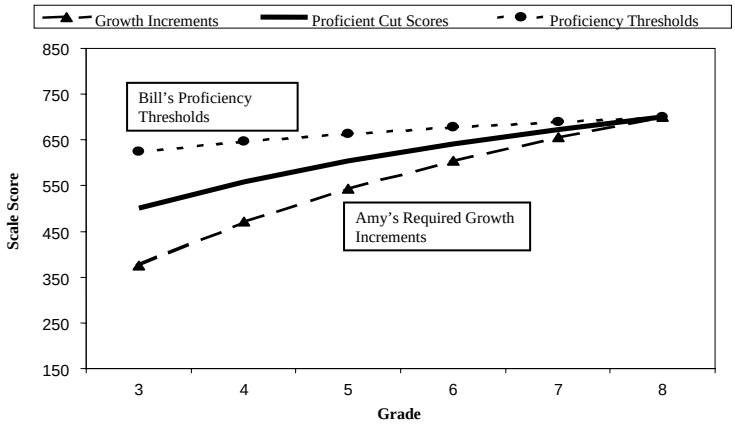
Arizona

Growth Model Description: Arizona tests students annually in reading and mathematics in grades 3-8 and 10. The statewide test in Arizona has a vertical scale in grades 3-8 reading and mathematics, and these grades and subjects are used in the growth model. In Arizona, the growth model includes growth targets and predicted scores. The growth target part of the model requires that students reach proficiency within three years or by the eighth grade, whichever comes first. To calculate student growth targets, Arizona subtracts a student's current year scale score from the scale score for proficiency three grades later and divides by the number of remaining grades. As an example (note that information was adapted from Arizona's proposal, which has been posted at http://www.ed.gov/admins/lead/account/growthmodel/az/azgmpadd.doc), suppose a student scores 362 on the 3rd grade math test. The proficient cut score on the 6th grade math test is 496. The student's math score must improve 45 points each year—(496-362)/(6-3) = 45—for her/him to reach proficiency by 6th grade. A regression is used to estimate a predicted score for each student based on previous grade 3 to 4 growth in each school. The predicted score for the above student is 417—this score is adjusted to the lower bound of the 97.5th percentile using the standard error of the prediction and equals 409. In order to count as proficient this student's lower bound of the predicted score in 4th grade must be at least (362+45) 407. Given the lower bound of the predicted score is 409 and the target score is 407, this student made AYP.

Student Growth in AYP Calculations: For schools and subgroups to meet AYP proficiency targets, the number of students meeting growth targets is added to the number of students who are proficient but who did not meet growth targets. Then, this sum is divided by all students in the analysis. This proportion is compared with the proficiency target for the year. If the value meets or exceeds the target, that school or subgroup meets the AYP performance requirements.

Arkansas

Growth Model Description: Arkansas administers language arts and mathematics exams in grades 3-8, and scores on these exams are on vertical scales. Arkansas also administers a literacy exam in grade 11 and an algebra I and geometry end-of-course exam. Arkansas implements nonlinear growth trajectories for students in grades 4-8 with the expectation that students will reach proficiency by eighth grade. Growth increments required to reach proficiency vary across the years as shown by the graph below (graph taken from Arkansas proposal at http://www.ed.gov/admins/lead/account/growthmodel/ar/argmp.doc).



Student Growth in AYP Calculations: In AYP calculations, a school or district is expected to meet the proficiency target that year based on the status model, the safe harbor model provision, or the target for individual growth. A district, school, or subgroup can meet safe harbor if the proportion of students in that school or subgroup scoring below proficient decreases by at least 10% from the prior year. The Arkansas growth model uses the proficiency targets from its status model to determine AYP. For example, under the Arkansas status model, the proficiency target for grades K-5 mathematics is that 64.08% of the students should be proficient in 2007-2008 in each school and subgroup within a school. Under the proposed growth model, 64.08% of the students in these grades in a school and in each subgroup in the school must make assessment gains in mathematics for the school to be making AYP for 2007-08, or the percentages of such students must meet the safe harbor standard. Required growth is calculated for all students, including those currently below proficient and for those at proficient or above.

