

# **THE COLORADO GROWTH MODEL: HIGHER EXPECTATIONS FOR ALL STUDENTS**

**Submitted by Dwight D. Jones, Commissioner of Education**

**Colorado Department of Education**

**October 15, 2008**

## **EXECUTIVE SUMMARY**

### **Introduction**

Commissioner Dwight D. Jones on behalf of the State of Colorado is pleased to submit this proposal to the United States Department of Education to allow the Colorado Department of Education to incorporate measures of student longitudinal growth into Colorado's Adequately Yearly Progress (AYP) determinations. Colorado's growth model meets all seven principles outlined by Secretary Spellings and spelled out in the peer review guidance.

Briefly, Colorado developed the Growth Model to answer three essential questions about student, school and district performance:

- What is? What is the growth rate of a student, a school and a district?
- What should be? What should the growth rate be for a student to reach a desired level of achievement within a period of time?
- What could be? What are the highest sustained growth rates to date and under what conditions could growth rates improve?

To answer these questions, the Colorado Growth Model uses a common measure to describe how much growth each student has made and how much growth is needed to reach state standards. The Colorado Growth Model provides data that are understood by stakeholders as fair and transparent to support school, district, state and federal accountability purposes. It does this by applying the common measure of Student Growth Percentiles to school, district and state performance in a normative and standards-based manner.

Colorado is committed to focusing educational reform and school improvement efforts around the Colorado Growth Model and incorporating results from the growth model into Colorado's District Accreditation System and School Accountability Reports (SAR). Incorporating the results of Colorado's Growth Model into AYP determinations will allow the Colorado Department of Education (CDE) to achieve a coherent state system of accountability and support that can reinforce the goals of both the state and federal systems.

This executive summary describes the Colorado context, the policy rationale, and an overview of the proposed model. The details of the model are described in the main body of the proposal, which is organized according to Secretary Spellings' seven guiding principles.

## **The Colorado Context and Background**

### **Policy Support and Rationale**

Colorado's educational accountability system is undergoing a transformation focused on alignment around individual student progress and relevance for educational improvement. This work builds on a bipartisan history of valuing the measurement of individual student progress toward state established academic standards, culminating in the Colorado Growth Model. Importantly, Colorado approached the measurement of student longitudinal growth thoughtfully and deliberately even before No Child Left Behind was enacted in 2001. Starting with instituting a rigorous system of unique student identifiers, CDE supported a program of research and development to explore several different approaches to measuring growth before adopting the Colorado Growth Model.

The proposed use of the Colorado Growth Model for AYP determinations aligns well with Colorado's overall education policy direction. The Colorado Department of Education (CDE) has actively pursued the analysis of student longitudinal data, including aggregations to the school and district levels, from the Colorado Student Assessment Program (CSAP) data for at least the past decade. Legislation enacted in 2004 (HB 2004-1433) led to the establishment of a technical advisory panel (appointed by then-Governor Bill Owens) and required that CDE established growth analysis techniques for diagnostic purposes. Building on this initiative, legislation enacted in 2007 (HB 2007-1048) directed CDE to refine the methodology and produce more useful information for schools and parents, while expanding its use for accountability purposes. A technical advisory panel was appointed by current Governor Bill Ritter and was tasked with recommending a model to the State Board of Education.

The intent of HB 2008-1048 was to make longitudinal growth of students the cornerstone of the state's accountability system. Colorado's accountability system includes the state's accreditation of school districts, the School Accountability Report, and the determination of Adequate Yearly Progress. In addition to unifying the accountability system, this legislation requires that the longitudinal growth data used for accountability also provide information to students, parents, teachers and administrators that support improved academic achievement.

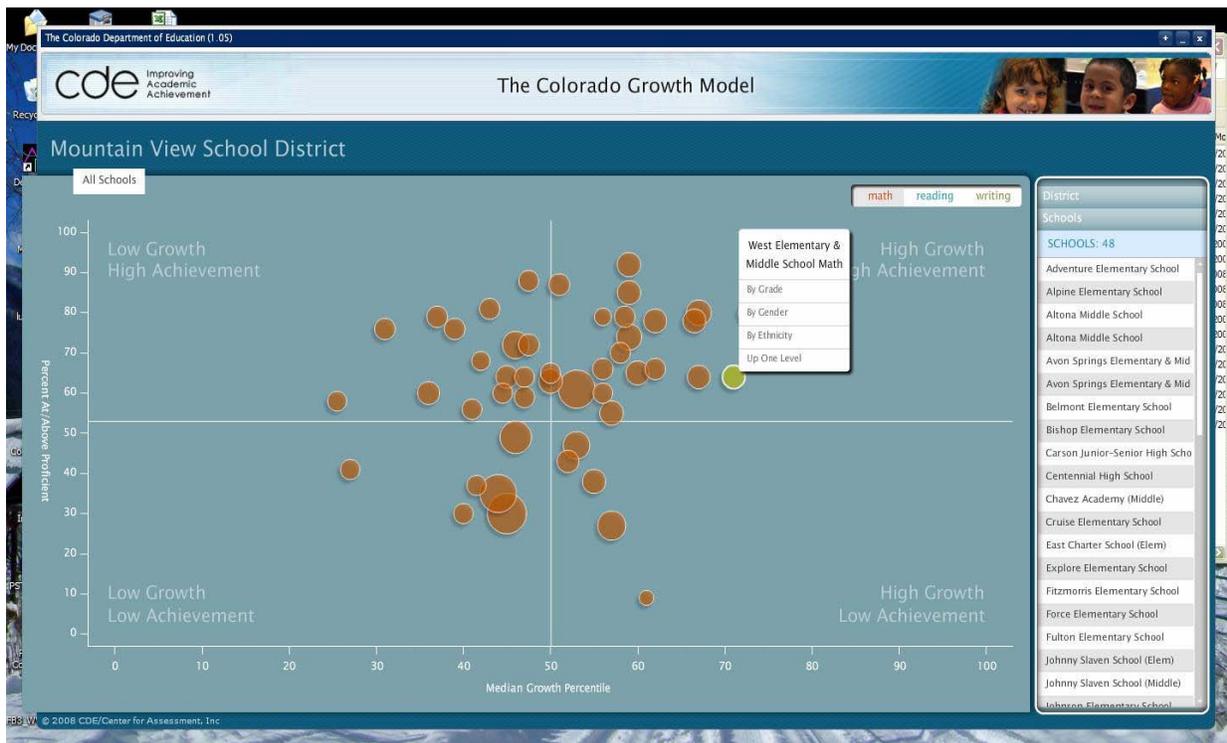
The State Board of Education required, based upon the advisory panel recommendations, that the Colorado Growth Model calculate a growth percentile for each student relative to all other students in the state with the same prior academic history (academic peers). Subsequently, Colorado's District Accreditation System and School Accountability Reports were revised to incorporate longitudinal growth for students, as measured by Colorado's Growth Model.

There is widespread support in Colorado to include measures of student longitudinal growth in AYP determinations. A joint resolution was passed February 1, 2006 by the Colorado House and Senate urging the Colorado Department of Education to apply for this growth pilot. The Denver Area School Superintendent's Council has also called for amending AYP to include growth measurements. A joint white paper from the Colorado Association of School Executives, the Colorado Association of School Boards, the Colorado Education Association, and Colorado

BOCES Association, also called for the inclusion of an individual student growth measure in AYP.

### Connection to other state initiatives

The Colorado Growth Model will allow educators from the state, district and school levels to focus service and support on raising student achievement for every student and closing achievement gaps. It offers a means for schools and districts to learn from one another, and allows CDE to target its limited resources toward the greatest need. The Colorado Growth Model's data visualization tools allow an unprecedented level of public disclosure of and interaction with information about school and district performance. The screenshot below shows the common framework Colorado uses to display school performance (for a view of the full reporting tool, see: <http://www.cde.state.co.us/cdeassess/growthmodel.html>). This approach has been embraced by stakeholders as an understandable and fair presentation of data. The software provided to districts allows educators to drill-down through districts, schools, grades, to individual students, showing their entire CSAP performance history.



From the perspective of state policy and practices, the Colorado Growth Model supports a common understanding of how individual students and groups of students progress from year to year toward proficiency on state standards based on where each student begins. The model reveals where, and among which students, the greatest growth is happening. It also identifies areas of least growth. It recognizes that effective schools produce higher sustained rates of student growth. Those schools may or may not be schools with the highest test scores every year. Given this framework, the Colorado Growth Model can serve as an effective tool for program evaluation and is therefore; central to several state initiatives.

Further, in response to the NCLB requirement that state education agencies provide technical assistance to districts identified for improvement, CDE developed a district evaluation process called Comprehensive Appraisal for District Improvement (CADI) and a parallel process for schools called the School Support Team (SST) review in 2003. Both CADI and SST are federally-funded school improvement resources that support districts and schools in the development of effective improvement plans. The state uses CADI and SST findings in determining the effectiveness of improvement plans and in supporting implementation with allocation of Title I improvement grants. In 2008, CDE introduced a new, state-funded pilot program called “Closing the Achievement Gap” (CTAG). This program incorporates the CADI process and extends it to the provision of additional resources focused on closing achievement gaps between disaggregated groups of students based on race and income.

