COLORADO DEPARTMENT OF EDUCATION

201 East Colfax Avenue • Denver, Colorado 80203-1799
303.866.6600 • www.cde.state.co.us

Dwight D. Jones
Commissioner of Education

Robert K. Hammond
Deputy Commissioner

Kenneth R. Turner
Deputy Commissioner

October 15, 2008

Secretary Margaret Spellings
U.S. Secretary of Education
U.S. Department of Education
400 Maryland Ave., SW
Washington DC. 20202-6132

Honorable Secretary Spellings:

The Colorado Department of Education (CDE) appreciates the opportunity to submit its longitudinal growth model proposal for your consideration. CDE has worked with stakeholders across the state and with national experts to develop this model. We believe it meets the standards you have set forth at the U.S. Department of Education. The approach proposed by Colorado focuses on maximizing all students' progress toward meeting state standards, including those students scoring below proficient and those already proficient.

Again, thank you in advance for your consideration of our proposal.

Sincerely,

Dwight D. Jones
Commissioner of Education
State of Colorado

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 deliberatively 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.

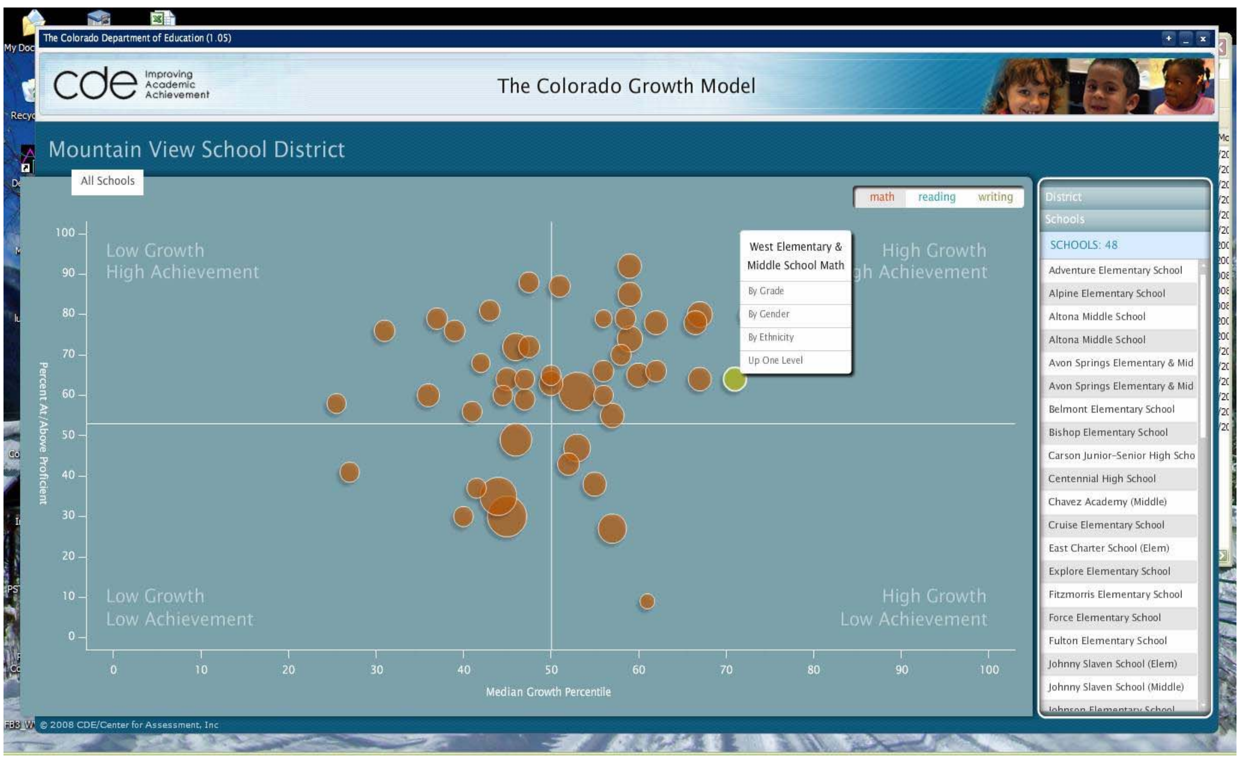
The State Board of Education required, based upon the advisory panel recommendations, that the Colorado Growth Model calculate a growth percentile for each student relative to all other students in the state with the same prior academic history (academic peers). Subsequently, Colorado's District Accreditation 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 20th 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 “20th 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 “60th percentile” school using an adaptation of the methodology described for the 20th 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.

Year	Grade Span	Math	Reading
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%
2014	Elementary	**100.00%**	**100.00%**
	Middle	**100.00%**	**100.00%**
	High	**100.00%**	**100.00%**

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 10th 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 21st 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). 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 1st 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 10th grade, whichever comes first,*
- *The student is NCLB-proficient and is on track to maintain NCLB-proficient for the upcoming 3 years or by 10th grade, whichever is first,*
- *The student is Colorado-proficient and is on track to maintain Colorado-proficient for the upcoming 3 years or by 10th grade, whichever is first, or*
- *The student is advanced and is on track to maintain Colorado-proficient for the upcoming 3 years or by 10th 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[1].	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

[1] 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 <u>and</u> 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 20th 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 "20th 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 "60th percentile" school using an adaptation of the methodology described for the 20th 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 60th 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.

Year	Grade Span	Math	Reading
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%
2014	Elementary	**100.00%**	**100.00%**
	Middle	**100.00%**	**100.00%**
	High	**100.00%**	**100.00%**

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: NO	Made proposed growth target: YES	Total
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 NO	Made Growth Target YES	Total
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).

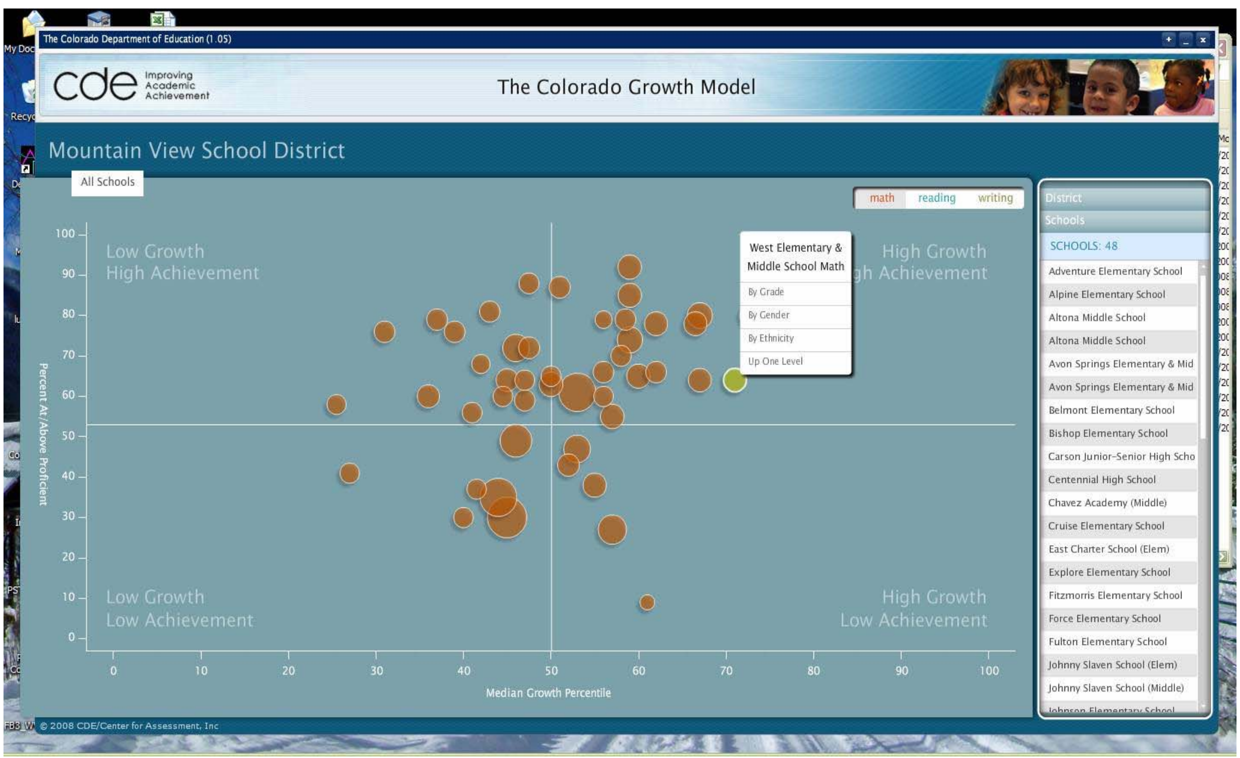
Made Performance target	Made Safe Harbor	Made Matched Safe Harbor		Made Growth Target NO	Made Growth Target YES	Total
NO	no	Met Target	NA	3		3
				100.0%		100.0%
			NO	48		48
				100.0%		100.0%
		Total		51		51
				100.0%		100.0%
	yes	Met Target	NA	1	0	1
				100.0%	0.0%	100.0%
			NO	37	2	39
				94.9%	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	8
				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 NO	Made Growth Target YES	Total
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/scriptscfpu/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[2]. If a student grew as well or better than 60 percent of her academic peers, she

[2] 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.

<u>Description of how Colorado established "sound criteria" for student level growth targets</u>

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^{th} 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

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.

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

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

[3] 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 6th, but especially 7th, 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

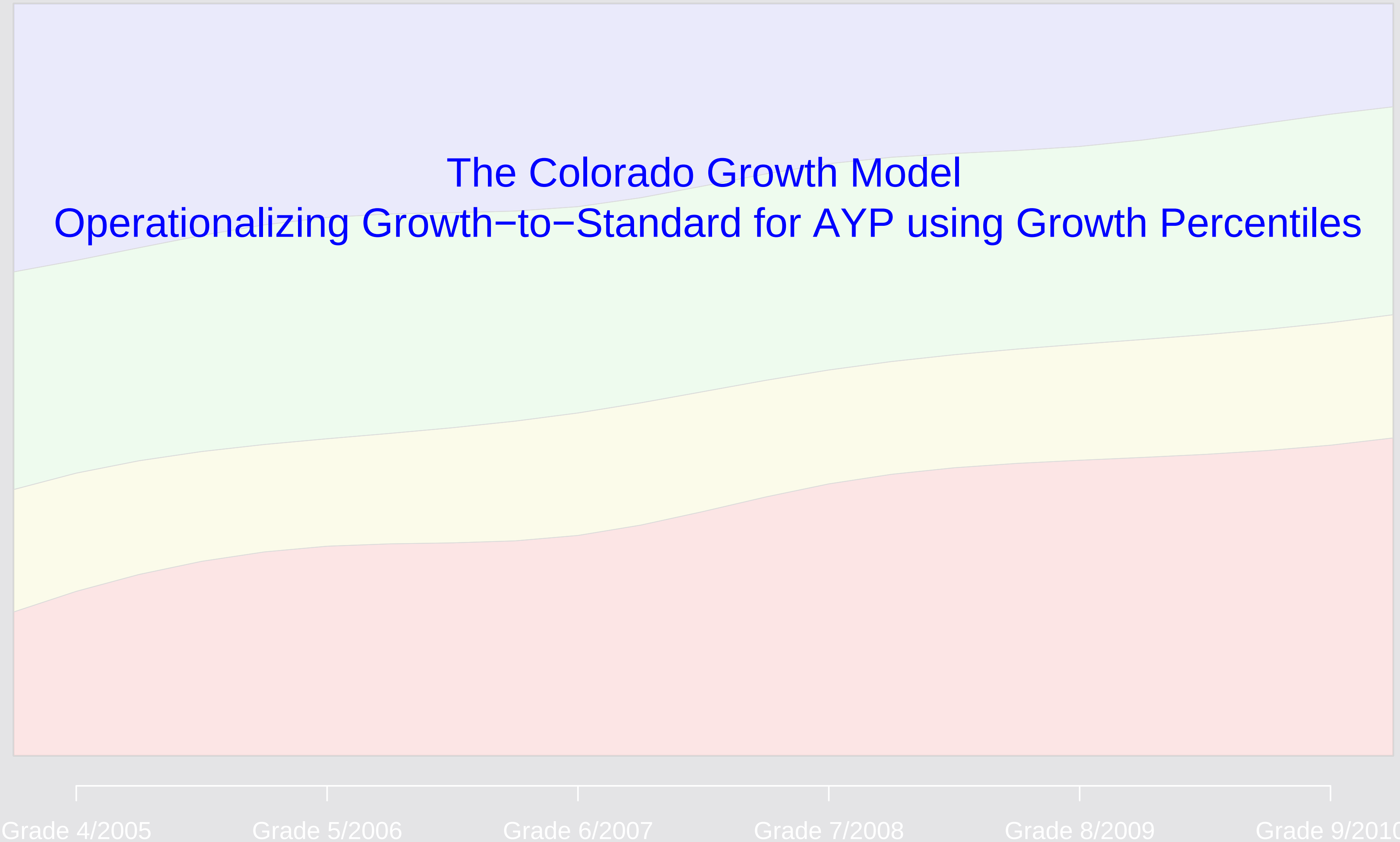
All schools, districts, and disaggregated groups with 30 or more students, need to meet the 95% participation rate in both reading and math in order to make AYP. Basically, it is a pre-requisite for making AYP. The inclusion of the growth model does not change the pre-requisite of meeting the 95% participation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