Delaware

Growth Model Description: Delaware tests students in reading and mathematics in grades 2-10, writing in grades 3 – 10, and science and social studies in four

grades each (grades 4, 6, 8, and 11). Scores on these assessments are not on vertical scales.

Delaware implements a value tables approach to growth-based accountability using their reading and math scores from grades 2 - 10. Each student in a subgroup earns points for moving across proficiency levels. The number of points increases as students move from levels below proficiency to levels at proficiency. Students who regress contribute zero points.

Delaware has five proficiency levels. In the growth model, the two lowest levels are subdivided into two groups each. For example, if a student scored at Level 1A in Year 1 and Level 1B in Year 2, that result would earn the student's school a certain number of points (e.g., 150 according to the table below). The points in the value tables were set by committee.

An example of a value table used in Delaware (taken from a slide show at link http://www.ccsso.org/content/PDFs/DE_Model_2006_032706.ppt#392,7,Slide 7) is shown below. As shown in the table, the maximum average score for a group of students is 300, which is equivalent to 100% proficient. An average score of 300 would indicate that all students are meeting the standards.

Value Table for Grades 4-10

Year 1 Level Grades 3 -10	Year 2 Level Grades 4 - 10						
	Level 1A	Level 1B	Level 2A	Level 2B	Level 3	Level 4	Level 5
Level 1A	25	150	225	250	300	300	300
Level 1B	25	75	175	225	300	300	300
Level 2A	0	25	125	200	300	300	300
Level 2B	0	0	50	125	300	300	300
Level 3	0	0	25	100	300	300	300
Level 4	0	0	0	50	300	300	300
Level 5	0	0	0	0	300	300	300

Student Growth in AYP Calculations: In Delaware, a school or subgroup meets AYP if that school or subgroup meets three conditions:

1. proficiency targets in reading and mathematics (meets growth targets)
2. meets participation target
3. meets other academic indicator requirements

The growth target for a school or subgroup in any one year is calculated as the proficiency target times 300. For example in 2007 the proficiency target for English language arts was 68%. The growth target was then calculated as 68% of 300 or 204. A school or subgroup needed to have an average growth value of at least 204 to meet growth expectations.

Florida

Growth Model Description: Florida administers its state assessment in grades 3-11. The test contains criterion-referenced tests (measuring selected benchmarks in mathematics, reading, science, and writing) and norm-referenced tests in reading