These three state support systems share common features. First, they all involve an appraisal visit to the district or school from an external group of experts (selected and trained by CDE) to review relevant documents and materials; interview staff, leadership, parents and community members; observe various operations; and develop a final report. The final report presents the team’s findings about the district or school’s current level of functioning against nine research-proven performance standards and cites commendations in the areas where the district/school is doing well. The report also sets out recommendations that are tied to themes; these themes are, in turn, tied to strategic actions that the district/school can implement to begin the improvement process (Colorado Department of Education, Consolidated Federal Programs, 2008).

Second, the three support systems are anchored in the same set of evidence-based standards of effective practice that form the basis of the appraisal process and inform strategic planning for improvement. Each of the nine standards has a set of indicators and accompanying rubrics to guide determinations about the degree to which the standard is currently evidenced in a specific district/school. The standards are divided into three strands: academic performance, learning environment and organizational effectiveness. CDE systematically revises the rubrics to improve them and align them with the needs of the various stakeholders in the process. Use of the Colorado Growth Model will enhance the work of providing support to schools and districts. The model will assist the department in honing its prioritization process, the accuracy of the reviews and the potency of the recommendations that follow.

The Colorado Growth Model creates an incentive for educators to focus on maximizing the growth of all students toward reaching state standards and provides tools to educators that allow them to quickly know which students need to make the most growth and how the growth they have been seeing measures up to the best progress in the state. The benefit of the Colorado Growth Model is in the logic underlying it. The growth model provides a tool to identify upwardly-trending schools and more importantly the possibility to learn from them. Principals, teachers, and the public can see trends, find others who are improving, learn from them, improve practice, and ultimately post improved results. The greatest value of the model may be that it is not just another way to sort but it is a way for us to learn from each other. We provide a few brief descriptions below of key state initiatives for which the Colorado Growth Model will provide useful information and support.

### Developing a coherent P-20 accountability system

The Colorado Growth Model is a core element of Senate Bill 2008-212, Colorado's P-20 alignment legislation. This legislation requires the revision of P-12 standards and assessments and the addition of early school readiness and postsecondary and workforce readiness (PWR) assessments. The Colorado Growth Model will be applied to new and revised assessments and used to identify and support the academic growth of Colorado students.

### Conducting program evaluation

The Department is implementing an ambitious research agenda focused on evaluating key aspects of its system of support for districts and schools and conducting a return on investment analysis of key programmatic elements. The Colorado Growth Model provides a common measure of student growth rates that is ideal for understanding program efficacy.

### Comprehensive Appraisal for District Improvement (CADI)

CADI is a federally-funded, state-supported, research-based appraisal that provides a foundation from which districts can engage in systematic and strategic planning for implementation of change. Since its statewide launch in 2005, 23 of Colorado's 178 local school districts have completed the CADI review process. A CADI review is voluntary and must be requested by district leadership. Eligible districts fall into one or more of the follow categories, but they must be a district in Title I Program Improvement status: (a) district was recommended because of district accreditation status, (b) district has an increased percentage of schools with declining achievement, or (c) district has gaps in achievement between disaggregated groups of (i.e., minority students and students living in poverty). Districts provide data for the CADI review by compiling a portfolio of specified district information and providing access to district personnel, selected parents and community members. The results from the Colorado Growth Model will become a critical component of the data review.

### School Support Team (SST) Review

This federally-funded statewide program was launched in 2003 as a way of delivering focused technical assistance to Title I schools. SST review uses a process like that of CADI, but moves the level of focus to the school rather than the district. To date, 93 schools (in 25 districts) have completed an SST review. The SST review is also voluntary and is requested by school leadership. Schools are eligible to apply for a school improvement grant and a SST review if they are identified for School Improvement (i.e., Title I schools that have not made AYP for two consecutive years in the same content).

### Closing the Achievement Gap (CTAG) Initiative

The Colorado Growth Model allows us to measure the existence of growth gaps along with "status" gaps to determine which groups of students are actually growing faster. Emerging research (e.g., Ho, 2008) is demonstrating the efficacy of using student longitudinal growth measures for evaluating achievement gaps compared with status approaches. Growth gap measures have already been incorporated into the state's district and school accreditation process. To close the achievement gaps that plague our education system, we must eliminate gaps in how children are growing academically and ensure that our neediest students grow faster — enough so that they catch up and keep up.

To this end, Colorado has a major initiative underway to close achievement gaps. Closing the Achievement Gap (CTAG) is a state-funded program being piloted in six Colorado school districts in the current (2008–2009) school year. Districts are eligible for participation by meeting two criteria: achievement gaps in reading and/or math larger than the state average for two consecutive years, and a demonstrated willingness and capacity to participate.

Achievement gaps are defined for this program as gaps in student achievement between white and minority students and between students who live in poverty and those who do not. Participating districts in the CTAG program collaborate with CDE to undergo a CADI review, as previously described.

## **An Overview of the Colorado Growth Model**

The Colorado Growth Model uses quantile regression methodology combined with all available prior test score data for each student to determine students' growth-to-standard targets. We explain the methodology in considerable detail in the main body of the proposal and in the technical appendices. Among its many uses, the Colorado Growth Model determines whether each student scoring in the unsatisfactory level is growing sufficiently to score partially proficient (NCLB-proficient in Colorado) within three years and whether each student scoring at the NCLB proficient level or higher is growing sufficiently to at least maintain their current level for the next three years. These growth-to-standard determinations form the basis of Colorado's application for the incorporation of growth into AYP determinations.

Colorado's assessment program begins with testing in grade 3. Therefore, growth-to-standard determinations will be performed for all full academic year (FAY) students in grades 4-10 with at least two valid CSAP scores. All previous test scores will be used in the calculations for students with more than one prior test score. Students participating in Colorado's alternate assessment (CSAPA) are not included in this growth metric, but they are included in performance, safe harbor and Colorado's already approved matched safe harbor.

The AYP growth calculation adds the number of unsatisfactory students on track to be NCLB-proficient to the number of proficient students on track to stay proficient. This numerator is divided by the total number of full academic year students in grades 4-10 with at least two valid CSAP scores in the Districts/schools/disaggregated groups. The resulting percentage is then compared to the growth AMO (described below) to determine whether or not the Districts/schools/disaggregated groups made AYP.

As described above, the Colorado Growth Model sets growth targets for all students, whether they are proficient or not. Further, the Colorado Growth Model holds all disaggregated groups accountable for reaching the same growth targets. As the central piece of Colorado's school and district accountability system, CDE is committed to reporting growth-to-standard results as well as the normative information generated from the growth percentile calculations. Colorado believes that characterizing student growth in learning academic content is a fairly high level inference and reporting growth using both standards-based and normative referents can greatly aid stakeholders' level of understanding and enhance subsequent actions.

## **Incorporating growth model results into AYP determinations**

Colorado currently evaluates schools based on status, safe harbor, and matched safe harbor, along with participation and the other indicator. We propose to include the Colorado Growth Model in our AYP calculations as an “equal partner” to the status determinations. All Districts/schools/disaggregated groups will be evaluated for both growth and status so that AYP calculations are coherent with other Colorado accountability systems where schools’ growth and status are reported for every school. Districts/schools/disaggregated groups meeting either the status or growth annual measurable objectives (AMO) will make AYP. Those Districts/schools/disaggregated groups that do not meet either the growth or status AMO will be eligible to make AYP through safe harbor or matched safe harbor.

The calculations for determining the percentage of students in a disaggregated group meeting growth targets are described in detail in the proposal and appendices. Briefly, each student is evaluated each year to determine whether they are on track to catch up (for unsatisfactory students) or keep up (for all other students). A percentage is calculated by dividing all students on track to catch up and keep up divided by the number of continuously enrolled students in the disaggregated student group, school, or district. These quantities must be compared to a target—or an AMO in NCLB terms—to determine whether the districts/schools/disaggregated student groups made enough growth to make AYP. NCLB provides only one method for establishing AMOs and intermediate goals. This is typically referred to as the 20<sup>th</sup> percentile approach where schools are rank-ordered based on the percentage of students scoring proficient, or meeting growth targets in this case, and identifying the school that cuts off the bottom 20 percent of the students. The percent of students scoring proficient (or meeting growth targets) in this school becomes the starting point for setting AMOs. This approach worked well in the early days of NCLB with 12 years until 100% of students were required to be proficient. Colorado believes that starting at the “20<sup>th</sup> percentile” at this point in the life of NCLB does not make sense because it would set low expectations in the first year, but would require unrealistic increases for schools in the coming years. Since we are just past the halfway point between 2002 and 2014, Colorado proposes setting the initial growth AMO at the “60<sup>th</sup> percentile” school using an adaptation of the methodology described for the 20<sup>th</sup> percentile approach. Applying this methodology results in the following initial targets for percentages of students meeting growth targets for the Districts/schools/disaggregated groups to make AYP. Further, in order to avoid too many changes to the AYP system at once and to aid communication with Colorado’s school and district leaders, Colorado proposes including one intermediate goal in 2011 to parallel the 2011 intermediate goal for the status calculations.