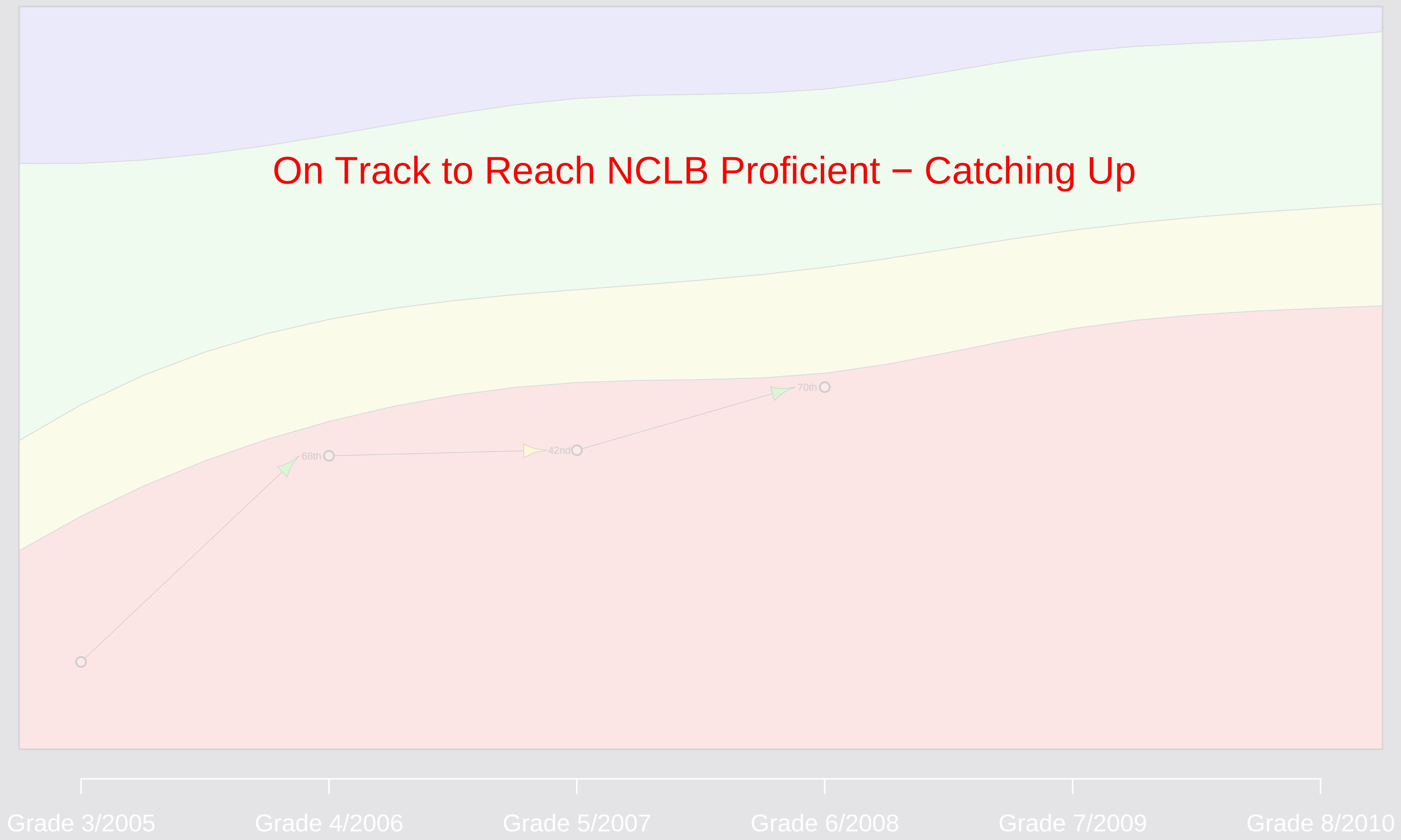
The Colorado Growth Model

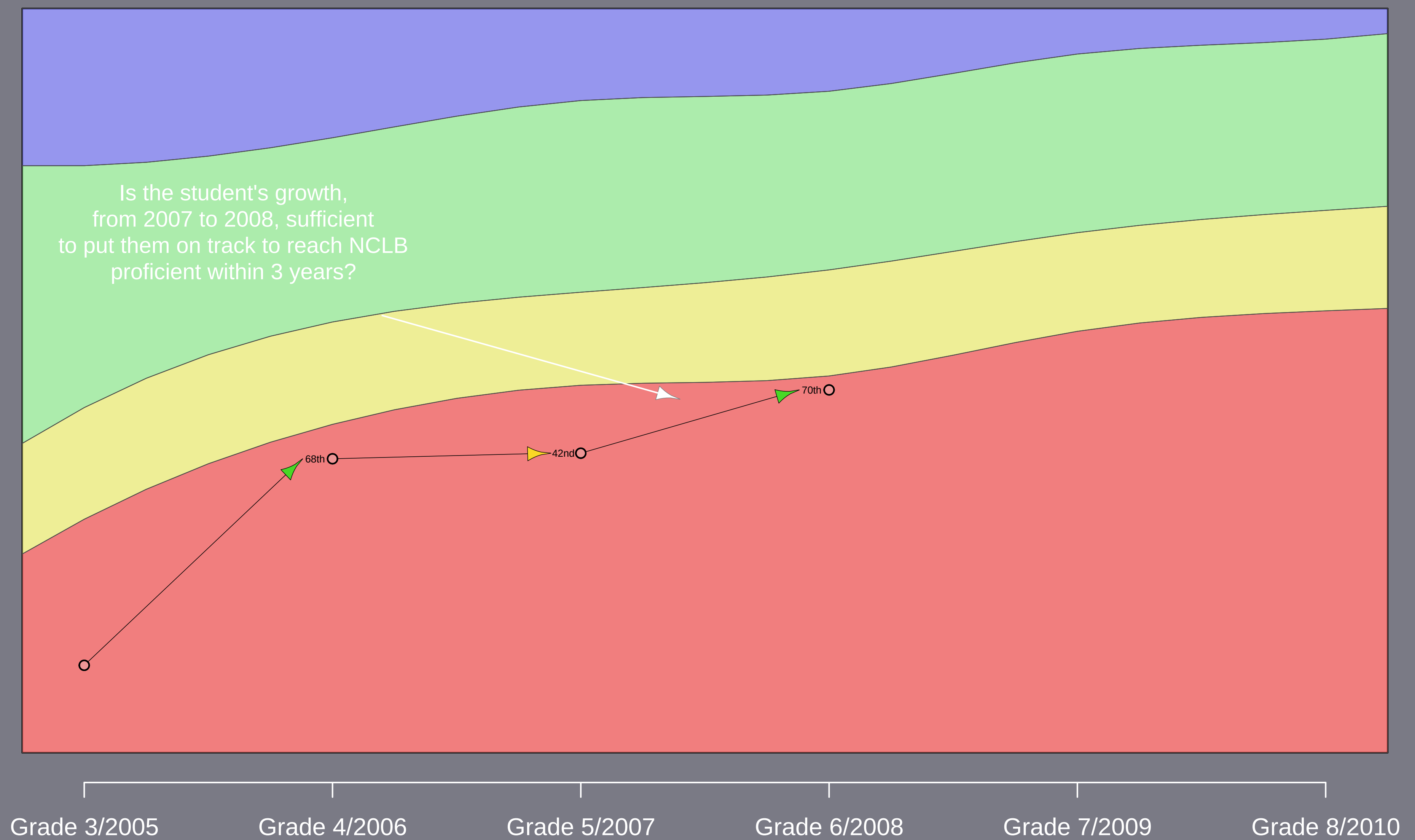
Operationalizing Growth−to−Standard for AYP using Growth Percentiles



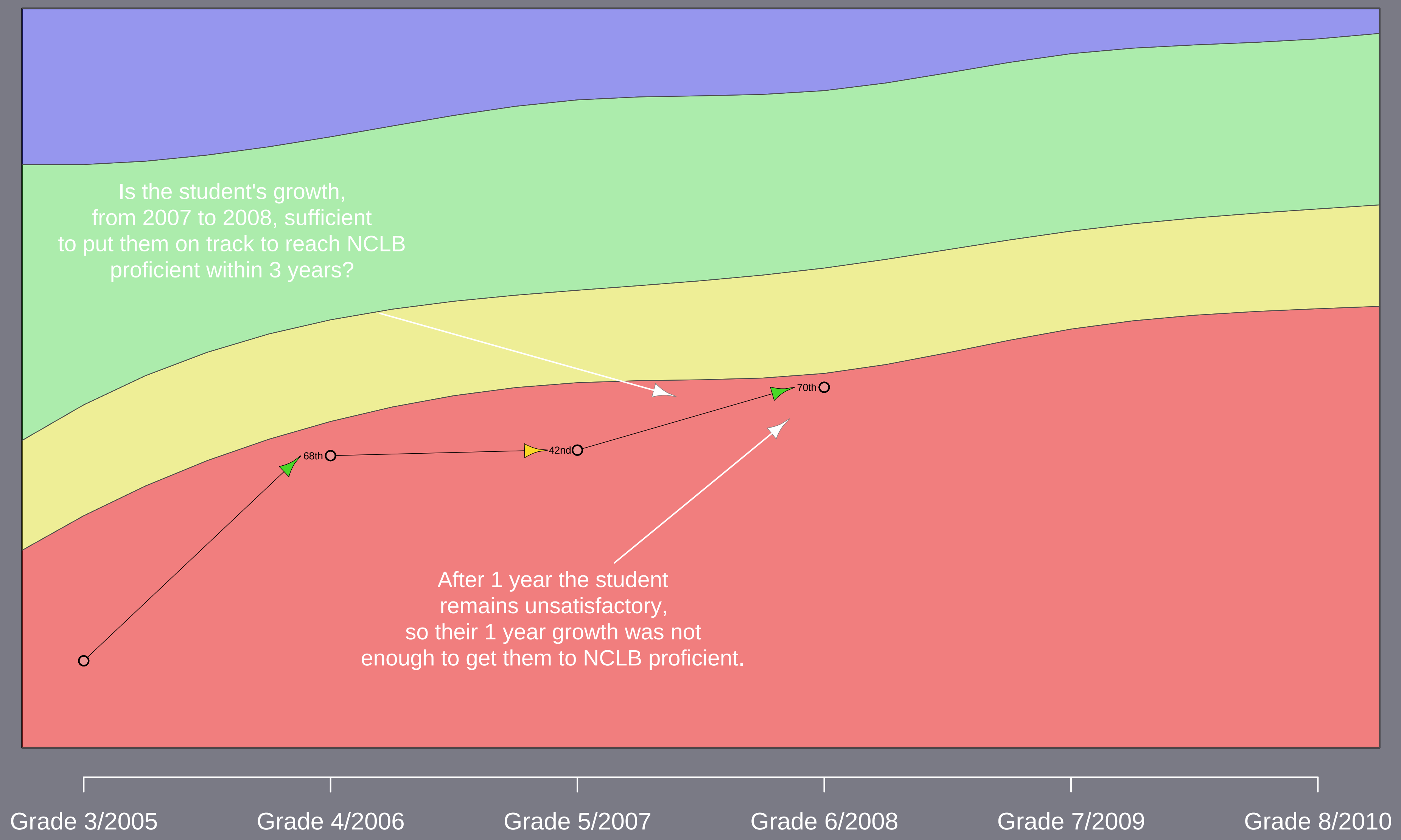
The Colorado Growth Model uses each student's growth percentile in two ways:
First, the growth percentile is used to describe how much a student has grown
during the last year. Second, the growth percentile is used to determine whether the
student is on track to reach/maintain proficiency. The following slides demonstrate,
for individual students, how the Colorado Growth Model is used to determine
whether the student is On Track to either Reach or Maintain Proficiency.
That is, whether the student is either "Catching Up" or "Keeping Up".

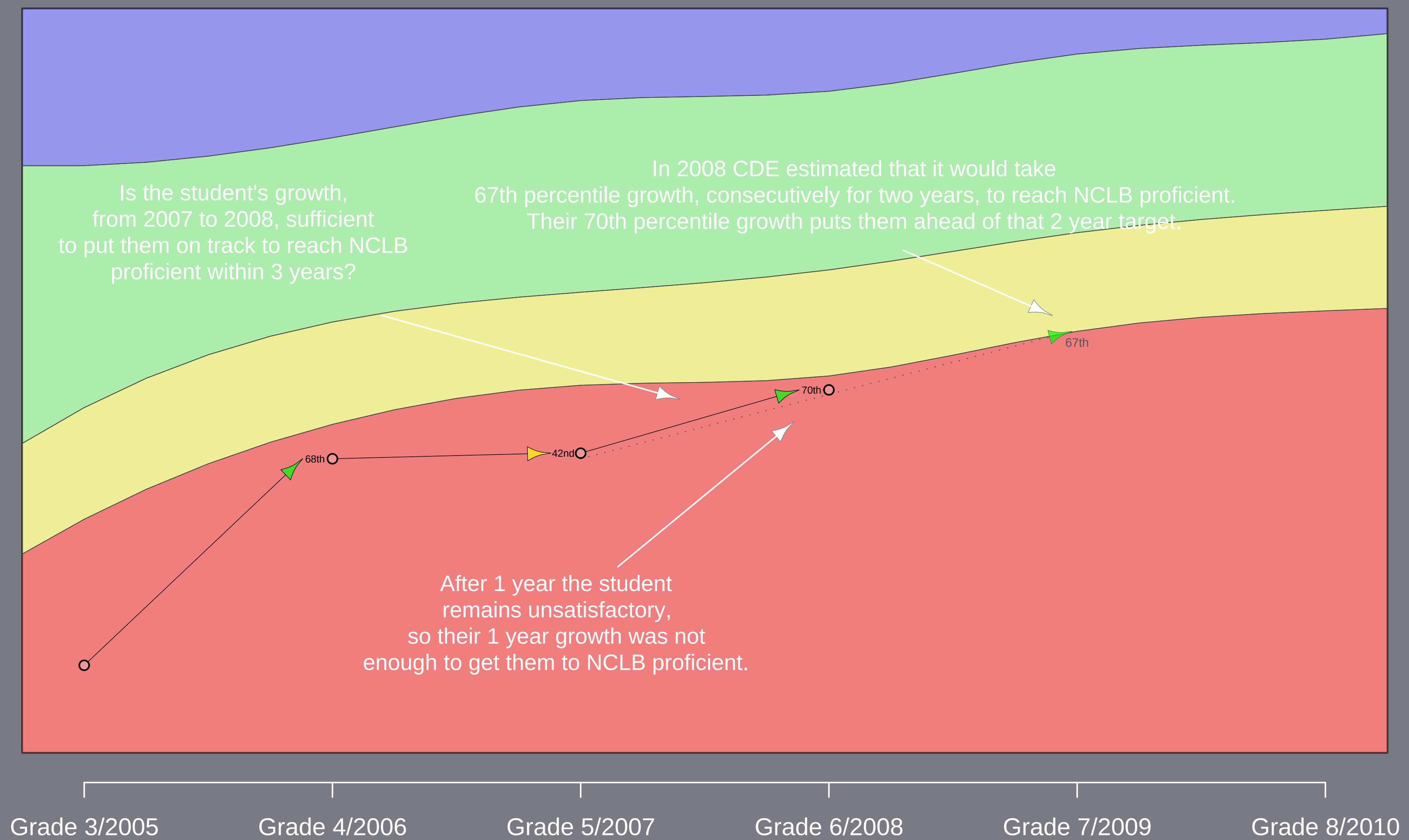
Grade 4/2005 Grade 5/2006 Grade 6/2007 Grade 7/2008 Grade 8/2009 Grade 9/2010