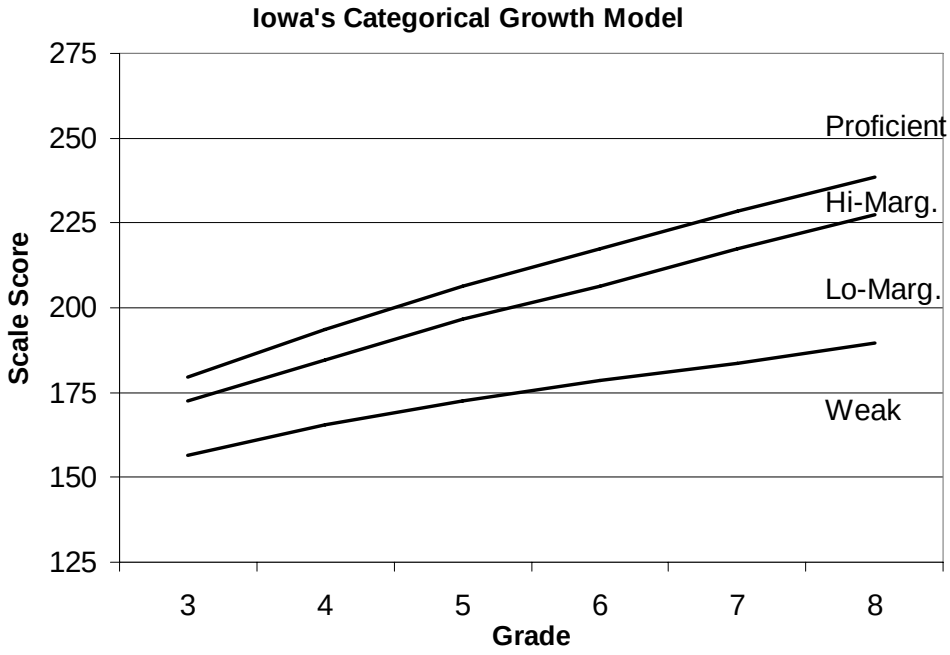
and mathematics (measuring individual student performance against national norms). Florida's state assessment is reported on a vertical scale from grades 3 to 10 with scale values ranging from 0 to 3000. Florida expects students who are not proficient to reach proficiency in three years. Growth targets are set by taking the difference between the proficiency score three years later and the student's initial test score. Students are expected to score at least a third of that difference greater each year. For example, the amount of improvement in terms of decreasing the score difference between the initial score and the proficiency point three years in the future is shown in the table (table information taken from Florida's proposal at link http://www.ed.gov/admins/lead/account/growthmodel/fl/flrevisions2006.doc) below.

Year In State-Tested Grade	Decrease From Baseline Assessment In Performance Discrepancy
1	33% of original gap
2	66% of original gap
3	Student must be proficient

Student Growth in AYP Calculations: Under Florida's growth proposal, each subgroup will have AYP calculated using the status model, safe harbor criteria and a growth model calculation. A school or subgroup makes AYP if it meets participation rate requirements; meets the writing requirement and the graduation rate requirement (as applicable); and meets the proficiency target that year based on the status model, satisfies the safe harbor model provision, or meets the target for individual growth.

Iowa

Growth Model Description: Iowa districts test all grade 3-8 students using the Iowa Tests of Basic Skills (ITBS) which has a vertical scale. Districts also test all grade 11 students using the Iowa Tests of Educational Development (ITED). Iowa's growth model was approved for grades 3-8 only. To set growth targets for non-proficient students, Iowa divides the below proficient scale score range for grades 3-8 into three categories and has established category boundaries on the scale score system for non-proficient students across grades. A student's growth trajectory must cross a category boundary in order to be considered to have met growth expectations. In their US ED growth proposal, Iowa decided to label the meeting of growth expectations as Adequate Yearly Growth. Adequate Yearly Growth is defined as the score improvement that non-proficient students are expected to make from one year to the next. The figure below (taken from Iowa's proposal at link http://www.ed.gov/admins/lead/account/growthmodel/ia/iagmp07.doc) shows the category boundaries for non-proficient students across grades. A student's growth trajectory must cross a category boundary in order to be considered for Adequate Yearly Growth.



Student Growth in AYP Calculations: In AYP calculations, Iowa adds the number of students who meet proficiency and those who meet growth targets and divides by all students in the analysis. The proportion is compared with the proficiency target for that year. When safe harbor is applied, students who meet growth targets are included in the analyses. For example, when safe harbor is examined, the number of students meeting AYP is combined with the number of students meeting growth for each year. This combined percentage is subtracted from 100% to determine the percent of non-proficient students, in the prior year and current year. If the percent of non-proficient students is reduced by 10% or more, from the prior year to the current year, the group meets safe harbor.