Table 1. Annual growth targets and intermediate goals.

<b>Year</b>	<b>Grade Span</b>	<b>Math</b>	<b>Reading</b>
2009	Elementary	67.69%	82.06%
	Middle	57.64%	81.90%
	High	58.62%	86.06%
2011	Elementary	83.84%	91.03%
	Middle	78.82%	91.45%
	High	79.31%	93.03%
<b>2014</b>	Elementary	<b>100.00%</b>	<b>100.00%</b>
	Middle	<b>100.00%</b>	<b>100.00%</b>
	High	<b>100.00%</b>	<b>100.00%</b>

### **The relationship between Colorado’s accountability system and AYP**

Colorado’s accountability system has been characterized by disparate elements that are now being brought into alignment through use of the Colorado Growth Model. Colorado’s accountability system includes a District and School Accreditation process, School Accountability Report Ratings, and AYP measures under NCLB. CDE revised the District Accreditation process for the 2008-09 school year to focus on individual student progress toward reaching state standards, as measured by Colorado’s Growth Model, the reduction of achievement gaps, and student readiness for postsecondary success. Also for the 2008-09 school year CDE has incorporated the Colorado Growth Model into its School Accountability Report ratings as the measure of student growth. The next step in our alignment strategy is to address Federal AYP requirements using the Colorado Growth Model. Once approved for this use CDE will be in a position to request that the Colorado General Assembly make necessary legislative changes to fully bring the three systems into alignment and deliver on the statutory mandate that the Colorado Growth Model be the cornerstone of the state’s educational accountability system.

The use of the Accreditation process provides our broadest perspective about district and school performance, given its use of annual student-level achievement and growth measures, gap measures, as well as measures of postsecondary readiness. Annual Accreditation reviews provide a signaling mechanism to direct state resources and attention to the districts with the greatest needs.

The Colorado Growth Model provides a Growth Percentile ranging from 1 to 99 for every student and provides the growth-to-standard percentile needed for a student to reach Partially Proficient, Proficient and Advanced levels within one, two, or three years. The growth-to-standard criteria supplied by the Colorado Growth Model provide information on the adequacy of growth to reach state-defined performance levels — we refer to these as Catch Up and Keep Up. On Track to Catch Up identifies low-achieving students in the prior year who demonstrated sufficient growth to reach desired performance levels within three years or by 10<sup>th</sup> grade. On Track to Keep Up identifies students already scoring NCLB-proficient or higher who demonstrated sufficient growth to stay at their current levels over three years. It is this articulation of criterion-referenced growth or “adequate growth” that is the focus of this AYP proposal. Once approved, Colorado will be positioned to have a unified articulation of student growth to standard for accountability purposes.

In addition to criterion referenced growth-to-standard results, the Colorado Growth Model also provides Median Growth Percentiles that are useful for benchmarking purposes and analysis of gaps in growth rates among groups of students. The overall State Median Growth Percentile for every grade is 50, so it is useful to look for differences from the 50th percentile when benchmarking the growth of the typical student. CDE is able to review median growth percentiles for every school and student group to understand how student progress is distributed across schools and districts.

The Colorado Department of Education and all of the key education stakeholder groups in Colorado support this move toward a unified view of school and student performance. The combination of status, growth-to-standard, and normative growth measures provides a comprehensive picture of student achievement allowing Colorado educators to better design strategies for ensuring that all Colorado students are prepared to participate in the 21<sup>st</sup> century economy.

It is this multi-dimensional perspective on student achievement and school quality that Colorado has focused on making widely available to all stakeholders through user-friendly and highly interactive data display tools and resources (see [www.cde.state.co.us/growthmodel.asp](http://www.cde.state.co.us/growthmodel.asp)). The goal is strengthened public and professional accountability. Through improved public disclosure and understanding of school performance, parents will become more knowledgeable choosers of their schools, taxpayers will become more knowledgeable and stronger advocates for education reform, and educators will have the tools to target their improvement efforts and know whether they are working.

**THE COLORADO GROWTH MODEL:  
HIGHER EXPECTATIONS FOR ALL STUDENTS**

**Submitted by Dwight D. Jones, Commissioner of Education**

**Colorado Department of Education**

**October 15, 2008**

**COLORADO'S RESPONSE TO THE SEVEN CORE PRINCIPLES**

**Core Principle 1: 100% Proficiency by 2014 and Incorporating Decisions  
about Student Growth into School Accountability**

**1.1 How does the State accountability model hold schools accountable for universal  
proficiency by 2013-14?**

Grades and content areas

Schools and districts will be held accountable for student growth in mathematics and reading for students with at least two consecutive valid Colorado Student Assessment Program (CSAP) scores who have been continuously enrolled for a Full Academic Year (FAY) in that school. In the case of transition grades (e.g., moving from elementary to middle schools), students will need to have been continuously enrolled in the district for at least one year and enrolled in the school before October 1<sup>st</sup> of the current school year. The CSAP program tests students in reading and math every year from grade 3 through grade 10. This gives Colorado the advantage of being able to carry this growth model seamlessly through the middle-high school transition. Therefore, schools and districts in Colorado will be accountable for student growth in all grades 4-10. While not part of this proposal, Colorado is moving towards implementing a series of college readiness assessments to become operational in 2010-2011. The Colorado Department of Education is committed to including these college readiness assessments in the growth model so that by 2011 Colorado expects to hold schools accountable for student growth from grade 4 through the end of high school.

100% by 2014

As can be seen in Table 1 below, districts/schools/student groups will be required to have **100% of their students either scoring proficient or be on track to score proficient within three years by the 2013-2014** school year. This requirement applies equally to all schools, districts, and student groups in Colorado.

Further, as we discuss in more detail in subsequent sections of the proposal, Colorado is applying the same relatively low minimum-n of 30 students that is currently used in Colorado's AYP determinations for status and safe harbor. CDE is convinced that n=30 ensures the appropriate

balance of inclusion, reliability, and validity, especially considering the diverse demographics of Colorado from major cities on the Front Range to small rural schools on the eastern plains.

## **1.2 Has the State proposed technically and educationally sound criteria for “growth targets” for schools and subgroups?**

### Calculating growth targets

As described in more detail in Section 2.1 below, Colorado uses all available test score data to describe students’ growth trajectories over time. Student growth targets for the assessment year (e.g., 2008) are established based on these analyses such that each student’s growth target is based upon that student’s unique testing history. By using all available data, Colorado is establishing the most valid possible set of targets for each student. After considerable deliberation, Colorado has established the following criteria for determining whether or not students have met their growth targets.

*A student will meet his/her growth target if:*

- *The student is unsatisfactory and on track to be NCLB-proficient within 3 years or by 10<sup>th</sup> grade, whichever comes first,*
- *The student is NCLB-proficient and is on track to maintain NCLB-proficient for the upcoming 3 years or by 10<sup>th</sup> grade, whichever is first,*
- *The student is Colorado-proficient and is on track to maintain Colorado-proficient for the upcoming 3 years or by 10<sup>th</sup> grade, whichever is first, or*
- *The student is advanced and is on track to maintain Colorado-proficient for the upcoming 3 years or by 10<sup>th</sup> grade, whichever is first.*

It is crucial to consider Colorado’s achievement standards when evaluating these growth targets. Colorado has purposefully designed a system of increasing expectations, particularly in mathematics, for students as they progress through the grade levels. It becomes increasingly more challenging for students to maintain proficient and advanced performance as they progress through the grades. Therefore, Colorado claims that these are appropriately rigorous growth targets for students all along the achievement continuum.

Students will be judged “on track” or not depending on whether they have met a specific growth target each year. This target will be the score that, depending on the student’s assessment history, is on or above the trajectory from the student’s previous scores to reaching/maintaining proficient in three years. Importantly, this is not a projection model. Instead, all students are evaluated each year to determine whether or not they made enough growth over the current school year to be on pace to reach or maintain proficiency in three years or less. Further, *achievement* targets are not reset each year, but *growth rates* necessary for students to reach these fixed, future achievement targets are updated annually based upon the latest student and state performance data targets are fixed until the student reaches/maintains proficiency.

For example, in 2007, the state of Colorado established the growth rates necessary for all non-proficient students to reach proficiency. If that student exceeds the first year growth necessary to reach proficiency in 3 years (but still didn’t reach proficiency), then to reach proficiency by 2010, the rate of growth necessary over the next two years can be updated to reflect the two year

goal confronting the student. The purpose is to present the most current data on progress and goals to the student without altering the achievement goals that have been established.