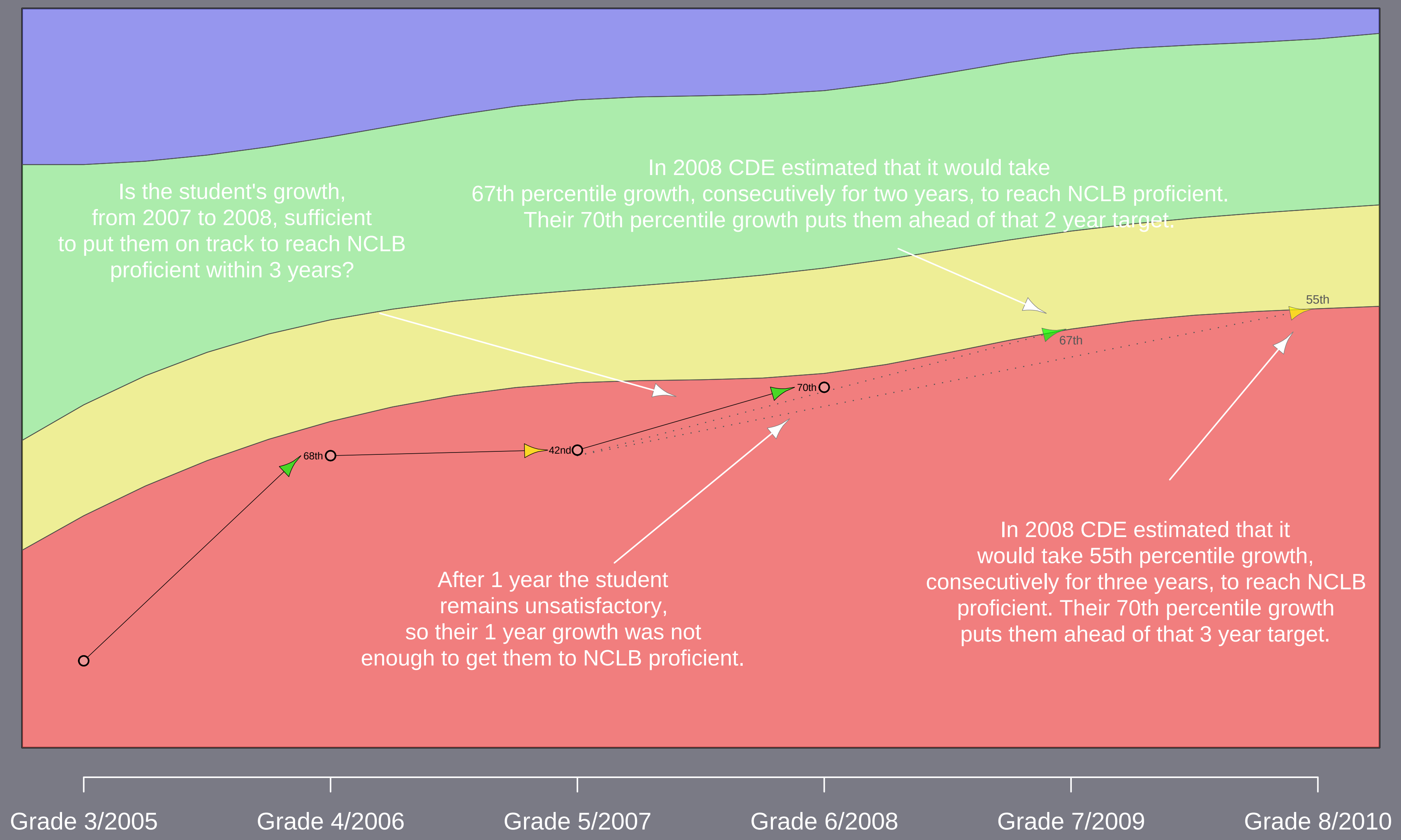

On Track to Reach NCLB Proficient − Catching Up
68th
42nd
70th
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010



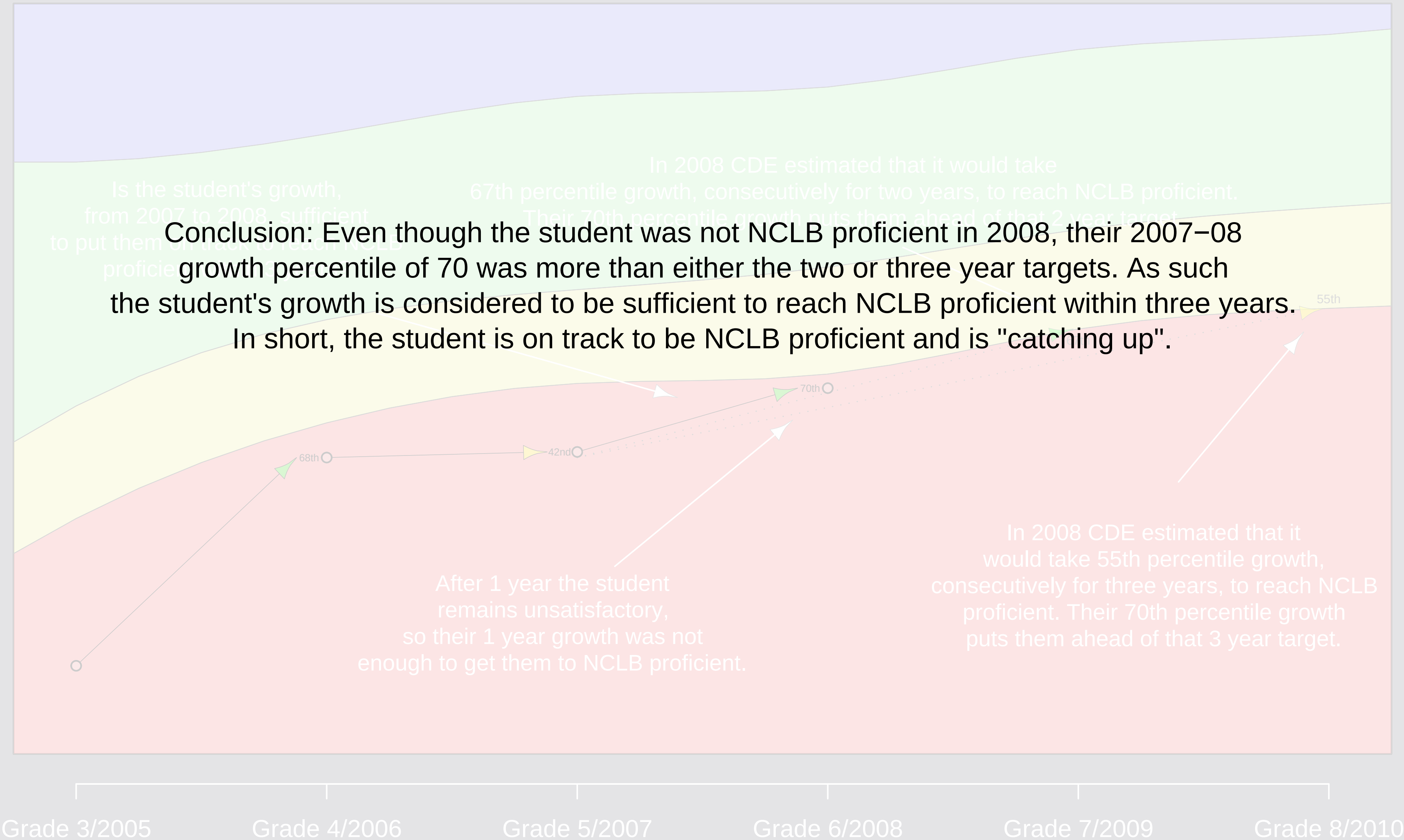
Is the student's growth,
from 2007 to 2008, sufficient
to put them on track to reach NCLB
proficient within 3 years?
68th
42nd
70th
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010


Is the student's growth,
from 2007 to 2008, sufficient
to put them on track to reach NCLB
proficient within 3 years?
68th
42nd
70th
After 1 year the student
remains unsatisfactory,
so their 1 year growth was not
enough to get them to NCLB proficient.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

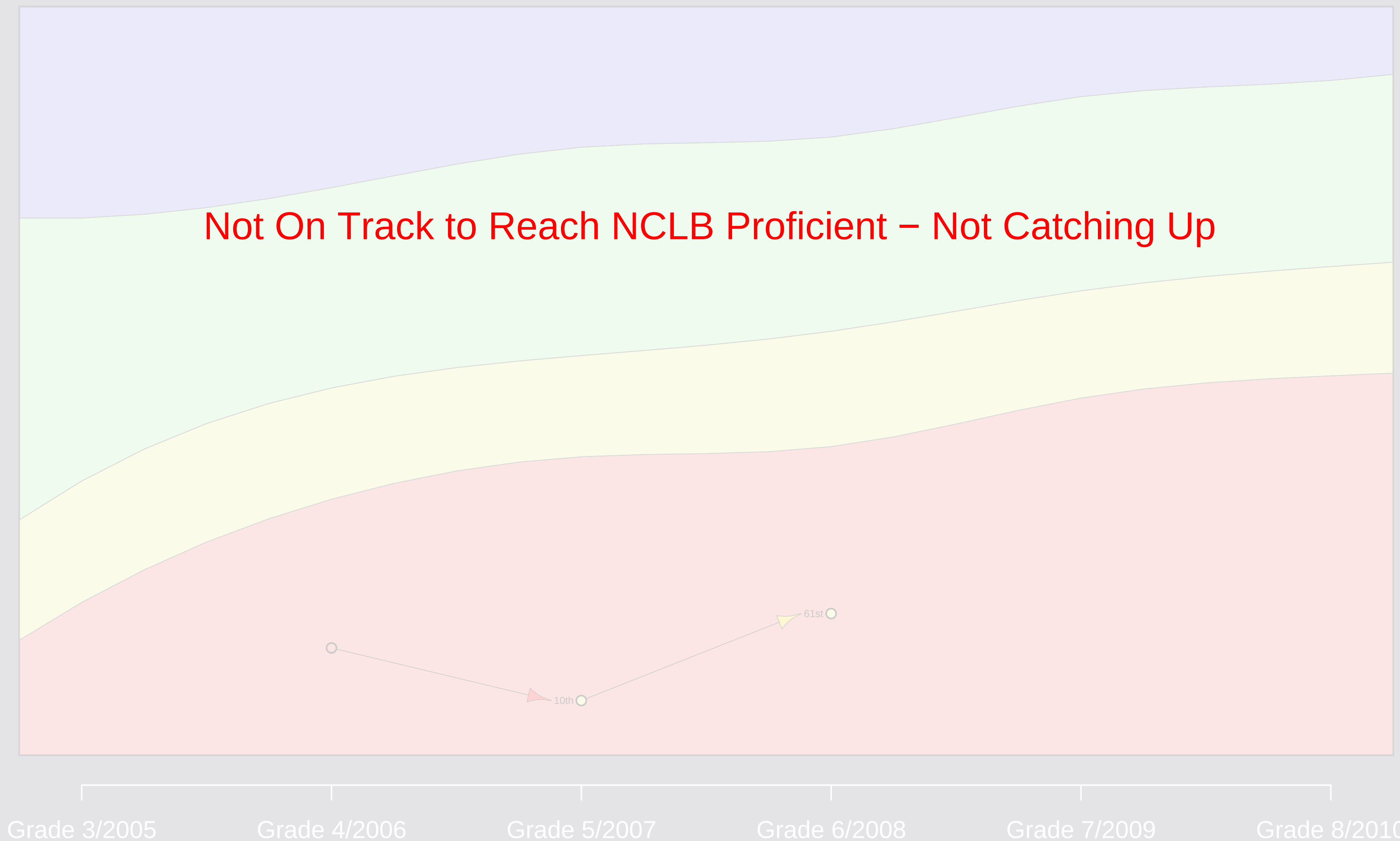

In 2008 CDE estimated that it would take
67th percentile growth, consecutively for two years, to reach NCLB proficient.
Their 70th percentile growth puts them ahead of that 2 year target.
Is the student's growth,
from 2007 to 2008, sufficient
to put them on track to reach NCLB
proficient within 3 years?
67th
70th
42nd
68th
After 1 year the student
remains unsatisfactory,
so their 1 year growth was not
enough to get them to NCLB proficient.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