Michigan

Growth Model Description: Michigan administers the Michigan Educational Assessment Program (MEAP) in grades 3-8, the Michigan Merit Examination to all students in grade 11, and the MI-Access for students with disabilities. Beginning in 2005-2006, the MEAP and MI-Access were expanded to grades 3-8. The Michigan growth model divides each of the four MEAP performance levels (not proficient, partially proficient, proficient, and advanced) into three sub-levels (low, middle, and high). A similar process divided the MI-Access levels into sub-levels, yet fewer sub-levels were used with this alternate assessment. Then, the model sets expectations and tracks student transitions from one year to the next. Based on the numbers of transitions a student must make and the number of years to achieve proficiency, each student is given an improvement target. For example, a student in the low Not Proficient category who needs to make 6 transitions in 3 years to reach proficiency would need to make 2 transitions each year. The improvement target for this student would be 2. See table below (note that this table was taken from Michigan's proposal) for student improvement targets.

The tracking mechanism is called a transition value table. The transition tables compare student performance to increasing expectations and indicate whether students are declining, whether they are exhibiting no change, or whether they are improving their standing. Students' change in performance level is classified into five categories (significant decline, decline, no change, improvement, significant improvement) with accompanying abbreviations (SD, D, N, I, SI, respectively).

Assessment	Previous Performance		Number of Sub-Levels Improvement Needed to Achieve Proficiency	Number of Years to Achieve Proficiency	Improvement Target	
	Level	Sub-Level			Unrounded	Rounded
MEAP	Not Proficient	Low	6	3	2.00	2
		Mid	5	3	1.67	2
		High	4	3	1.33	2
	Partially Proficient	Low	3	3	1.00	1
		Mid	2	3	0.67	1
		High	1	3	0.33	1
MI-Access Functional Independence	Emerging	Low	3	3	1.00	1
		Mid	2	3	0.67	1
		High	1	3	0.33	1
MI-Access Participation & Supported Independence	Emerging	No Sub Divisions	1	3	.33	1

Student Growth in AYP Calculations: Michigan determines whether districts, schools, and subgroups within districts and schools meet the AYP proficiency targets using the following:

$$100 * \left(nProficient + nOnTrajectory + nProvisional \right) / nValidAssessments$$

where *nProficient* is the number of proficient students, *nOnTrajectory* is the number of students on trajectory to proficiency within the next three years, *nProvisional* is the number of provisionally proficient students (i.e., students who meet proficiency when a confidence interval is applied), and *nValidAssessments* is the number of students receiving valid scores on the assessment (non-valid scores cannot count toward participation rates and are not used in proficiency rate calculations).

Missouri

Growth Model Description: The Missouri Assessment Program (MAP) includes assessments in grades 3-8, and in one grade at the high school level annually in communication arts, mathematics, and science. Beginning in 2008-2009, high school assessments will be replaced by end-of-course assessments in English II, Algebra I, and Biology. Students whose significant cognitive disabilities prevent them from participating in MAP subject-area assessments are assessed with the Missouri Assessment Program-Alternate (MAP-A).

The first year a student tests in a Missouri public school is the baseline year. Student growth targets are established using that baseline score to determine the scores that the student should achieve in each subsequent year of testing to be proficient at the end of four years, or by the end of grade 8, whichever occurs first. Growth targets are calculated as the numeric difference between the student's scale score in the baseline year (grade 3 for the majority of students) and the scale score defining proficiency at the end of the target grade level.

Growth targets represent the amount of improvement (in terms of scale score changes) the student should show each year in order to reach proficiency by the target grade level (the earlier of grade 8, or four years from the baseline score). The table below (taken from Missouri's proposal at link http://www.ed.gov/admins/lead/account/growthmodel/mo/mogmp.doc) illustrates growth targets for baseline scores determined in each grade 3-7.