We recognize that the Colorado Growth Model, employing growth percentile methodology, will be a relatively new approach for evaluating student longitudinal growth. The technical appendix (Attachment 2) presents a very comprehensive and technically-detailed description of these methods. Further, the growth percentile methods have received widespread accolades at national conferences because of the model's capacity for avoiding many of the limitations of other growth modeling techniques. To gain a better understanding of how the Colorado Growth Model is implemented, we have created the Colorado Growth Model Tutorial (see Attachment 1) and urge all readers to review this ten minute tutorial before proceeding with this proposal.

#### Growth targets and 100% proficient

As documented in Table 1, Colorado's growth targets require 100% of students to be at or on track to be proficient. We fully recognize that schools/student groups that have not achieved universal proficiency by 2014 could still make AYP if they have 100% of their students on track to reach proficiency within three years. Colorado strongly believes that while this would require the Secretary's flexibility, 100% of student on track or proficient by 2014 is still an ambitious goal and fully aligned with the intentions of NCLB.

### **1.3 Has the State proposed a technically and educationally sound method of making annual judgments about school performance using growth?**

#### Introduction

The current AYP status and safe harbor calculations capture two ways to measure the "quality" of a school. Safe harbor looks at how the school improves with different cohorts of students from one year to the next. It answers the question, "Does the school have fewer non-proficient third grade students this year than last?" This is an important question to answer, as school improvement requires that schools get better at doing what they do every year. However, school effectiveness can best be measured by looking at the growth of individual students. The proposed growth measure would answer "have the students in the school shown improvement during the time they have attended the school?" All three measures add to our understanding of school effectiveness, and both have their limitations. In combination, the three approaches can help us understand if a school is showing improvement for purposes of accountability and school support. Figure 1 below provides a brief overview of this complementary information.

Figure 1. Questions, benefits, and limitations of various measures of school performance.

	Performance	Safe Harbor	Growth
Questions Asked	Are schools/districts meeting a target for proficiency?	Are schools/districts showing decreases in the percent of non-proficient students?	Are students in schools/districts showing progress towards proficiency? Is the school/district able to help all student grow and maintain proficiency?
Benefits	Compares the school/district to a standard level of performance.	Describes if a school/district has fewer low performing students this year compared to last year.	Describes the progress school/district is making with individual students over time.
Limitations	Does not account for where the school/district/student started. Can lead to a focus on “bubble” students <sup>1</sup> .	Does not account for the cohort of students entering or leaving the school. May also lead to a focus on “bubble” students.	While intuitive for many people, the complexities of growth models are not always well understood.

Accountability measures work best when those that are required to do the real work believe both in the goal and in their capacity to achieve the goal. The growth proposal is structured in a way that is difficult for any teacher, parent or citizen to argue with either the intention or the actual targets.

Incorporating growth into AYP determinations

Colorado currently evaluates schools based on status, safe harbor, and matched safe harbor, along with participation, and the additional academic indicators of percent of students scoring advanced (for elementary and middle schools) and graduation rate (for high schools). We propose to include the Colorado Growth Model in our AYP calculations as an “equal partner” to the status determinations in order to support the message that growth is the cornerstone of Colorado’s accountability efforts. All Districts/schools/disaggregated groups will be evaluated for both growth and status so that AYP calculations are coherent with Colorado’s overall educational accountability system where growth and status are reported for every school. Districts/schools/disaggregated groups meeting either the status or growth annual measurable objectives (AMO) will make AYP. Those Districts/schools/disaggregated groups that do not

---

<sup>1</sup> The term “bubble” refers to an unhealthy focus on students who are closest to the cut point.

meet either the growth or status AMO will be eligible to make AYP through safe harbor or matched safe harbor.

The calculations for determining the percentage of students in a disaggregated group meeting growth targets are described in Section 2 below. Briefly, each student is evaluated each year to determine whether they are on track to catch up (for unsatisfactory students) or keep up (for all other students). A percentage is calculated by dividing all students on track to catch up and keep up divided by the number of continuously enrolled students in the disaggregated student group, school, or district. These quantities must be compared to a target—or AMO in NCLB terms—to determine whether the districts/schools/disaggregated groups students made enough growth to make AYP.

NCLB provides only one method for establishing AMOs and intermediate goals. This is typically referred to as the 20<sup>th</sup> percentile approach where schools are rank-ordered based on the percentage of students scoring proficient, or meeting growth targets in this case, and identifying the school that cuts off the bottom 20 percent of the students. The percent of students scoring proficient (or meeting growth targets) in this school becomes the starting point for setting AMOs.

While this approach worked well in the early days of NCLB with 12 years until 100% of students were required to be proficient, Colorado believes that starting at the “20<sup>th</sup> percentile” at this point in the life of NCLB does not make sense because it would set low expectations in the first year, but would require unrealistic increases for schools in the coming years. Because we are just past the halfway point between 2002 and 2014, Colorado proposes setting the initial growth AMO at the “60<sup>th</sup> percentile” school using an adaptation of the methodology described for the 20<sup>th</sup> percentile approach. In this approach, CDE rank-ordered schools by level (elementary, middle, and high school) and content area according to the percentage of students in the school meeting growth targets. The associated number of students continuously enrolled in each school was included in the rank-ordered list. We then “counted up” from the school with the lowest percentage of students meeting growth targets until we accumulated 60% of the students. The percentage of students meeting the growth target associated with this 60<sup>th</sup> percentile school became the initial annual growth measurable objective (Growth AMO). This methodology results in the following initial targets for percentages of students meeting growth targets in order for the districts/schools/disaggregated groups to make AYP. Further, in order to avoid too many changes to the AYP system at once and to aid communication with Colorado’s school and district leaders, Colorado proposes including one intermediate goal in 2011 to parallel the 2011 intermediate goal for status calculations.

Table 1. Annual growth targets and intermediate goals.

<b>Year</b>	<b>Grade Span</b>	<b>Math</b>	<b>Reading</b>
2009	Elementary	67.69%	82.06%
	Middle	57.64%	81.90%
	High	58.62%	86.06%
2011	Elementary	83.84%	91.03%
	Middle	78.82%	91.45%
	High	79.31%	93.03%
<b>2014</b>	Elementary	<b>100.00%</b>	<b>100.00%</b>
	Middle	<b>100.00%</b>	<b>100.00%</b>
	High	<b>100.00%</b>	<b>100.00%</b>

Potential Impact on AYP Determinations

Colorado is proposing to include growth-to-standard determinations in AYP determinations to ensure coherence with Colorado’s focus on measures of individual student growth in state and local initiatives and accountability efforts. The impact on AYP determinations is much less important to CDE leaders than the opportunity to promote a coherent message about the importance of ensuring that every Colorado student is supported to grow to their full potential. Nevertheless, CDE recognizes the importance of examining and reporting the anticipated impact on AYP determinations as a result of implementing the Colorado Growth Model and including the results in AYP determinations.

First, we examined the simple school-level bivariate correlations between the performance/status (i.e., % partially proficient or greater) and the growth-to-standard results (i.e., % of students on track to catch up or keep up). These results are found in Table 2 below. The relationships were calculated for the whole school and for the free and reduced lunch students. We present the results for free and reduced lunch price students, in the interest of parsimony, because of the strong overlap between these students and many of the other student groups. In general, there is a moderate-to-strong relationship between the performance and growth results, but the correlations indicate that there is still a fair amount of variance unique to each metric. The strength of the correlations is certainly related to the requirement that the growth-to-standard metric meets Colorado’s interpretation of NCLB and associated regulations and guidance for all students on track to catch up or keep up within three years and to have all students proficient or on track to proficient by 2014.

The correlations were quite strong for reading at the whole school level for all grade levels and for math in middle school and high school. The correlations were moderate for reading when analyzing the performance of free and reduced price lunch students at all three grade spans and for elementary math. However, the correlations were quite strong for middle and high school math when analyzing free and reduced price lunch students.

Table 2. Pearson correlation coefficients between percent of students scoring at least partially-proficient (NCLB proficient) and percent of students meeting growth-to-standard criteria for whole school and free-reduced lunch students.

Grade Span	Reading		Math	
	Whole School	Free/reduced lunch	Whole School	Free/reduced lunch
Elementary	0.71	0.57	0.59	0.43
Middle	0.61	0.43	0.75	0.76
High	0.69	0.57	0.84	0.81

These moderate-to-strong correlations—but far from perfect—suggest that we might expect a fair proportion of schools to meet AYP requirements as a result of incorporating the growth measure into AYP determinations. This is not the case, however, in part because many of Colorado’s schools are already meeting the performance (status) or safe harbor targets, so there are relatively few schools that “need” to employ the growth measures to make AYP.

There are only a handful of Colorado schools that, if the Colorado Growth Model had been employed in 2008, would have made AYP because of the growth model results alone. That is, these schools would have missed the performance and safe harbor targets, but made the growth targets. In general, the growth targets are more rigorous so that far fewer schools meet the growth targets compared with the performance targets. The “outcomes” tab on the Microsoft Excel appendix (Attachment-3) documents this quite clearly for the different grade spans and for each of the disaggregated student groups.

Table 3 below provides an example of this phenomenon. Several things can be seen from this table. First, more schools meet the performance targets than the growth targets. Second, very few schools that miss the performance targets meet the growth targets. On the other many of the schools that miss the growth targets are able to meet the performance targets.