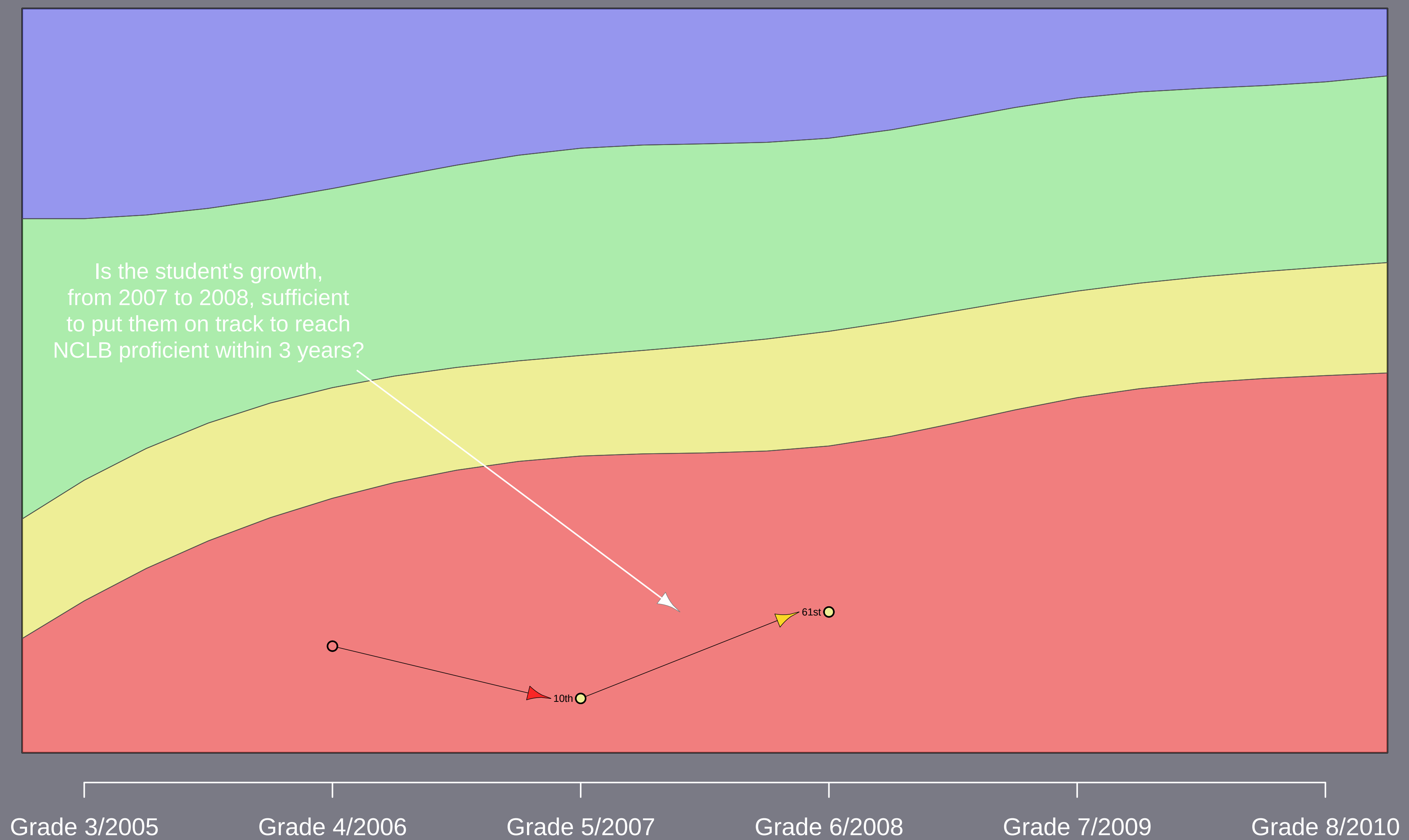


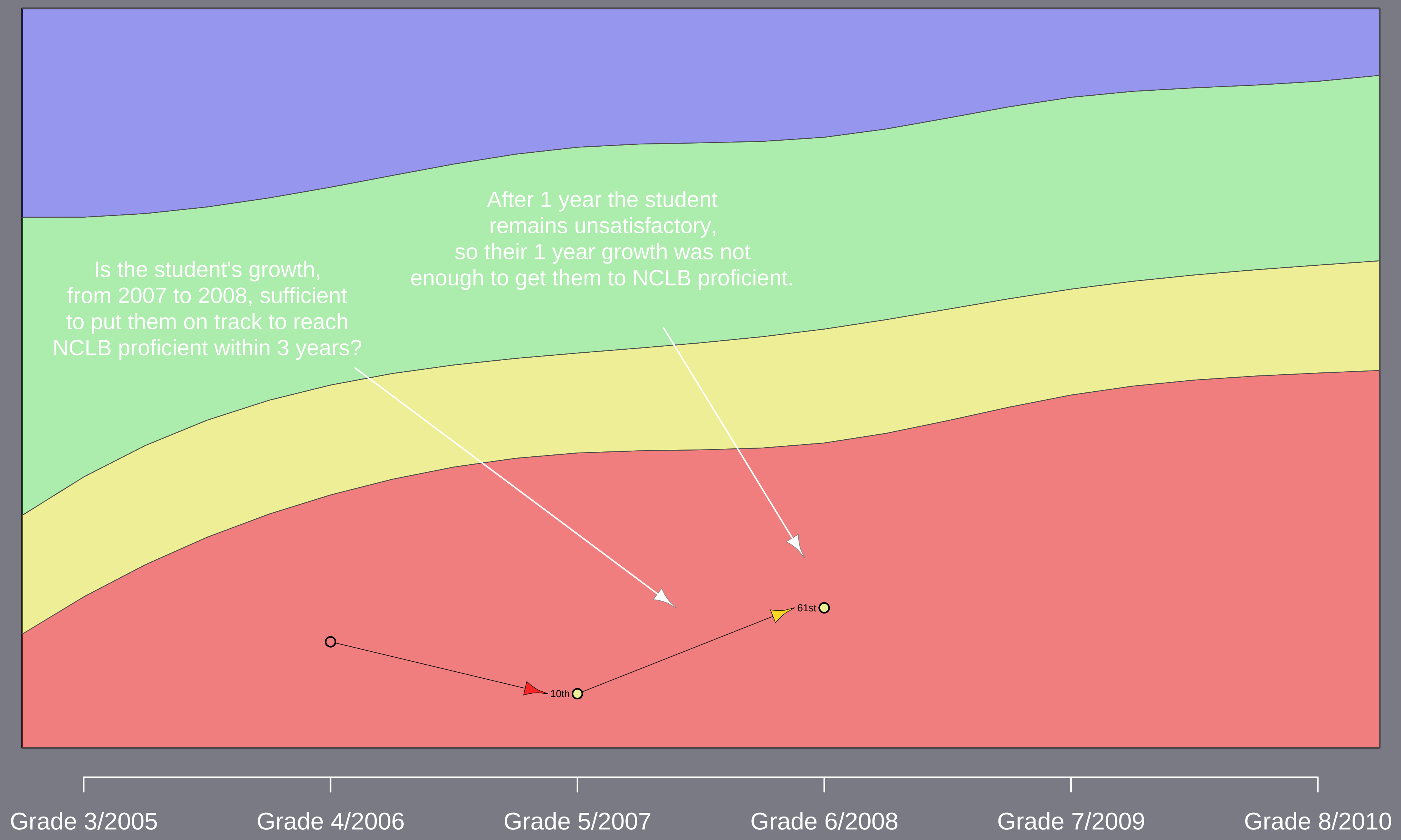
Is the student's growth, from 2007 to 2008, sufficient to put them on track to reach NCLB proficient within 3 years?
In 2008 CDE estimated that it would take 67th percentile growth, consecutively for two years, to reach NCLB proficient. Their 70th percentile growth puts them ahead of that 2 year target.
55th
67th
68th
42nd
70th
After 1 year the student remains unsatisfactory, so their 1 year growth was not enough to get them to NCLB proficient.
In 2008 CDE estimated that it would take 55th percentile growth, consecutively for three years, to reach NCLB proficient. Their 70th percentile growth puts them ahead of that 3 year target.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

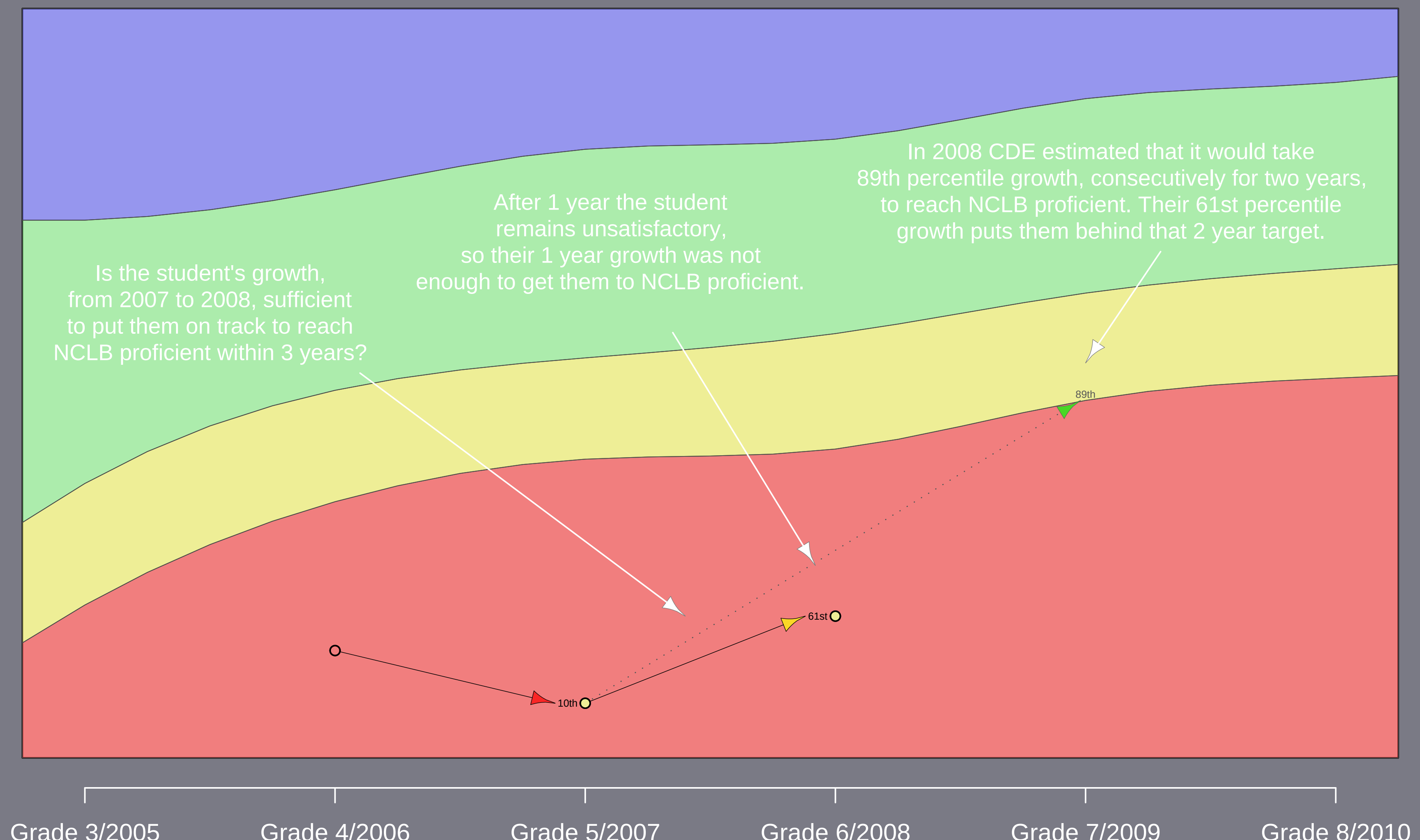


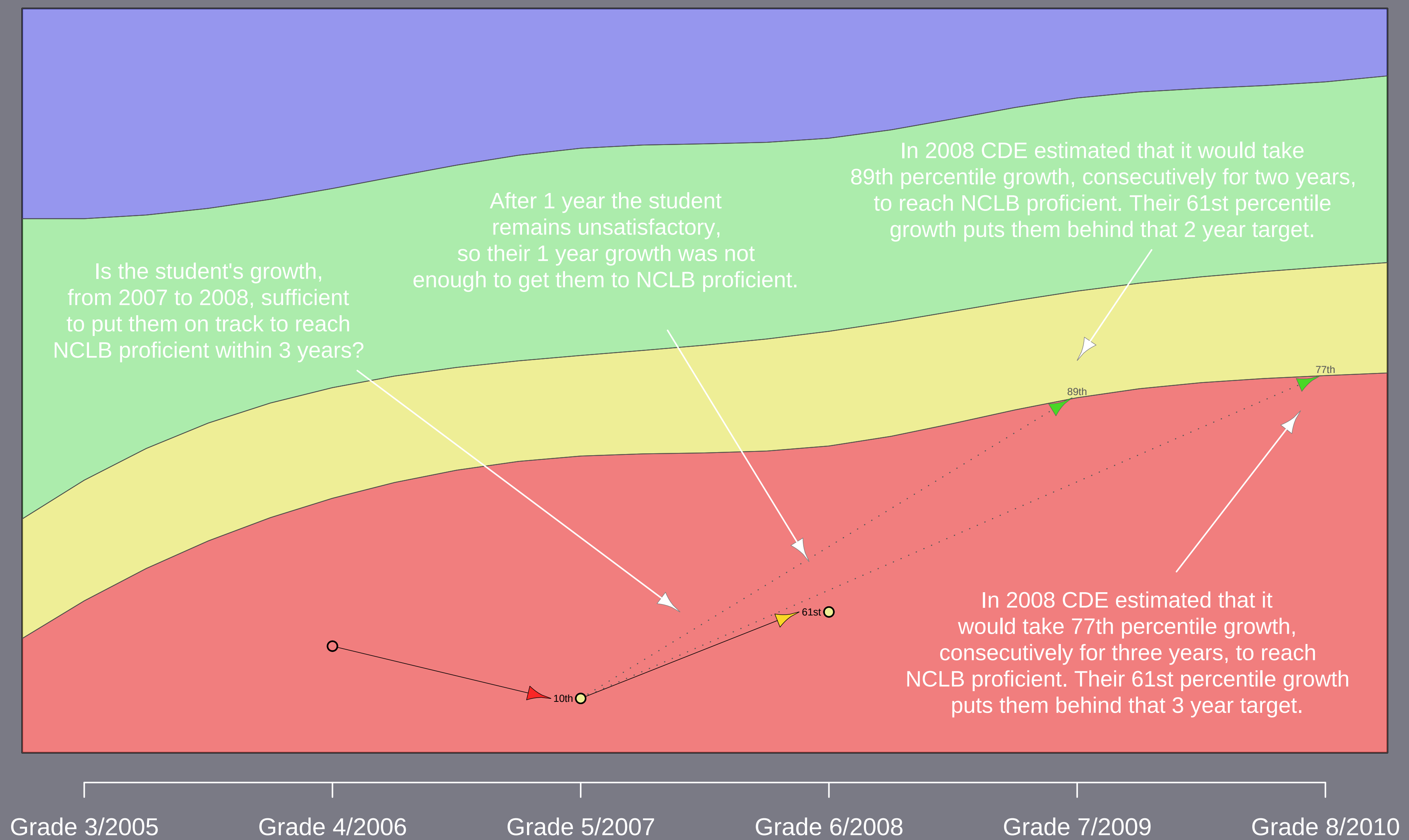
Conclusion: Even though the student was not NCLB proficient in 2008, their 2007−08
growth percentile of 70 was more than either the two or three year targets. As such
the student's growth is considered to be sufficient to reach NCLB proficient within three years.
In short, the student is on track to be NCLB proficient and is "catching up".
In 2008 CDE estimated that it would take
67th percentile growth, consecutively for two years, to reach NCLB proficient.
Is the student's growth,
from 2007 to 2008, sufficient
After 1 year the student
remains unsatisfactory,
so their 1 year growth was not
enough to get them to NCLB proficient.
In 2008 CDE estimated that it
would take 55th percentile growth,
consecutively for three years, to reach NCLB
proficient. Their 70th percentile growth
puts them ahead of that 3 year target.
68th
42nd
70th
55th
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010


Not On Track to Reach NCLB Proficient – Not Catching Up
10th
61st
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