Baseline Grade (Status)	Year 1 Benchmark	Year 2 Benchmark	Year 3 Benchmark	Target Grade (Status)
3	Grade ¼ distance from baseline to grade 7	Grade 5 ½ distance from baseline to grade 7	Grade 6 ¾ distance from baseline to grade 7	**7**
4	Grade 5 ¼ distance from baseline to grade 8	Grade 6 ½ distance from baseline to grade 8	Grade 7 ¾ distance from baseline to grade 8	**8**
5	Grade 6 1/3 distance from baseline to grade 8	Grade 7 2/3 distance from baseline to grade 8	N/A	**8**
6	Grade 7 ½ distance from baseline to grade 8	N/A	N/A	**8**
7	Status for grade 7	N/A	N/A	**8**

Student Growth in AYP Calculations: Missouri incorporates a growth model calculation into its accountability system at grades 3-8, establishing unique growth trajectories that will ensure that, by 2014, all students will either be proficient or "on-track to be proficient" by the end of grade 8, or within four years of the baseline score, whichever is reached first. For buildings that do not make AYP based on status (as defined in Missouri's current approved Accountability Workbook), assessment data will be analyzed at the student level to determine which students are "on track to be proficient." For each student, a growth trajectory will be calculated that will ensure that the student is "on- track to be proficient" in each content area within four years, or by the end of grade 8, whichever comes first, depending upon the grade level in which the student's baseline score is determined. The number of students that are "on track to be proficient" will be added to the numerator of the "Percent Proficient" calculation to determine AYP (based on the state's established AMOs identified in the approved Accountability Workbook) for each subgroup, school, district, and the state. All

students that have been enrolled in the district for at least one full academic year (as defined in Missouri's current approved Accountability Workbook, Reference 1, p. 14) will be included in the denominator of this calculation.

North Carolina

Growth Model Description: North Carolina incorporates student test results from grades 3-8 in reading and mathematics. Both end of grade (EOG) and end of course (EOC) assessments are used in the growth model. North Carolina has a vertical scale in the grades 3-8 assessments, but it does not use that vertical scale in its growth model. To set student growth targets, North Carolina transforms students' scores on to a common scale (using standard deviation units). Then, students who are not proficient are expected to lower by a minimum percent (e.g., 25%) each year the difference between the first test and the proficiency standard, typically four years later. North Carolina also gives a third grade pretest and uses that pretest in growth calculations. For example, a student in grade 3 who scores below proficiency on the pretest is expected to score, by the end of grade 3, 25% closer to the proficiency score by the end of grade 6. The North Carolina grades and tests used to define growth and the percent of score difference expected to be closed each year is presented in the table below (taken from North Carolina's proposal at link http://www.ed.gov/admins/lead/account/growthmodel/nc/ncgmp.doc).

Grade Of First Enrollment	Test Used As The Basis For Prediction	Test Used As Target For Proficiency	Years In Trajectory	Percent Of Difference Closed Per Year
3	3rd grade pretest	6th grade EOG	4	25%
4	3rd grade EOG	7th grade EOG	4	25%
5	4th grade EOG	8th grade EOG	4	25%
6	5th grade EOG	Algebra I or English I EOC	4	25%
7	6th grade EOG	Algebra I or English I EOC	4	25%
8	7th grade EOG	Algebra I or English I EOC	3	33%

Student Growth in AYP Calculations: In its AYP calculations, North Carolina adds the number of students who score at or above proficiency to those students (below proficiency) meeting growth targets and divides by the number of students in the analysis. The proportion is compared with the proficiency target (i.e., annual measurable objective) for that year. The state runs the AYP growth calculations only after a school or subgroup misses an AYP target using safe harbor or through the use of the confidence interval.

Ohio

Growth Model Description: Ohio tests lower grade level students in reading and mathematics in grades 3-8, writing grades 4 and 7, science in grades 5 and 8, and social studies in grades 5 and 8. Ohio also tests grade 10 students in reading, mathematics, writing, science, and social studies. Ohio works with Dr. Bill Sanders to calculate student growth. Ohio implements a multivariate longitudinal statistical model which projects student growth based on up to five years of prior assessment data. For example, a student's projected score in mathematics will be computed from that student's prior mathematics, reading, science, and social studies scores. Ohio's model counts as proficient (for AYP purposes) students who are on a path to reach proficiency within two years (with the exception of 7th grade, where students must reach proficiency by 8th grade).