Table 3. Comparing the number/percent of schools meeting performance and growth targets for reading (whole school results only).

School overall	Made current performance target	Made proposed growth target		Total
		NO	YES	
Elementary Reading	NO	196	7	203
		96.6%	3.4%	
	YES	418	403	821
		50.9%	49.1%	
Total	614	410	1,024	
	60.0%	40.0%		
Middle School Reading	NO	75	1	76
		98.7%	1.3%	
	YES	221	196	417
		53.0%	47.0%	
Total	296	197	493	
	60.0%	40.0%		
High School Reading	NO	61	1	62
		98.4%	1.6%	
	YES	178	158	336
		53.0%	47.0%	
Total	240	159	398	
	60.2%	39.8%		

For a different, slightly more complex look at these interactions, we repeated the analyses presented in Table 3, but included the safe harbor and matched safe harbor results as well. We present these findings in Tables 4, 5, and 6 below. The results presented in these tables corroborate the results presented in Table 3 above, **but adding in safe harbor for math results in no additional schools meeting math AYP requirements as a result of including growth model results.**

Table 4. Comparing the number/percent of schools meeting performance, growth, and safe harbor targets for elementary math (whole school results only).

Made Performance target	made Safe Harbor	Made Matched Safe Harbor		Made Growth Target		Total
				NO	YES	
NO	no	Met Target	NA	3		3
				100.0%		100.0%
			NO	35		35
		100.0%			100.0%	
		YES	9	1	10	
			90.0%	10.0%	100.0%	
	Total		47	1	48	
			97.9%	2.1%	100.0%	
	yes	Met Target	NA	3		3
				100.0%		100.0%
			NO	60		60
		100.0%			100.0%	
		YES	21	5	26	
			80.8%	19.2%	100.0%	
Total		84	5	89		
		94.4%	5.6%	100.0%		
YES	no	Met Target	NA	18	24	42
				42.9%	57.1%	100.0%
			NO	131	71	202
		64.9%		35.1%	100.0%	
		YES	5	20	25	
			20.0%	80.0%	100.0%	
	Total		154	115	269	
			57.2%	42.8%	100.0%	
	yes	Met Target	NA	20	30	50
				40.0%	60.0%	100.0%
			NO	169	99	268
		63.1%		36.9%	100.0%	
		YES	78	121	199	
			39.2%	60.8%	100.0%	
Total		267	250	517		
		51.6%	48.4%	100.0%		

Table 5. Comparing the number/percent of schools meeting performance, growth, and safe harbor targets for middle school math (whole school results only).

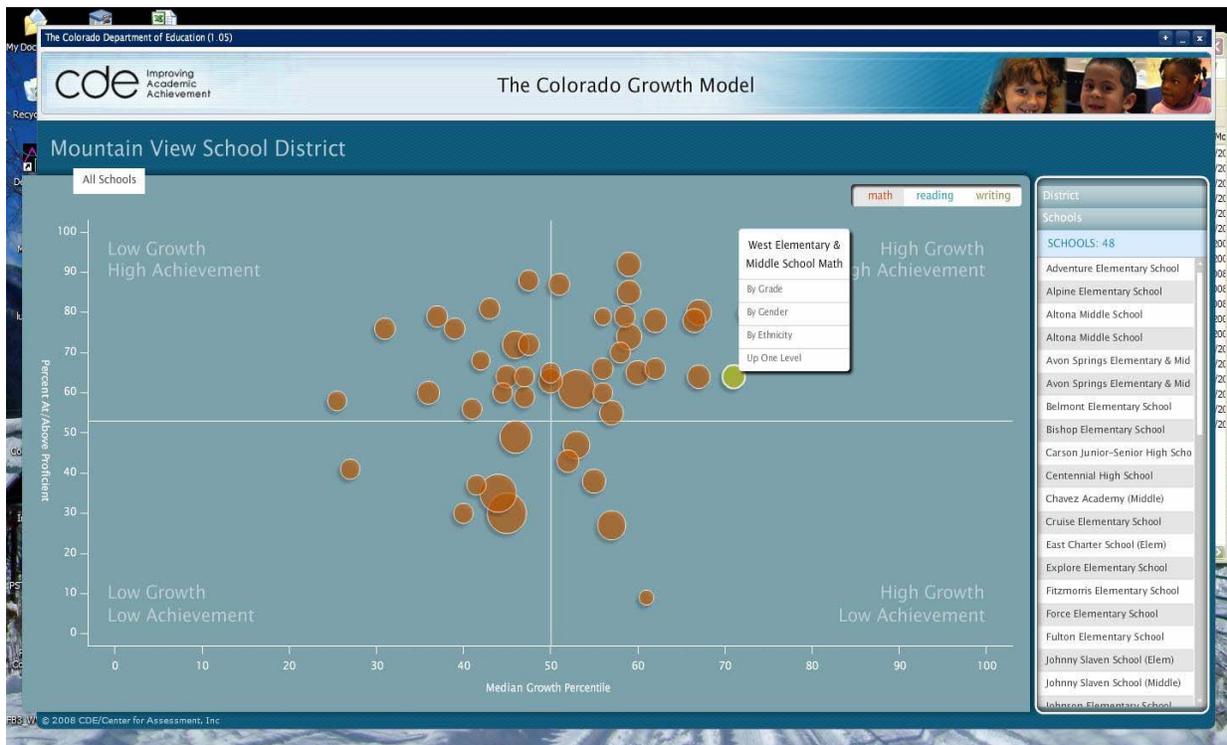
Made Performance target	Made Safe Harbor	Made Matched Safe Harbor		Made Growth Target		Total
				NO	YES	
NO	no	Met Target	NA	3		3
				100.0%		100.0%
			NO	48		48
		100.0%		100.0%		
		Total		51		51
				100.0%		100.0%
	yes	Met Target	NA	1	0	1
				100.0%	0.0%	100.0%
			NO	37	2	39
		94.9%	5.1%	100.0%		
		Total		38	2	40
				95.0%	5.0%	100.0%
YES	no	Met Target	NA	5	6	11
				45.5%	54.5%	100.0%
			NO	79	66	145
		54.5%	45.5%	100.0%		
		YES	0	1	1	
		0.0%	100.0%	100.0%		
		Total		84	73	157
				53.5%	46.5%	100.0%
		yes	Met Target	NA	1	7
	12.5%				87.5%	100.0%
	NO			54	63	117
	46.2%		53.8%	100.0%		
	YES		2	10	12	
	16.7%		83.3%	100.0%		
	Total		57	80	137	
			41.6%	58.4%	100.0%	

Table 6. Comparing the number/percent of schools meeting performance, growth, and safe harbor targets for high school math (whole school results only).

Made Performance target	made Safe Harbor	Made Matched Safe Harbor		Made Growth Target		Total
				NO	YES	
NO	no	Met Target	NA	1		1
				100.0%		100.0%
		NO	14		14	
			100.0%		100.0%	
	Total		15		15	
			100.0%		100.0%	
	yes	Met Target	NA	7	0	7
				100.0%	0.0%	100.0%
		NO	72	2	74	
			97.3%	2.7%	100.0%	
Total		79	2	81		
		97.5%	2.5%	100.0%		
YES	no	Met Target	NO	12	19	31
				38.7%	61.3%	100.0%
		Total		12	19	31
			38.7%	61.3%	100.0%	
	yes	Met Target	NA	2	1	3
				66.7%	33.3%	100.0%
		NO	31	80	111	
			27.9%	72.1%	100.0%	
		YES	2	10	12	
			16.7%	83.3%	100.0%	
Total		35	91	126		
		27.8%	72.2%	100.0%		

## Reporting growth results

Noted psychometrician Ron Hambleton likes to say, “the test score reports are the major way that we communicate with the public about the testing system, but the last thing we attend to when designing an assessment or accountability system.” Colorado has taken Hambleton’s words to heart and has designed a sophisticated (based on Adobe FLEX technology) yet user-friendly reporting system to present growth results to the public and key education stakeholders. Just released this year, this reporting system has been met with widespread praise. More importantly, the system enables educational leaders to see and understand the most important aspects of the growth and performance results. The following link provides an overview of growth percentiles and the reporting structure that Colorado has currently deployed (<http://www.cde.state.co.us/cdeassess/growthmodel.html>). A screenshot captured from this reporting tool is presented below. This view makes it clear that Colorado values both performance (status) and growth, demonstrating a method for reporting these results to the public in ways that can convey important messages about school effectiveness. We are currently adding the growth-to-standard results to this reporting structure so that Colorado’s educators will be able to receive a comprehensive view of school quality and improvement indicators.



Additionally, Colorado uses a straightforward approach for reporting AYP results. Currently, CDE provides school and district profiles for all schools in the state with detailed AYP

information.