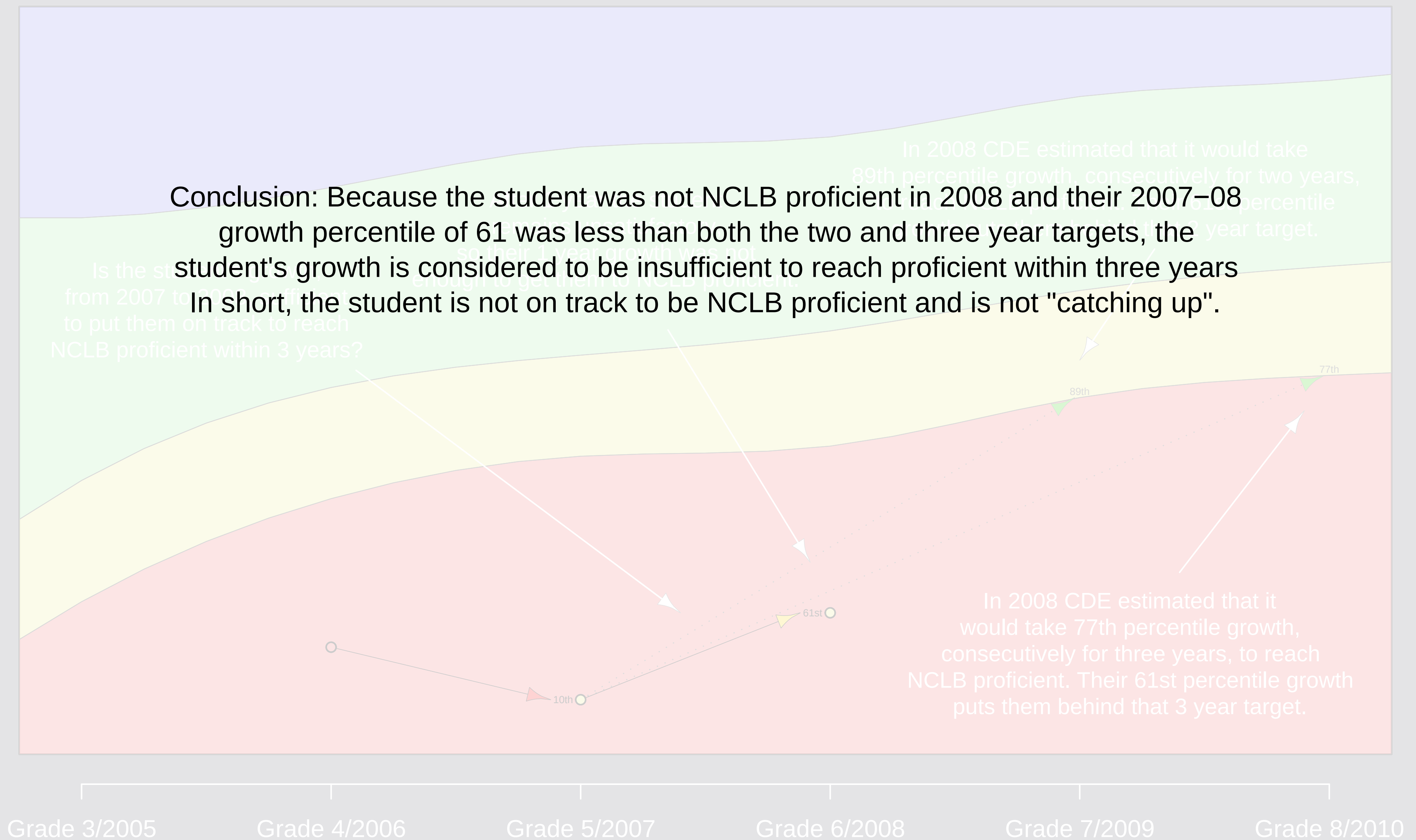

Is the student's growth,
from 2007 to 2008, sufficient
to put them on track to reach
NCLB proficient within 3 years?
10th
61st
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

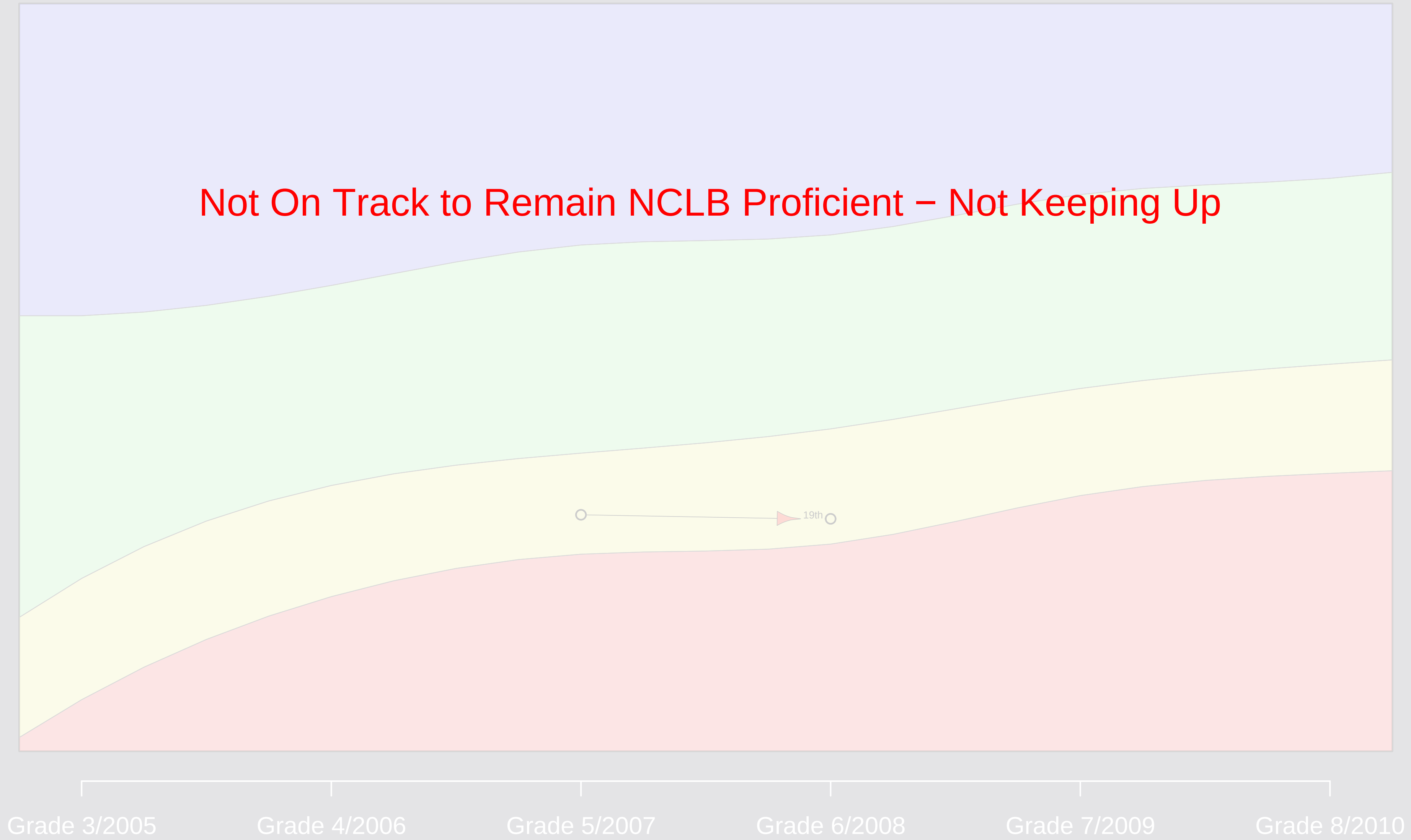

After 1 year the student
remains unsatisfactory,
so their 1 year growth was not
enough to get them to NCLB proficient.
Is the student's growth,
from 2007 to 2008, sufficient
to put them on track to reach
NCLB proficient within 3 years?
10th
61st
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

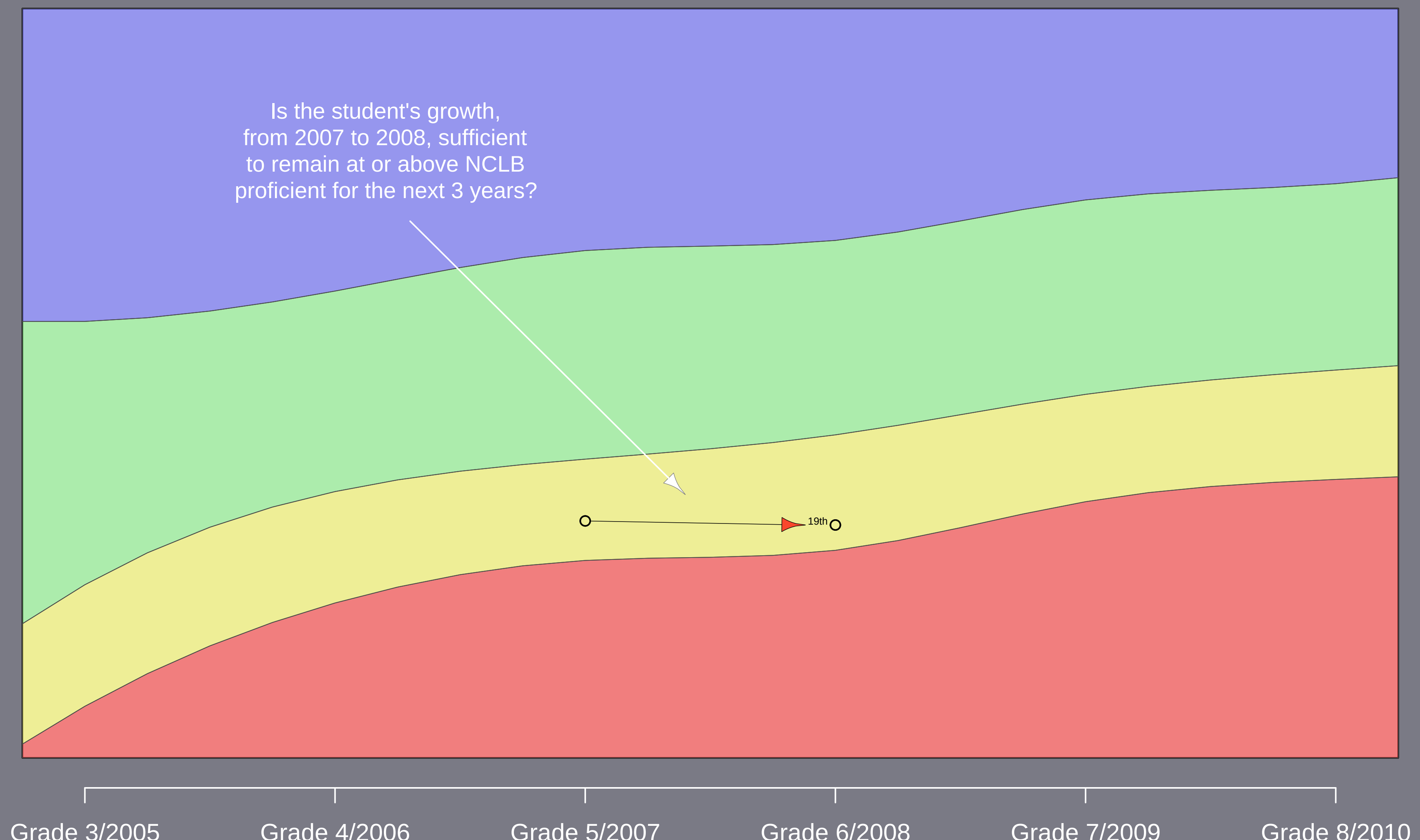

Is the student's growth, from 2007 to 2008, sufficient to put them on track to reach NCLB proficient within 3 years?
After 1 year the student remains unsatisfactory, so their 1 year growth was not enough to get them to NCLB proficient.
In 2008 CDE estimated that it would take 89th percentile growth, consecutively for two years, to reach NCLB proficient. Their 61st percentile growth puts them behind that 2 year target.
89th
61st
10th
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