Student Growth in AYP Calculations: Ohio evaluates projected scores separately for reading/language arts and mathematics. It should be noted that projection scores only are used in the determination of AYP for subgroups that were not proficient based on Ohio's other methods of determining proficiency (meeting or exceeding the target, 2-Year Combined Results and Safe Harbor). When determining whether a subgroup has met the annual proficiency target in reading/language arts, Ohio determines if the subgroup has the sufficient percentage of students who are projected to be proficient within two years. The resulting percentage is compared with the proficiency target for the current year, and if the target is met, the subgroup is considered as having met AYP in the current year.

Tennessee

Growth Model Description: Tennessee tests students in grades 3-8 in reading/language arts, mathematics, science, and social studies. In high school grades, Tennessee assesses students in end-of-course assessments, including Algebra I, Mathematics Foundations II, English I, English II, and Biology I, Physics, and US History. Students must pass Algebra I, English II, and Biology I in order to earn a high school diploma. Tennessee works with Dr. Bill Sanders and implements a statistically complex projection model such that students projections,, based on average growth in future years, for the targeted grade are compared to the proficiency standard (three years in the future). The growth model applies to students in grades 4-8. For example, a fourth grade student must be projected to be at or above proficiency in grade 7 to meet the growth target in Tennessee's AYP calculations. For example, a third grade student must be projected to be at or above proficiency in grade 6 to meet the growth target in Tennessee's AYP calculations. Tennessee's criterion referenced assessments in grades 3-8 are on a vertical scale, though the vertical scale is not used in the projection calculations. To compare scores for purposes of growth, Tennessee transforms all assessment scores to a normal curve equivalent scale.

Student Growth in AYP Calculations: Schools and subgroups in Tennessee have three options for meeting elementary and middle school AYP proficiency targets. The first way is to have the percent of students scoring at proficiency or higher at least as great as the proficiency target for Tennessee that year. The second option is a safe harbor option. The third option is to have the percent of students with scores in both reading and mathematics projected three years later as proficient or

higher at least as high as the proficiency target for that year. One feature of the second option is that a student who scores at least proficient in the current year but has a projected score three years later that is below proficient will not add to the numerator of the percent of students compared with the proficiency target for that year.

Impact on Schools and Districts

Table 5 summarizes the impact of incorporating growth into states' AYP calculations. The number and percent of districts and schools who met AYP because of growth in the noted assessment year are those districts and schools who would not have met AYP had growth not been included. Results suggest that the impact of adding the growth model to state's AYP calculations varies. For some states, such as Alaska and North Carolina, the impact has been small. For other states, the impact has been much greater. The Ohio projection model, for example, has made a large impact on the numbers of districts and schools meeting AYP. The reasons for the variability in impact are many and complex, including the conditions under which the model was reviewed, the way the model is calculated, the number of years for which students are able to meet growth targets yet not meet proficiency expectations, and the ways in which growth is incorporated into the AYP calculations. The differential impact relates as much to the type of model implemented as to the way in which growth information is integrated into the AYP calculations.