(<http://www.cde.state.co.us/scripts/cfpu/NCLBProfiles0708/SchlDataEle.asp?SchlCode=1878&level=E&DISTCODE=0020>). These reports would be amended to include the growth data. See the Microsoft Excel appendix, the “Reporting Template” tab for a sample report of how to include growth information into the AYP report.

#### **1.4 Does the proposed growth model include a relationship between consequences and rate of student growth consistent with Section 1116 of ESEA?**

Because Colorado intends to fully incorporate the proposed growth model into its AYP framework (as described above), the Section 1116 consequences for being identified in need of improvement apply consistently no matter what the reason for the school or district’s failure to make AYP. However, Colorado is currently building different actions into its state initiatives, which recognize the constitutional role of districts in Colorado balanced by the State’s interest in ensuring continuous improvement and the need for all students to graduate ready for postsecondary and workforce success.

Colorado’s objective is a coherent accountability system that provides consistent performance incentives for all schools to maximize each student’s annual progress toward proficiency and postsecondary and workforce readiness. The “four-quadrant” depiction presented earlier in this document provides a helpful heuristic for Colorado to organize its responses to schools/districts depending on the quadrant in which a district or school is located. This depiction provides a useful point of departure for a more in-depth analysis of district and school performance supported by the drill-down functionality of the data display tools, district accreditation results, and the school performance measures anchored by the Colorado Growth Model.

District Accreditation measures, which focus on individual student growth and status, gaps in growth and status among subgroups, and postsecondary readiness, provide a balanced perspective on district-wide strengths and weaknesses. School performance measures, as reflected in this AYP proposal, also center on growth and status and the performance of disaggregated student groups. Together, these district and school measurement systems provide the cumulative performance information that stakeholders need to understand where improvements are necessary and the intensity and kind of support required.

Armed with this understanding, CDE is positioned to allocate levels of service and support to districts and schools according to the severity and nature of their needs. The support provided by CDE for districts and schools is built around the Comprehensive Appraisal for District Improvement (CADI) and a parallel process for schools called the School Support Team (SST) review. Both CADI and SST are federally funded programs and followed up with implementation grants. In 2008, CDE introduced a new, state-funded pilot program -- Closing the Achievement Gap (CTAG) to provide intensive support to selected districts. This program incorporates the CADI process and extends it to the provision of additional resources focused on closing achievement gaps among disaggregated students groups.

## Core Principle 2: Establishing Appropriate Growth Targets at the Student Level

### 2.1 Has the State proposed a technically and educationally sound method of depicting annual student growth in relation to growth targets?

#### In-depth description of the model

The Colorado Growth Model provides a means to understand how individual students and groups of students progress from year to year toward state standards based on where each student begins and relative to the progress of other students in the state. The growth model focuses attention on maximizing student progress over time and reveals where, and among which students, the greatest growth is happening—and where it is not. It recognizes that the most effective schools are those that produce the highest sustained rates of growth in student progress. Those schools may or may not be schools with the highest test scores every year.

CDE developed the growth model to answer three essential questions about student, school and district performance:

- **What is?** What is the academic growth of a student?
- **What should be?** What should the academic growth be for a student to reach a desired level of achievement within a period of time?
- **What could be?** What are the highest sustained student growth rates to date and under what conditions could they improve?

The model addresses these questions in two ways: (1) by calculating individual *student growth percentiles* for each Colorado student with at least two consecutive year/grade CSAP scores and (2) by calculating individual *percentile growth trajectories* which indicate how much growth it will take to reach each of Colorado's three performance level cut-points in one, two, and three years. Student growth percentiles and percentile growth trajectories combine a description of growth (What is?) with individualized prescriptions for how much growth is required to reach future achievement goals (What should be?). Providing a common probabilistic, percentile-based metric for all descriptions identifies what is possible (What could be?) and aids in identifying exemplary student growth together with schools, programs, and district with which it is associated.

A *student growth percentile* defines how much relative growth a student made. The Colorado Growth Model serves as a way for educators to understand and communicate about how much growth a student makes from one year to the next relative to a student's "academic peers." More specifically, the Colorado Growth Model compares each student's performance to students in the same grade throughout the state who had similar CSAP scores in past years. The result is a student growth percentile, much like children's height and weight percentiles that pediatricians share with parents<sup>2</sup>. If a student grew as well or better than 60 percent of her academic peers, she

---

<sup>2</sup> In important distinction between the growth percentiles used by pediatricians and the growth percentiles calculated in the Colorado Growth Model is that pediatricians' growth percentiles are conditions on the child's age and sex,

would have a growth percentile of 60. Individual student growth percentiles are categorized as low (1 to 35), typical (36 to 65), or high (66-99).

Despite their utility in *describing* progress, student growth percentiles fail in and of themselves to define whether a student's growth is *adequate* or *good enough*. Judging the adequacy of a student's growth percentile is a standard setting process requiring external criteria such as whether the student's growth puts them on track to reach proficient. For example, a student might have a growth percentile of 75, but even though they are doing better than 75 percent of their academic peers, their growth might not be good enough to reach proficiency within the next three years. Establishing standards for growth requires an analysis of what level of growth is necessary to reach desirable achievement goals within a reasonable time frame.

To establish these growth standards, Colorado looks at the entire state's data to examine how much students have grown in the current year to determine how much students need to grow to reach pre-established achievement targets. Specifically, the Colorado Growth Model uses the most recent results from statewide analyses calculating each student's growth percentile to define *percentile growth trajectories* for each student. Percentile growth trajectories depict what a student's future achievement will be assuming annual growth percentiles from 1 to 99 compounded for one, two, and three years. Based upon these trajectories, it is possible to define what level of growth is necessary for a non-proficient student, for example, to reach proficient status within 3 years. Similarly, it is possible to define what level of growth is necessary for a proficient student to maintain proficient status or to progress to advanced status. Moreover, these standards are communicated with the same percentile metric used for student growth percentiles. Again, please refer to the Colorado Growth Model Tutorial (Attachment 1) for several examples of how this works in practice.

The NCLB growth criteria presented in this application represent one set of growth standards that can be used to judge the quality of schools. Colorado's state accountability model is currently testing a set of more ambitious standards requiring students in higher achievement levels to be on track to move up. Eventually, these growth standards will form the basis of adequacy determinations regarding student growth. Details on the methodology used to calculate student growth percentiles and percentile growth trajectories are supplied in the Adobe PDF attachment written by Dr. Damian Betebenner.

#### Description of how Colorado established "sound criteria" for student level growth targets

As described in sections 1.2 and 2.1 above, Colorado believes it has set ambitious, yet achievable, growth targets for all students and that these growth targets are technically defensible. The analyses presented in this document indicate that very few schools which otherwise would have not made AYP, will make AYP as a result of the growth model. Given these results, some might view these targets as too ambitious, but the Colorado Department of Education and other key policy makers in the state are not willing to back down from these targets. Colorado is committed to focusing schools and districts on improving the performance of every student over time.

---

whereby the growth percentiles calculated here are conditioned on all available prior achievement test scores (in the same content area) so that the student is being compared to students with essentially the same academic history.

That being said, CDE will conduct ongoing analyses to evaluate the validity of the growth targets at the student, student group, and school levels. These analyses will entail using the growth results as useful information in Colorado's long-term plans to update its content and achievement standards to ensure that Colorado students graduate high school to be college and work ready.

## **Core Principle 3: Accountability for Reading/Language Arts and Mathematics Separately**

### **3.1 Has the State proposed a technically and educationally sound method of holding schools accountable for student growth separately in reading/language arts and mathematics?**

The Colorado Growth Model calculates growth targets and evaluates schools separately for reading and mathematics. Further, while the growth targets are calculated uniquely for each student based on all prior subject-specific test scores, all student groups, schools, and districts are held to the same growth AMOs for each reading and mathematics.

## Core Principle 4: Inclusion of All Students

### **4.1 Does the State’s growth model proposal address the inclusion of all students, subgroups and schools appropriately?**

#### Description of how missing data are handled

As with any analysis of large scale observational data, missing or incomplete data is an issue. Calculation of the student growth percentile corresponding, for example, with a student’s 2008 CSAP score, is based upon estimation of a conditional density. Because the purpose is to describe growth in the most recent year, at a minimum, it is necessary to have the student’s prior year’s score. Operational decisions are made that define which student scores are part of the conditioning population and which are not:

- Students must have at a minimum two CSAP scores in the same content area coming from consecutive years and grades (e.g., grade 4 math in 2007 and grade 5 math in 2008).
- Only CSAP data from consecutive grades are used to calculate student growth percentiles.
- Students repeating/skipping grades associated with the latest score (of which there are very few) are not included in the analyses at this time.
- Only CSAP data from consecutive years are used to estimate the conditional density for a student. For example, a student possessed 2005, 2007, and 2008 CSAP reading data, only scale scores from 2007 and 2008 are used to calculate their student growth percentile. The number of students with such “holes” in their data is extremely small (less than 1%). Moreover, examination of students with complete data indicates that even though the inclusion of prior scores can lead to different student growth percentiles, growth percentiles based upon fewer prior scores are unbiased with regard to those derived using the maximum possible prior scores.
- Operationally, students with two CSAP scores in consecutive grades and years receive a student growth percentile conditioning upon the single prior score; students with three CSAP scores in consecutive grades/years receive a student growth percentile conditioning upon the prior two scores; students with four CSAP scores in consecutive grades/years receive a student growth percentile conditioning upon the prior three scores; and so on. After all possible growth percentiles are calculated for students, following HB 07-1048 stipulations, the growth percentile based upon the maximum prior data is assigned to the student. Thus, students with exactly two consecutive CSAP scores are compared with all other students with at least 2 scores; students with exactly 3 consecutive CSAP scores are compared with all other students with at least 3 scores; and so on. Student growth percentiles are conditioned relative to all students possessing scores on the same number of tests as the student and conditioned on the prior score histories.