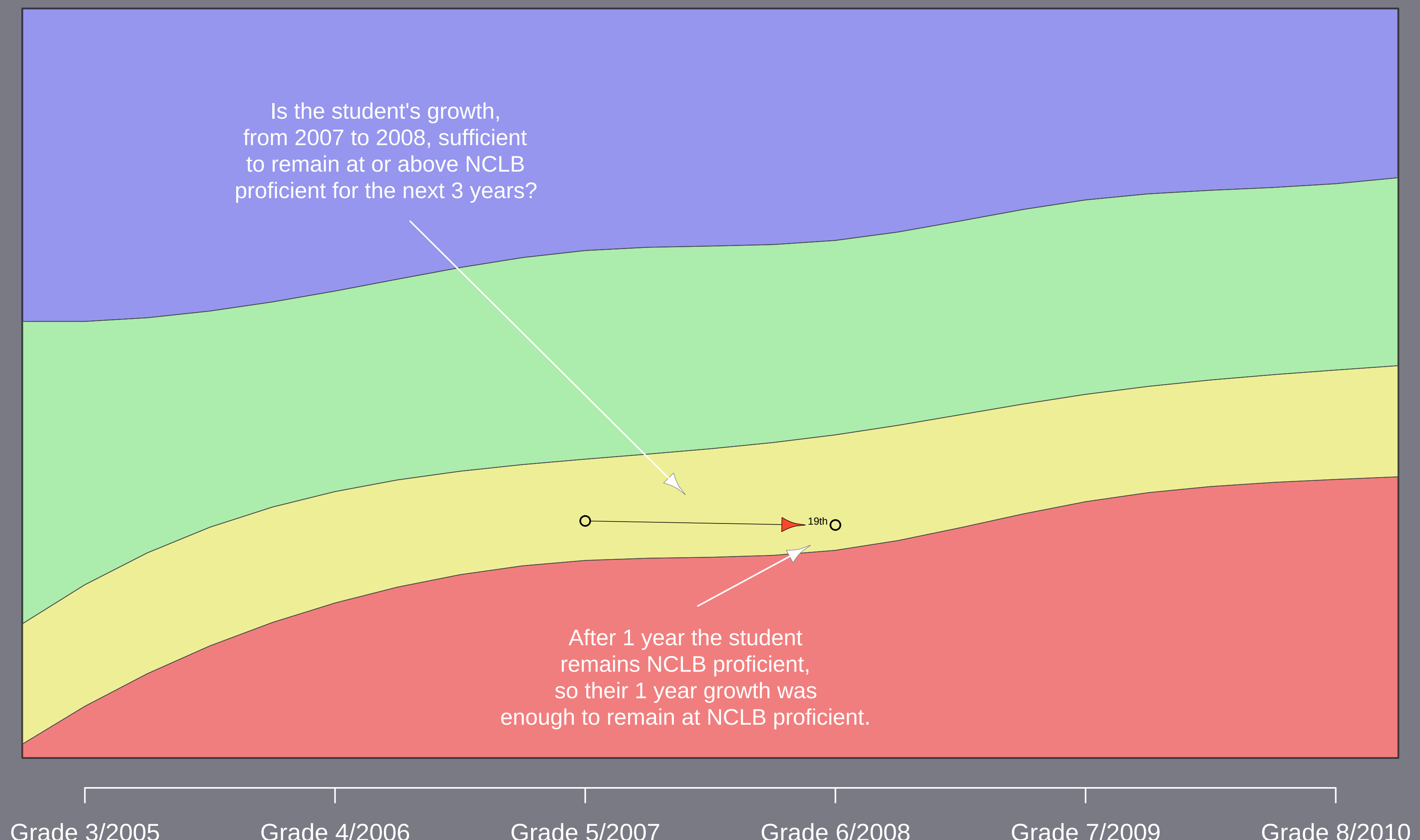


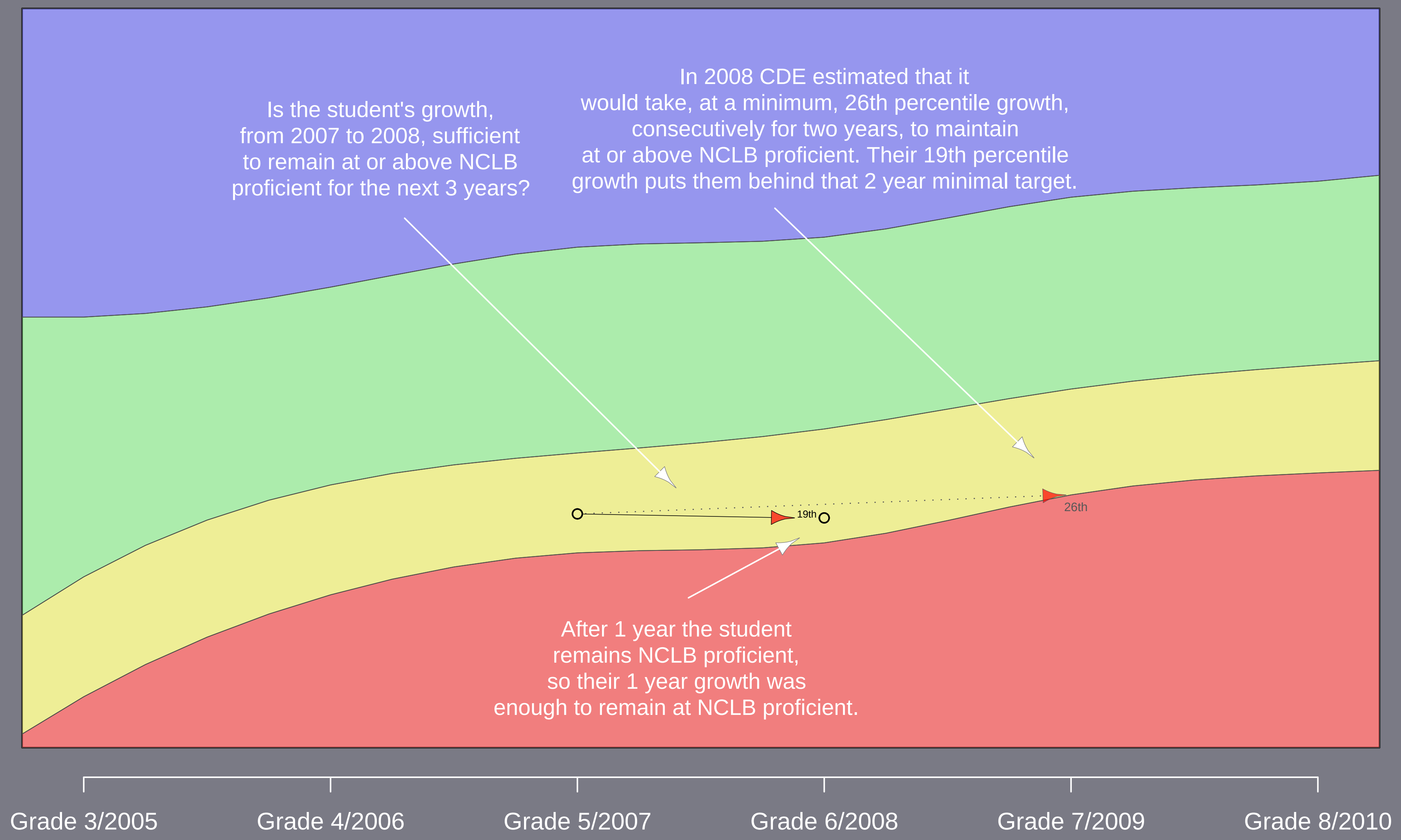
In 2008 CDE estimated that it would take 89th percentile growth, consecutively for two years, to reach NCLB proficient. Their 61st percentile growth puts them behind that 2 year target.
After 1 year the student remains unsatisfactory, so their 1 year growth was not enough to get them to NCLB proficient.
Is the student's growth, from 2007 to 2008, sufficient to put them on track to reach NCLB proficient within 3 years?
77th
89th
61st
10th
In 2008 CDE estimated that it would take 77th percentile growth, consecutively for three years, to reach NCLB proficient. Their 61st percentile growth puts them behind that 3 year target.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010


Conclusion: Because the student was not NCLB proficient in 2008 and their 2007−08 growth percentile of 61 was less than both the two and three year targets, the student's growth is considered to be insufficient to reach proficient within three years
In short, the student is not on track to be NCLB proficient and is not "catching up".
In 2008 CDE estimated that it would take 77th percentile growth, consecutively for three years, to reach NCLB proficient. Their 61st percentile growth puts them behind that 3 year target.
10th
61st
89th
77th
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

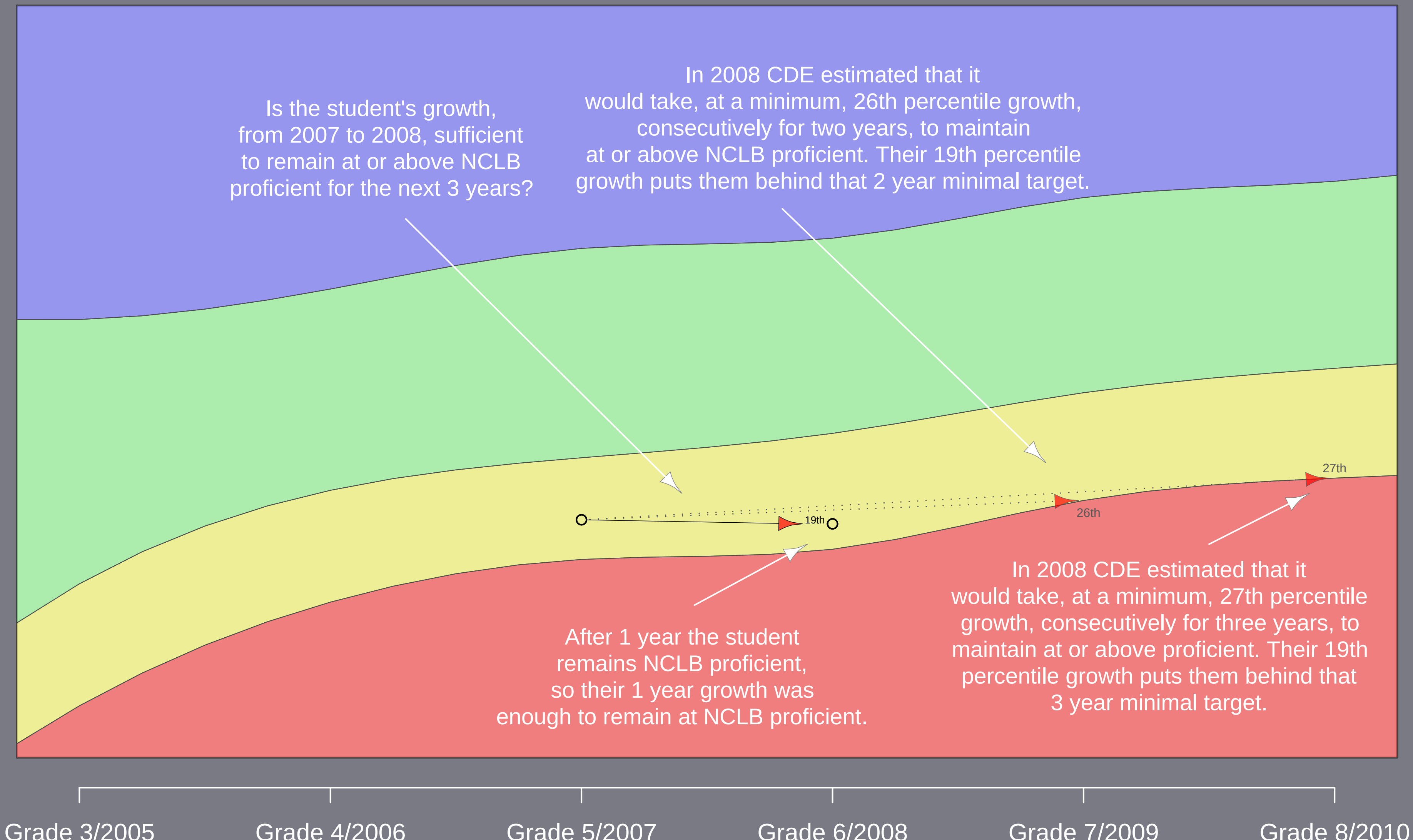

Not On Track to Remain NCLB Proficient − Not Keeping Up
19th
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

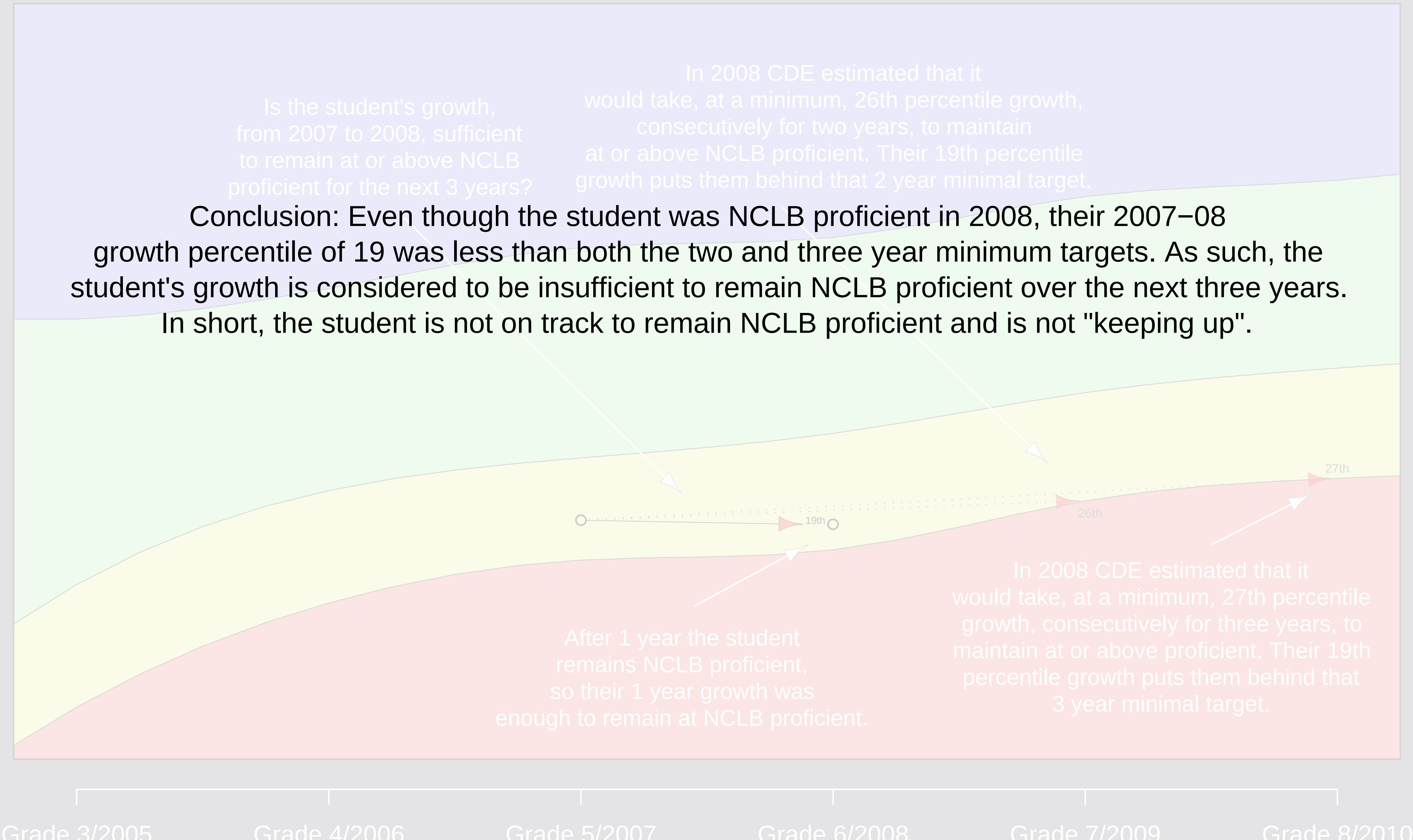

Is the student's growth,
from 2007 to 2008, sufficient
to remain at or above NCLB
proficient for the next 3 years?
19th
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010


Is the student's growth,
from 2007 to 2008, sufficient
to remain at or above NCLB
proficient for the next 3 years?
19th
After 1 year the student
remains NCLB proficient,
so their 1 year growth was
enough to remain at NCLB proficient.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

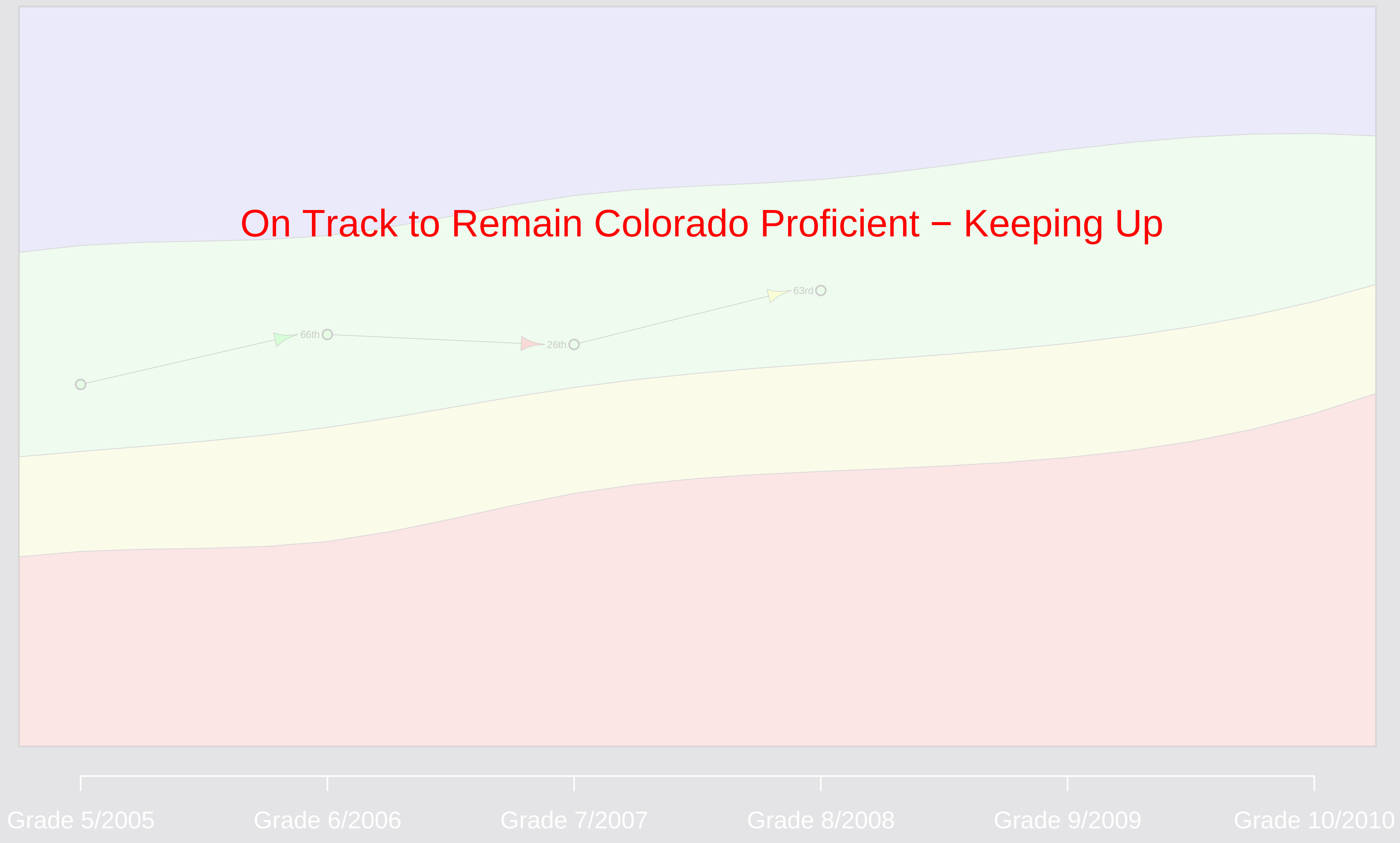


Is the student's growth,
from 2007 to 2008, sufficient
to remain at or above NCLB
proficient for the next 3 years?
In 2008 CDE estimated that it
would take, at a minimum, 26th percentile growth,
consecutively for two years, to maintain
at or above NCLB proficient. Their 19th percentile
growth puts them behind that 2 year minimal target.
19th
26th
After 1 year the student
remains NCLB proficient,
so their 1 year growth was
enough to remain at NCLB proficient.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