Table 5. Summary of Growth Impact on AYP for Approved States

State	Assessment Year	Students*	Districts		Schools		Comments
			Total*	Number (%) Met AYP Because of Growth	Total*	Number (%) Met AYP Because of Growth	
Alaska	2006-2007	133,288	54	0 (0%)	502	0 (0%)	
Arizona	2006-2007	1,094,494	218	0 (0%)	2,078	8 (0.7%)	
Arkansas	2006-2007	474,206	255	9 (3.5%) in 2007-2008	1,138	11 (1%)	
Delaware	2006-2007	120,937	19	N/A	222	7 (3.0%)	
Florida	2006-2007	2,675,024	67	4.7%	3,244**	151	
Iowa	2006-2007	482,584^	365^	77 (21.1%)	1491^	128 (8.6%)	
Iowa	2007-2008	480,609^	364^	9 (2.4%)	1477^	65 (4.4%)	
Michigan	2007-2008	1,741,845	552	32 (5.7%)	3,640^	111 (3.0%)	751 (20.6%) schools did not meet AYP
Missouri	2007-2008	917,705	524	14 (2.6%)	2,361	149 (6.3%)	
North Carolina	2005-2006	1,416,436	115	0 (0%)	2,353	0 (0%)	
North Carolina	2006-2007	1,417,426	115	1(0.9%)	2,350	12 (0.5%)	1047 schools met; 1298 schools not met
Ohio	2007-2008	1,839,683	615	249 (40.6%)	4,007	1028 (25.7%)	
Tennessee	2005-2006	933,688+	136+	N/A	1,373	7 (0.4%)	^^Growth is final of three options for AYP
Tennessee	2006-2007	978,368+	136+	N/A	1,373	19 (1.1%)	^^Growth is final of three options for AYP

Note. The numbers in this table depend on the sources from which the data were obtained. If data are taken from other sources, results may be slightly different. *Data from SchoolDataDirect.org website. **Represents total number of schools with a "Yes" or "No" AYP designation in Florida. ^Number determined from state AYP calculations or from the state's website. ^^ In Tennessee, AYP at the campus level is determined by status, safe harbor, and growth, where growth as defined by meeting projection of proficiency or above at three years in reading/language arts and math separately. Tennessee districts do not currently have the option to use growth for AYP.

Conclusions and Next Steps

The US ED growth pilot study reflects the increased national attention on students' longitudinal data. Instead of limiting the focus to a snapshot view of students' performance, educators and policymakers are evaluating multiple data points and taking into account students' patterns of scores over time to draw inferences about students' development, as facilitated by schools, through the educational system. Whereas previously a student's adequate performance at one point in time was interpreted as evidence that the student was on track to meeting proficiency in subsequent years, the longitudinal view of students' scores over years provides direct evidence of student learning over time.

Several conclusions might be drawn from the review of the US ED growth pilot study from 2005 to 2008. First, though there are three general models that have been approved (the growth to proficiency models, the value and transition tables, and the regression-based models), each state model is unique and has features that make it different from the other models. The set of growth to proficiency models includes models with many varying features. Some of these models, for example, set growth targets at a maximum of three years, whereas others set targets using four or even five years as a maximum. Even the two approved projection models, which are both implemented through a consulting agreement with SAS, have considerable differences. Ohio evaluates projections for most students two years in the future, whereas Tennessee evaluates projections for most students three years in the future. In addition, Ohio and Tennessee incorporate projections into their AYP calculations in different ways.

A second conclusion is that US ED has not provided consistent guidance around the rules for implementing a growth model in AYP calculations. This is likely due to changes in peer reviewers for each round of review. The panels have had different emphases and this has resulted in inconsistencies. For almost every core principle and guiding principle, there are approved states that have had their compliance with the guidance and principles interpreted differently. It seems as if the limitations and rules around the growth models emerged throughout the pilot study. Though the lack of clear guidance has been frustrating to states and may have led some states to wait to submit growth models until the "rules" were clearer, the newness of using growth models in the federal accountability system made it necessary for US ED to use the pilot study to refine the rules over time.

A third conclusion is that though growth models have provided some flexibility for states in the federal accountability system, the growth models by and large have not resulted in substantially more schools and districts meeting AYP. For most states, the addition of growth to their federal accountability systems has resulted in less than 5% of schools or districts meeting AYP due to the growth model. The two states that appear to have experienced the greatest increase from the growth model are Iowa and Ohio. For Iowa, 21.2% of districts met AYP due to the growth model the first year Iowa implemented the growth model in 2006-2007. However, since Iowa implements a growth to proficiency model, many students eligible to meet growth the first year were not eligible to meet growth the second year. Therefore, the impact of growth on Iowa's AYP in 2007-2008 was smaller than in the first year. For Ohio, the impact of the growth model was substantial. The large impact in Ohio is likely due in part to the way in which Ohio implements its

projection model and the way Ohio integrates the growth information into the AYP calculations.