#### Minimum-n and accountability for student groups

Colorado is applying the same relatively low minimum-n of 30 students that is currently used in Colorado’s AYP determinations for status and safe harbor. CDE is convinced that n=30 ensures the correct balance of inclusion, reliability, and validity, especially considering the diverse demographics of Colorado from major cities on the Front Range to small rural towns on the

eastern plains. As can be seen under the “NSize” tab in the Microsoft Excel attachment, there is very little difference in the number/percent of schools held directly accountable for specific student groups under performance (status) and under growth with a minimum-n of 30. This is due, in part to our extremely high match rates (discussed below). Growth models contain more error than status measures because of the use of multiple imperfectly correlated assessments. While we would have liked to use a minimum-n less than 30 to include more student groups in direct accountability, concerns about threats to the reliability when evaluating student groups suggests that we maintain a minimum-n of at least 30. Further, CDE believes that the consistency among the performance, growth, and safe harbor calculations is an important reason for maintaining the minimum-n of 30 students.

### Match rates

One of the risks with implementing a growth model where students are required to have at least two valid test scores in order to be included in the accountability is that there will never be more students included in the growth system than a status model. CDE has examined the match rates for students in elementary, middle, and high school (presented under the “match rate” tab in the Microsoft Excel attachment). As expected the match rates for elementary schools are spuriously low because third graders are not included in the growth calculations. Middle school match rates provide a more realistic picture of the match rates for students in grades 4-8. As seen in the attachment, the match rate for all student groups generally range from 95% to 98%. The match rates for high schools are only slightly lower than for middle schools. The one student group with noticeably lower match rates than the other groups is the students with disabilities student group. We suspect this is due to a variety of factors. First almost 10% of students with disabilities participate in the Colorado Student Assessment Program Alternate Assessment (CSAP-A) and at this time, CDE is unable to include CSAP-A scores in the Colorado Growth Model. Further, a small percentage of these students move in and out of the alternate assessment and generally have higher rate of school absence compared with other student groups. Importantly, except for this slightly lower match rate for students with disabilities, all other student groups had essentially the same match rates at all three grade spans.

### Full Academic Year

Colorado has been using the same definition of Full Academic Year (FAY) since NCLB was first enacted. A student in Colorado is considered continuously enrolled if they had been in the school/district from one CSAP administration to the next. As documented in the match rates presented above, this definition of FAY works very well with the Colorado Growth Model, because the school is held accountable for essentially all of the same students whether using a performance (status) or growth model.

### Alternate assessment students

As much as Colorado would like to include every student in the Colorado Growth Model, we have not yet determined how to appropriately include students participating in the CSAP-A in the growth calculations. The Colorado Growth Model takes advantage of the CSAP vertical scale for calculating and evaluating targets and since CSAP-A is on a different scale, we are unable to include the scores of students with significant cognitive disabilities at this time. CDE, however, is committed to finding a way to include all students in the growth model and this is at the top of CDE’s research agenda.

### Retained Students

Current year retained students, of which there are very few in Colorado—are not included in the growth model. Students retained in the most recent year are not included in the current year analysis due to difficulty in comparing the students' scores against other students taking the same exam. Students repeating grades in the past are included in the analyses using as much prior data coming from consecutive grades as possible. Retained students, however, are included in performance calculations as well as in safe harbor and matched safe harbor.

### Grade 3 Students

As noted earlier in this proposal, grade 3 students are not included in the Colorado Growth Model because they do not have at least two test scores.

### Small schools and uniquely configured schools

CDE will not be able to extend the growth component of AYP to K-2 schools. Currently, Colorado's state assessment system (CSAP and CSAP-A) does not include students until grade 3. AYP is defined differently for Colorado's K-1 and K-2 schools than for those schools containing grades 3 and higher. K-1 and K-2 school AYP is determined by using the third grade reading and math scores of students previously enrolled at the school. K-1 and K-2 schools will be held to the elementary school AYP targets for accountability purposes. All schools will be expected to yield annual results that meet the requirement of 100% proficiency in reading and math by 2013-2014. However, to wait until 4<sup>th</sup> grade growth scores are available and then to go back and attribute those scores to the student's K-1 or K-2 school is not appropriate. Schools with a single grade, except for third grade, will have the growth targets included in their AYP calculations.

## **Core Principle 5: State Assessment System and Methodology**

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

#### Description of state assessment system

Colorado's fully approved and stable assessment system is a key feature of this proposal. Colorado was one of the earliest states to receive full approval under the NCLB Standards and Assessment Peer Review. Moreover, Colorado did not wait for the passage of NCLB to implement every grade testing. Colorado's early and strategic commitment to the measurement of student longitudinal growth led Colorado to implement CSAP testing in grades 3-10 as early as 2002. Further, Colorado's Model Content Standards and CSAP have remained fairly stable since 2002.

The test development process for CSAP requires "up-front" alignment (as well as post-hoc alignment) for both content and process dimensions. These assessments are designed in a way that allows students to demonstrate content knowledge through activities described in the standards and, importantly, to allow the assessment of higher order thinking skills. Test items are developed within the range of Depth of Knowledge (DoK) specified in the content standards. The State's assessments have sufficient items at each achievement level to permit students to demonstrate the full range of the State's academic achievement standards. This is an important consideration for the implementation of a growth model.

### **5.2 How will the State report individual student growth to parents?**

#### Individual student reports

As described above, CDE has launched an ambitious and innovative reporting system so that educators, parents, and students will be able to understand and utilize CSAP results, both in terms of growth and performance (status). Further, school and district level data will be available to parents in the Colorado School and District Profiles. A template of this report is included in the Microsoft Excel attachment, on the "Reporting template" tab.

### **5.3 Does the Statewide assessment system produce comparable information on each student as he/she moves from one grade level to the next?**

The State assessment system – that is the achievement levels and content expectations – needs to make sense from one grade to the next, as well as within achievement levels, for an assessment system to support a growth model. The stability of Colorado's assessment design and the long standing use of a vertical scale contribute to the assessment system's capacity to support the measurement of student growth. While the vertical scale—discussed in more detail below—is an important feature of CSAP, maintaining the year-to-year comparability for each grade level test is critical to support the validity of the vertical scale and for supporting growth measures. Toward this end, Colorado has adopted a very robust equating design whereby approximately 35% of the items on a grade level assessment are part of the set of linking item. This ensures

that scores from any given year lead to valid inferences when placed on the previous year's test scale.

### The vertical scale

A vertical scale is not necessary for calculating growth percentiles. However, having the vertical scale permits more user-friendly representations of achievement over time, allowing more intuitive graphical depictions of growth and its quantification as a growth percentile with comparison growth-to-standard percentiles. A full description of the design, implementation, and equating for the vertical scale can be found in the CSAP technical manuals found on CDE's website at:

[http://www.cde.state.co.us/cdeassess/documents/res\\_eval/FinalLongitudinalGrowthTAPReport.pdf](http://www.cde.state.co.us/cdeassess/documents/res_eval/FinalLongitudinalGrowthTAPReport.pdf)

Horizontal equating within each grade is used to place each year's forms on the vertical scales that had been established previously for reading and mathematics. The vertical scale for reading, spanning grades 3 through 10, was established in 2001, while the scale for Mathematics, spanning grades 5 through 10, was established in 2002. The mathematics assessments in grades 3 and 4 were added to the vertical scale in 2005. Stocking and Lord's (1983) procedures were used to place each grade's test on the vertical scale that had been developed for each content area.

As noted above, each year's CSAP tests contain a linking set of approximately 17–25 multiple-choice items pre-selected from previous administrations for the same grade to ensure that the same scale is maintained for each grade level test. These repeated multiple-choice items served as anchors in the Stocking and Lord's (1983) equating procedure, which was used to place each test form on the previously established scale. By equating the yearly CSAP tests within each grade, the unique metrics of the CSAP reading and mathematics vertical scales are maintained.

### The achievement standards

The Colorado Student Assessment Program (CSAP) measures student performance in reading, writing, mathematics and science relative to the Colorado Model Content Standards. The content standards articulate challenging content that Colorado students are expected to know and be able to do.

All Colorado school districts are required by law to adopt content standards which meet or exceed the Colorado Model Content Standards and develop a plan for revising curriculum and programs of instruction to align with the Colorado Model Content Standards to ensure that each student will have educational experiences needed to achieve the adopted content standards (CRS 22-7-407).