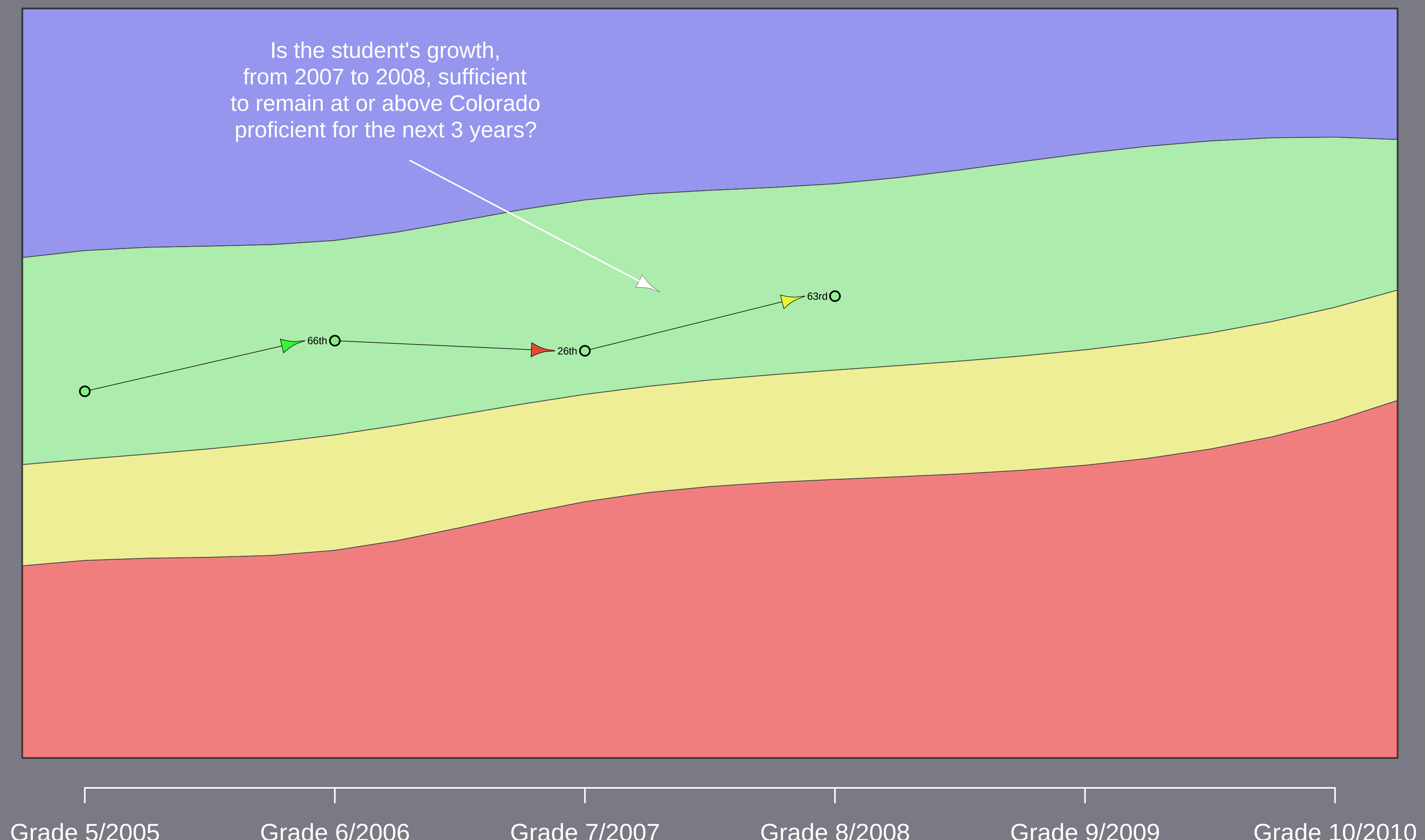

Is the student's growth, from 2007 to 2008, sufficient to remain at or above NCLB proficient for the next 3 years?
In 2008 CDE estimated that it would take, at a minimum, 26th percentile growth, consecutively for two years, to maintain at or above NCLB proficient. Their 19th percentile growth puts them behind that 2 year minimal target.
27th
19th
26th
After 1 year the student remains NCLB proficient, so their 1 year growth was enough to remain at NCLB proficient.
In 2008 CDE estimated that it would take, at a minimum, 27th percentile growth, consecutively for three years, to maintain at or above proficient. Their 19th percentile growth puts them behind that 3 year minimal target.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010



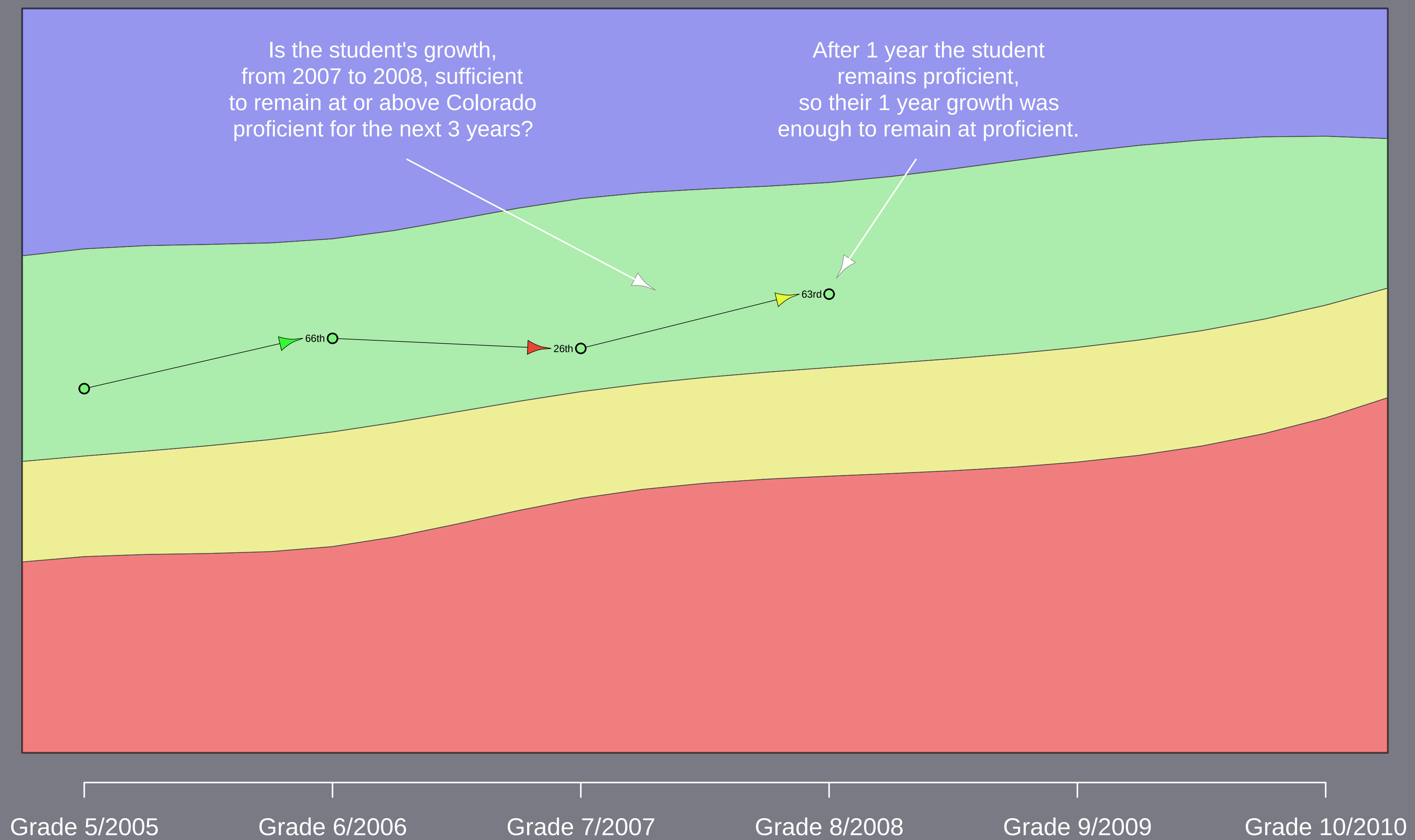
Conclusion: Even though the student was NCLB proficient in 2008, their 2007−08 growth percentile of 19 was less than both the two and three year minimum targets. As such, the student's growth is considered to be insufficient to remain NCLB proficient over the next three years. In short, the student is not on track to remain NCLB proficient and is not "keeping up".