A fourth conclusion is one related specifically to the growth to proficiency models. For most of the states with these types of models approved, the AYP impact will be greatly diminished after several years of implementation. Many of these states set a limit to the number of years a student can meet growth targets without meeting proficiency. For example, Florida allows students no more than three years to reach proficiency. After three years, students who are not proficient will not be able to count for AYP through the growth model. After the three years is up, the only students for whom growth can cause students to count positively for AYP are students new to the state, such as entering third graders or students who move to the state. Therefore, states implementing growth to proficiency models will experience a decrease in the benefit from the growth model. It will be interesting to see if any states with growth to proficiency models propose to switch to a model that does not limit the years for which growth can count for AYP. Of course, changing the time to proficiency has implications beyond simply providing more time for students to become proficient, such as shifting the accountability burden from elementary to middle or high schools.

Regarding next steps, several issues remain unresolved. One is the extent to which states can implement index systems and growth models. After US ED gathers input from the National Technical Advisory Committee, the parameters under which states may be able to implement both types of flexibility will likely be defined more clearly. Another unresolved issue relates to the caps on the percents of students taking the modified and alternate assessments in states that can count for AYP. It is unclear whether a state with 1.5% of students taking the alternate assessment can have 1% count for AYP due to meeting proficiency and the additional 0.5% of students count for AYP due to meeting growth expectations.

Another next step in growth modeling relates to the ways in which states will communicate growth model results to LEAs, schools, and parents. Many states are finding it challenging to explain how growth is defined and the many details related to growth models. The next several years should see states focus resources on reporting tools that include longitudinal student performance. In addition, over the next few years, states with approved growth models will be generating many documents and websites with supporting documentation on growth models. The types of information and reporting mechanisms that communicate most successfully will need to be identified and shared across states.

An additional next step will be to identify fruitful ways to integrate growth model results down to the classroom level. Ways in which growth results could be used to identify promising teacher practices and ways in which growth results can be shared effectively with teachers will need to studied.

In sum, the nation is at the beginning stages of implementing growth models. Researchers have begun to work out some of the technical psychometric issues, but state data and context still must be critically weighed in designing a specific growth model. Technical issues aside, the practical benefits of growth models are only now being realized and much work is still needed to maximize growth results. The US ED and states are learning about the effectiveness of growth models, their impact

on AYP, and communication strategies that work. As such, US ED and states have more questions than answers at this time.

Access to Other ASR Papers and Products

http://www.ccsso.org/projects/scass/projects/accountability_systems_and_reporting_consortium/8705.cfm

- Statewide Educational Accountability Systems Under the NCLB Act - A Report on 2008 Amendments to State Plans (Annual paper 2003-2008)
- Implementer's Guide to Growth Models, 2008
- Key Elements for Educational Accountability Models
- Quality Assurance Practices Associated With Producing Cohort Graduation Rates 2007
- Validity Threats: Detection and Control Practices for State and Local Education Officials, 2006
- Policymakers' Guide to Growth Models for School Accountability: How Do Accountability Models Differ? 2005
- A Framework for Examining Validity in State Accountability Systems,
- Revisiting Statewide Educational Accountability Under NCLB: A Summary of State Requests in 2003-04 for Amendments to State Accountability Plans, 2004
- Statewide Educational Accountability Under NCLB: Central Issues Arising from an Examination of State Accountability Workbooks and U.S. Department of Education Reviews Under the No Child Left Behind Act of 2001, 2003
- Making Valid and Reliable Decisions in Determining Adequate Yearly Progress (Joint CAS and ASR Publication), 2003, and Executive Summary
- Guide to Effective Accountability Reporting, 2002
- Designing School Accountability Systems: Towards a Framework and Process, 2002
- Accountability State Profiles http://accountability.ccsso.org/index.asp