Achievement levels describe the success a student has achieved on the Colorado Model Content Standards tested on the CSAP Reading/Writing and Mathematics (grades 3-10) and Science (grades 5, 8 and 10). The following table provides the broad policy definitions that serve as the first level of the achievement level descriptions. Each grade level test includes a set of content-specific achievement level descriptors to convey to the public the meaning of the achievement levels. These can be found at:

[http://www.cde.state.co.us/cdeassess/documents/csap/csap\\_plds.html](http://www.cde.state.co.us/cdeassess/documents/csap/csap_plds.html).

### Achievement Level Descriptions

Advanced (4)	A student scoring at the Advanced Level has success with the most challenging content of the Colorado Model Content Standards. These students answer most of the test questions correctly, including the most challenging questions
Proficient (3)	A student scoring at the Proficient Level has success with the challenging content of the Colorado Model Content Standards.
Partially Proficient (2)	A student scoring at the Partially Proficient Level has limited success with the challenging content of the Colorado Model Content Standards. These students may demonstrate inconsistent performance, answer many of the test questions correctly but are generally less successful with questions that are most challenging.
Unsatisfactory (1)	A student scoring at the Unsatisfactory Level has little success with the challenging content of the Colorado Model Content Standards.

These achievement levels have been stable since 2001 for reading and 2002 for math grades 5-10, with grades 3 and 4 added in 2005. The following table provides the state level results from the most recent (2008) CSAP administration.

Table 7. Percent of students scoring in each of the achievement levels on the 2008 CSAP.

Reading	Unsatisfactory	Partially Proficient <sup>3</sup>	Proficient	Advanced	No Score
3 <sup>rd</sup> Grade	10.88%	18.38%	63.82%	6.51%	0.41%
4 <sup>th</sup> Grade	10.25%	23.53%	61.70%	4.17%	0.36%
5 <sup>th</sup> Grade	11.49%	17.91%	60.73%	9.48%	0.38%
6 <sup>th</sup> Grade	9.10%	19.20%	59.06%	12.08%	0.56%
7 <sup>th</sup> Grade	11.86%	22.00%	56.16%	9.29%	0.69%
8 <sup>th</sup> Grade	10.89%	21.02%	57.43%	9.64%	1.02%
9 <sup>th</sup> Grade	8.65%	22.88%	59.96%	6.14%	2.37%
10 <sup>th</sup> Grade	9.86%	20.85%	55.22%	10.85%	3.22%

Math	Unsatisfactory	Partially Proficient	Proficient	Advanced	No Score
3 <sup>rd</sup> Grade	7.99%	22.04%	39.84%	29.81%	0.32%
4 <sup>th</sup> Grade	8.83%	22.50%	42.29%	26.06%	0.32%
5 <sup>th</sup> Grade	8.38%	26.16%	37.23%	27.96%	0.27%
6 <sup>th</sup> Grade	12.78%	25.81%	36.61%	24.42%	0.38%
7 <sup>th</sup> Grade	18.29%	35.00%	27.82%	18.37%	0.52%
8 <sup>th</sup> Grade	22.76%	29.50%	26.52%	20.41%	0.81%
9 <sup>th</sup> Grade	30.44%	29.83%	24.48%	13.32%	1.92%
10 <sup>th</sup> Grade	31.99%	34.88%	25.39%	4.98%	2.75%

<sup>3</sup> Note: Colorado's partially proficient level has been used as "NCLB-proficient" since NCLB was first enacted. This has been approved as part of Colorado's accountability and assessment peer reviews.

As seen in Table 7, the percent of students scoring within each achievement level is fairly consistent across grade levels in reading. However, this is not the case for mathematics where an increasing percentage of students score in the unsatisfactory level from 6<sup>th</sup>, but especially 7<sup>th</sup>, grade through high school. CDE recognizes this pattern and attributes it to two main causes. First, students' course-taking patterns in mathematics begin to differentiate in middle school, but Colorado has chosen to target the grade-level census test at a rigorous level instead of the "least common denominator" approach. As a local control state, CDE has very little to no control over local curriculum and students' course-taking patterns, but is determined to make the expectations clear through the rigor of the assessment. Therefore, the increasing percentage of students scoring at the unsatisfactory level reflects, in part, differences in opportunity to learn the required content. CDE, to the extent possible in a local control state, is trying to address this issue through its school improvement efforts. Second, the standard setting panelists in mathematics and Colorado education leaders believed in the need to ramp up middle and high school expectations in mathematics in order to adequately prepare students for college. The effect of the increasing rigor of the achievements standards in a growth-to-standard context means that, all things being equal, it is considerably more challenging for students to "catch up and keep up" in middle and high school mathematics than it is in reading and elementary math. CDE is fully aware of this issue and believes that implementing the Colorado Growth Model will help shine a light on the need for students to meet higher standards in math as they progress through high school.

#### **5.4 Is the Statewide assessment system stable in its design?**

As noted above, Colorado has one of the most stable every-grade assessment systems in the country. CSAP and the Colorado Model Content Standards, on which CSAP is based, have been in place since the early years of IASA and every-grade testing has been in place since 2001 and 2002 for reading and mathematics, respectively.

## Core Principle 6: Tracking Student Progress

### **6.1 Has the State designed and implemented a technically and educationally sound system for accurately matching student data from one year to the next?**

#### Unique Student Identifier

In addition to a stable, high quality assessment system and a powerful framework for calculating student growth metrics, Colorado's system for tracking individual students over time is a strength of this proposal. CDE is able to match essentially all of the students in the state with prior records, which is crucial for implementing a fair and valid growth model.

Colorado uses a numeric student identification system called the Record Integration Tracking System (RITS). The State Assigned Student Identification number (SASID) is a ten digit numeric identifier that is unique for each student. The SASID remains with the student as they move from one school to another within a district, or from one district to another within the State.

Each SASID is matched with over 20 demographic characteristics for each student in order to provide quality control/quality assurance of the SASID. Some of this demographic information is used when verifying matches of names to SASIDs during assessment data review and cleanup processes in which districts participate each spring. Once the SASID quality is assured, only five data elements; first name, middle name, last name, date of birth, and gender are maintained in the RITS.

RITS system performance reports are generated each quarter to examine total case volume and total number of each type of case, in order to detect unexpected fluctuations in volume. A report showing duplicate records on the Master Student Index (MSI) is generated monthly in order to detect and resolve instances of duplicate SASID assignments.

Examining the results from 2007 provides a useful case study to document the efficacy of RITS. In 2007, there were a total of 2,985,771 records submitted to RITS. Of these 2,930,244 (98.1%) were automatically matched to existing records. Another 9,933 (0.3%) records were automatically incorporated into RITS as new additions. Approximately forty-thousand (45,594 or 1.5%) cases were resolved manually. These rates have been fairly stable ever since RITS was first put in place. Therefore, it is safe to say that Colorado's student tracking system adds to the validity of the Colorado Growth Model.

#### Data Warehouse

Initiated in 2001, the Colorado Department of Education's Enterprise Data Warehouse continues to mature and evolve. At the present time, the data warehouse and associated reporting systems provide detailed information on these subjects:

- Accreditation (District-Level)
- Adequate Yearly Progress AYP (District, School, Student-Level)
- Assessment (District, School, Student-Level)
  - Annual Performance measurements (scale scores)
  - Longitudinal Growth measurements (growth percentiles)

- Graduation/Drop-out data
- Safety and Discipline
- School Accountability
- School/District Finance
- School/District Staff and Administrators
- Special Education student and staff data

In 2002 a state-wide student identifier system was implemented and the Department began collecting identifiable student-level data. The SASID (State Assigned Student Identifier) is the key for tracking students across district boundaries and throughout their PreK-12 history. Using the SASID, disparate student information is unified within the data warehouse, providing a complete historical record for each Colorado student.

CDE's Data Warehouse continues to evolve. Most recently, data infrastructure enhancements contributed to the successful generation and launch of the Colorado Growth Model. Built upon a star-schema architecture, the warehouse combines rich historical data with technical flexibility to accommodate the requirements of present and future data analysis projects.

## **CORE PRINCIPLE #7: Participation Rates and Additional Academic Indicator**

### Participation Rate

All schools, districts, and disaggregated groups with 30 or more students, need to meet the 95% participation rate in both reading and math in order to make AYP. Basically, it is a pre-requisite for making AYP. The inclusion of the growth model does not change the pre-requisite of meeting the 95% participation rate in order to make AYP.

### Additional Academic Indicator

Colorado holds elementary and middle schools to an increasing target of students scoring **advanced** in reading and math, as the “Additional Academic Indicator.” High schools are accountable for the graduation rate, as per NCLB. All schools, districts, and disaggregated groups with 30 or more students need to meet the additional academic indicator in both reading and math, or the graduation rate (for high schools), in order to make AYP. The additional academic indicator or “Other Indicator” is not affected as a result of this proposal.