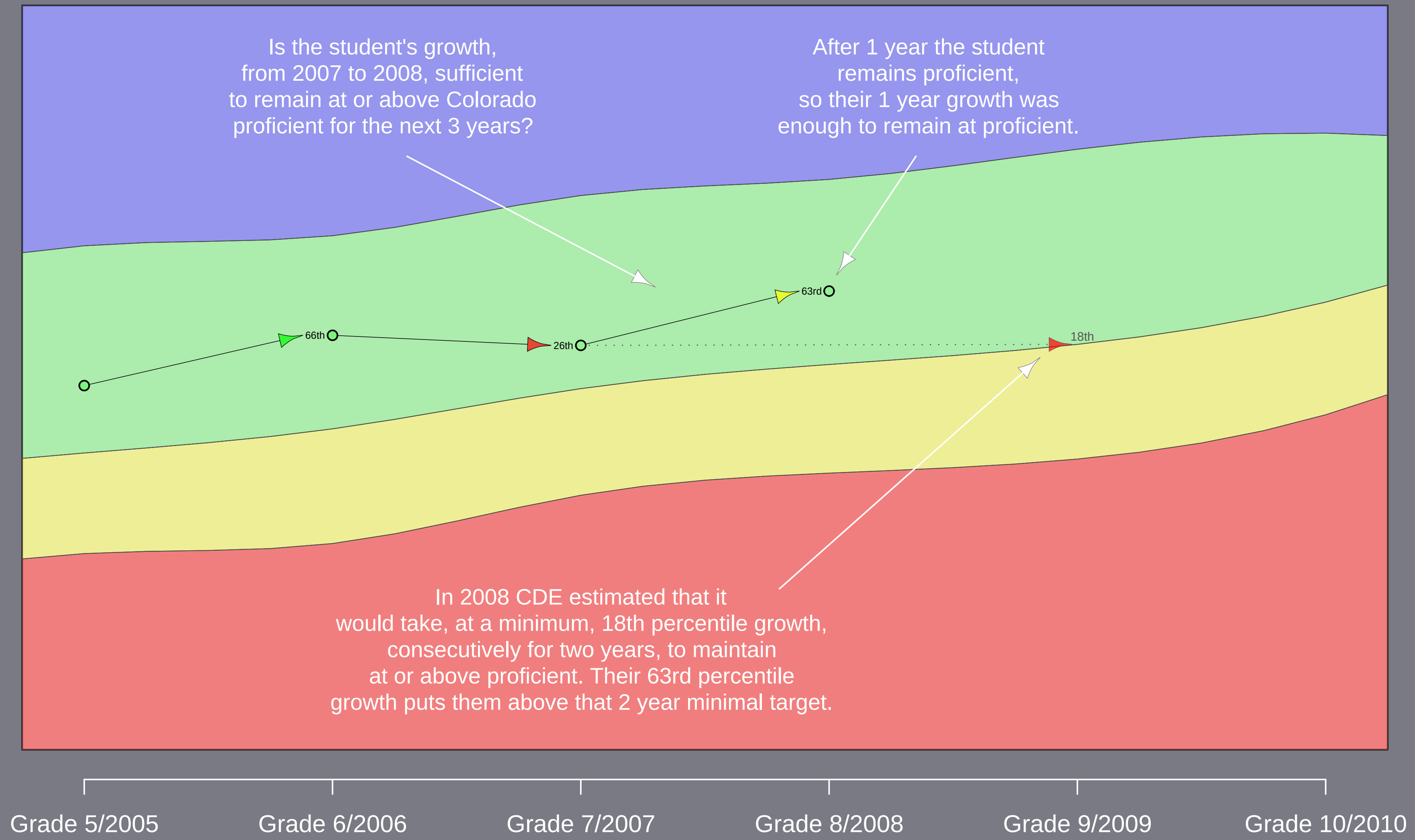


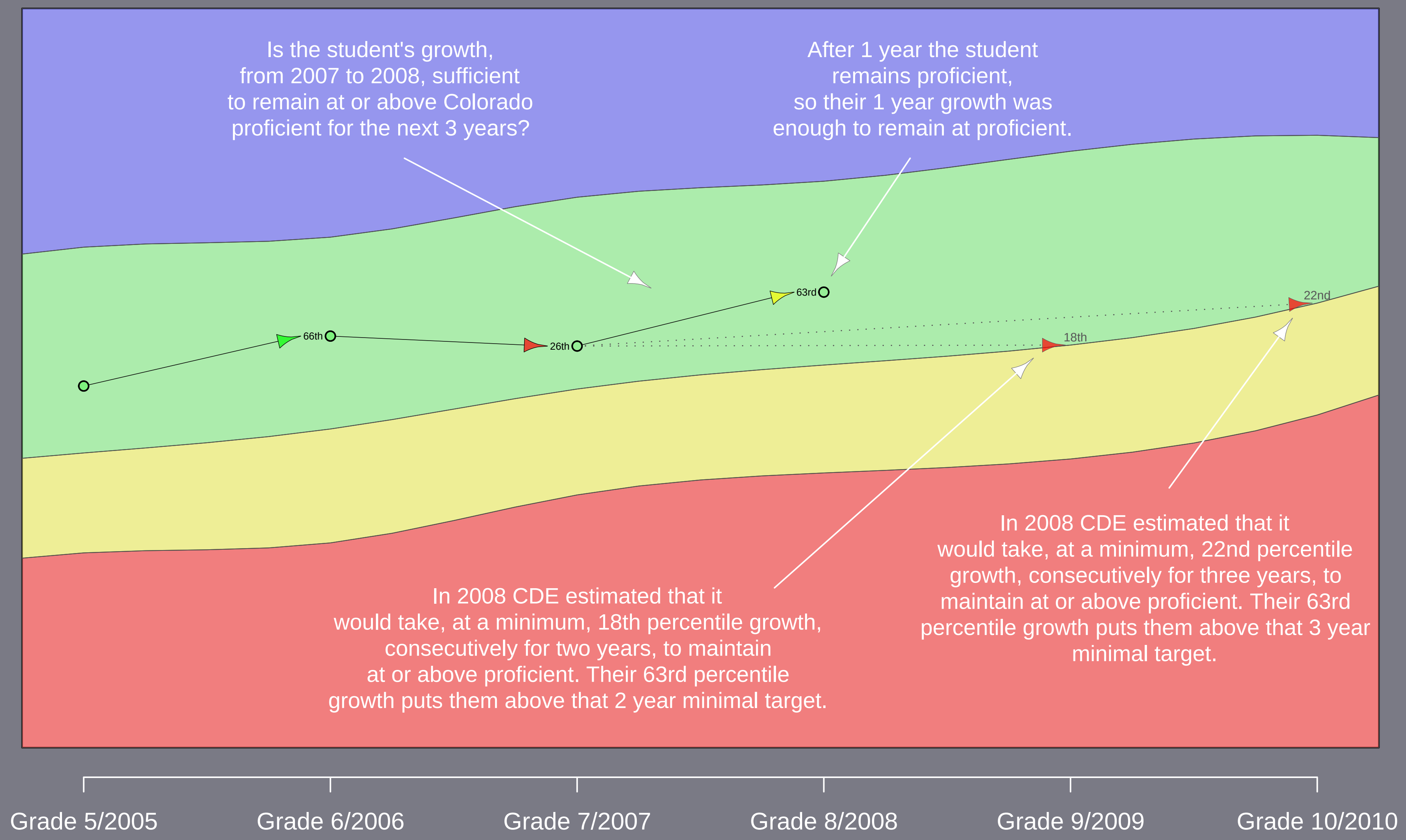
On Track to Remain Colorado Proficient − Keeping Up
66th
26th
63rd
Grade 5/2005
Grade 6/2006
Grade 7/2007
Grade 8/2008
Grade 9/2009
Grade 10/2010



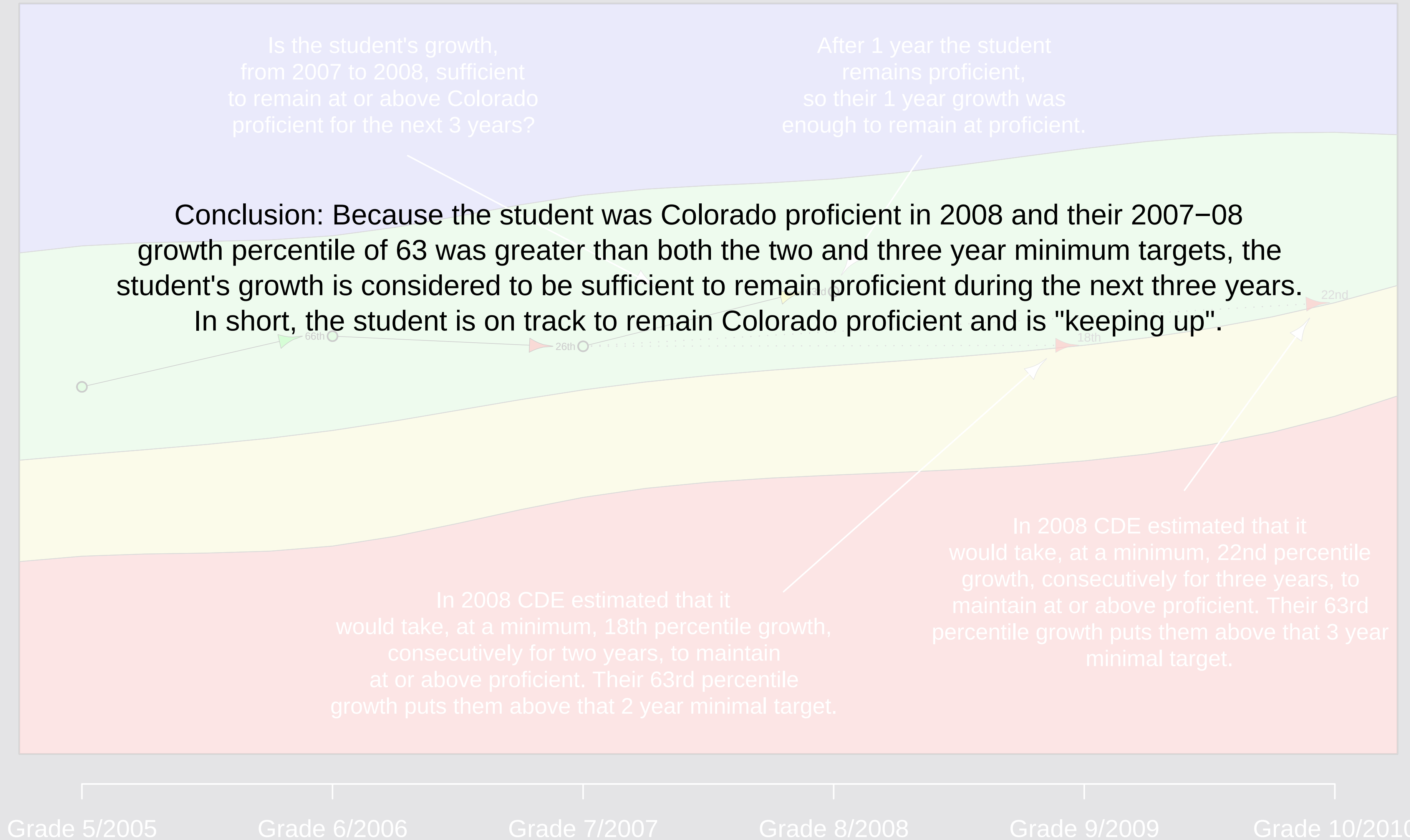
Is the student's growth,
from 2007 to 2008, sufficient
to remain at or above Colorado
proficient for the next 3 years?
66th
26th
63rd
Grade 5/2005
Grade 6/2006
Grade 7/2007
Grade 8/2008
Grade 9/2009
Grade 10/2010

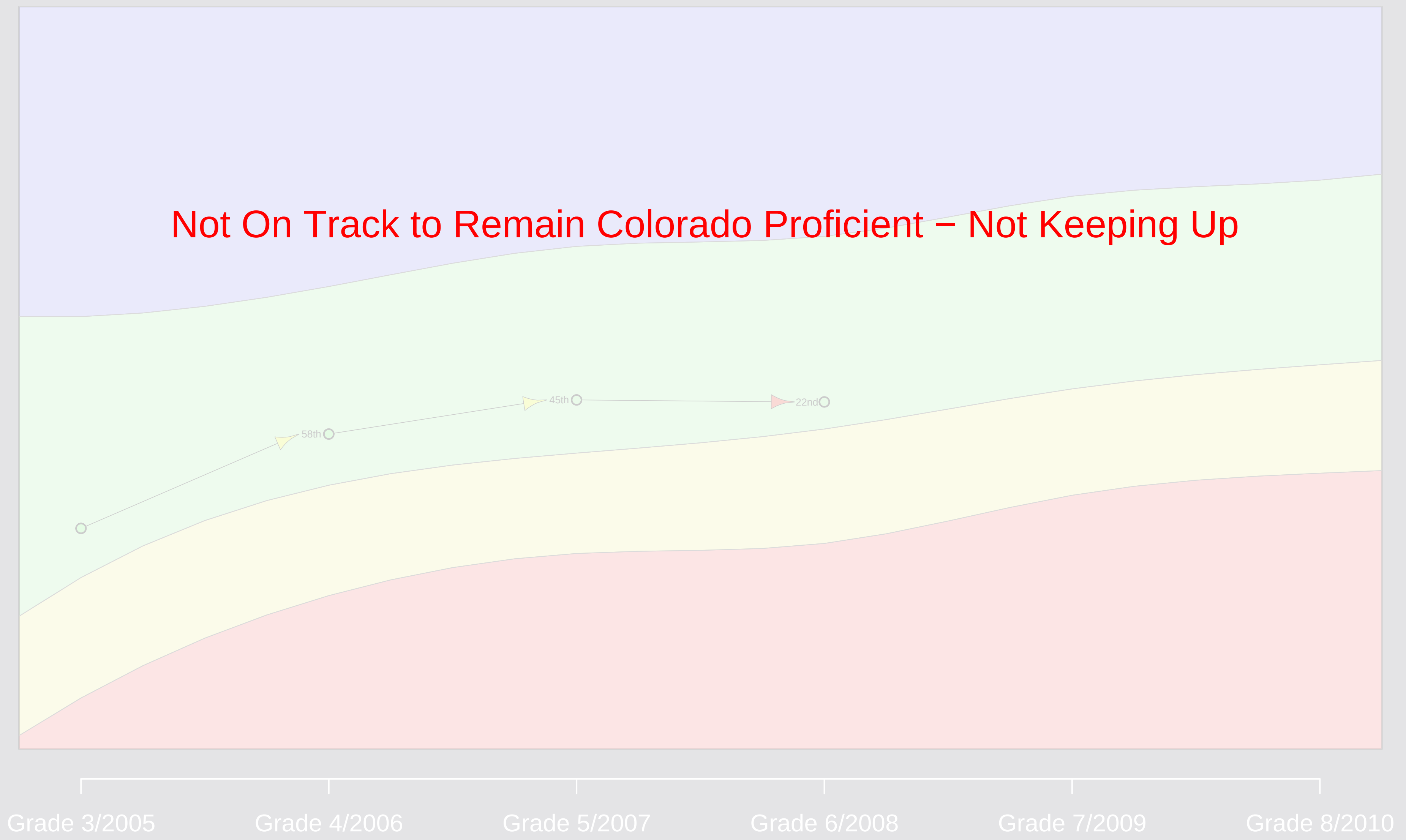

Is the student's growth,
from 2007 to 2008, sufficient
to remain at or above Colorado
proficient for the next 3 years?
After 1 year the student
remains proficient,
so their 1 year growth was
enough to remain at proficient.
66th
26th
63rd
Grade 5/2005
Grade 6/2006
Grade 7/2007
Grade 8/2008
Grade 9/2009
Grade 10/2010


Is the student's growth, from 2007 to 2008, sufficient to remain at or above Colorado proficient for the next 3 years?
After 1 year the student remains proficient, so their 1 year growth was enough to remain at proficient.
66th
26th
63rd
18th
In 2008 CDE estimated that it would take, at a minimum, 18th percentile growth, consecutively for two years, to maintain at or above proficient. Their 63rd percentile growth puts them above that 2 year minimal target.
Grade 5/2005
Grade 6/2006
Grade 7/2007
Grade 8/2008
Grade 9/2009
Grade 10/2010

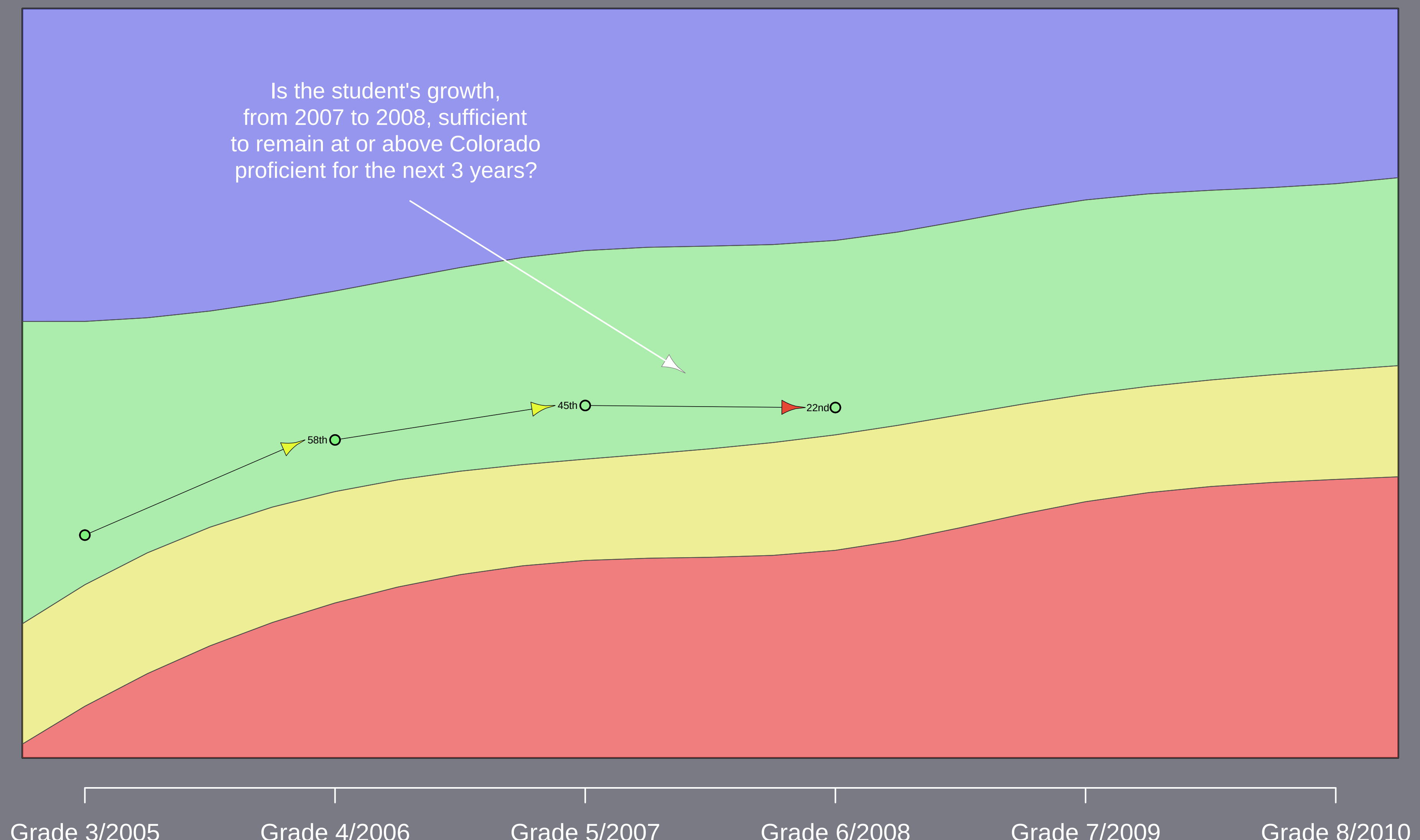


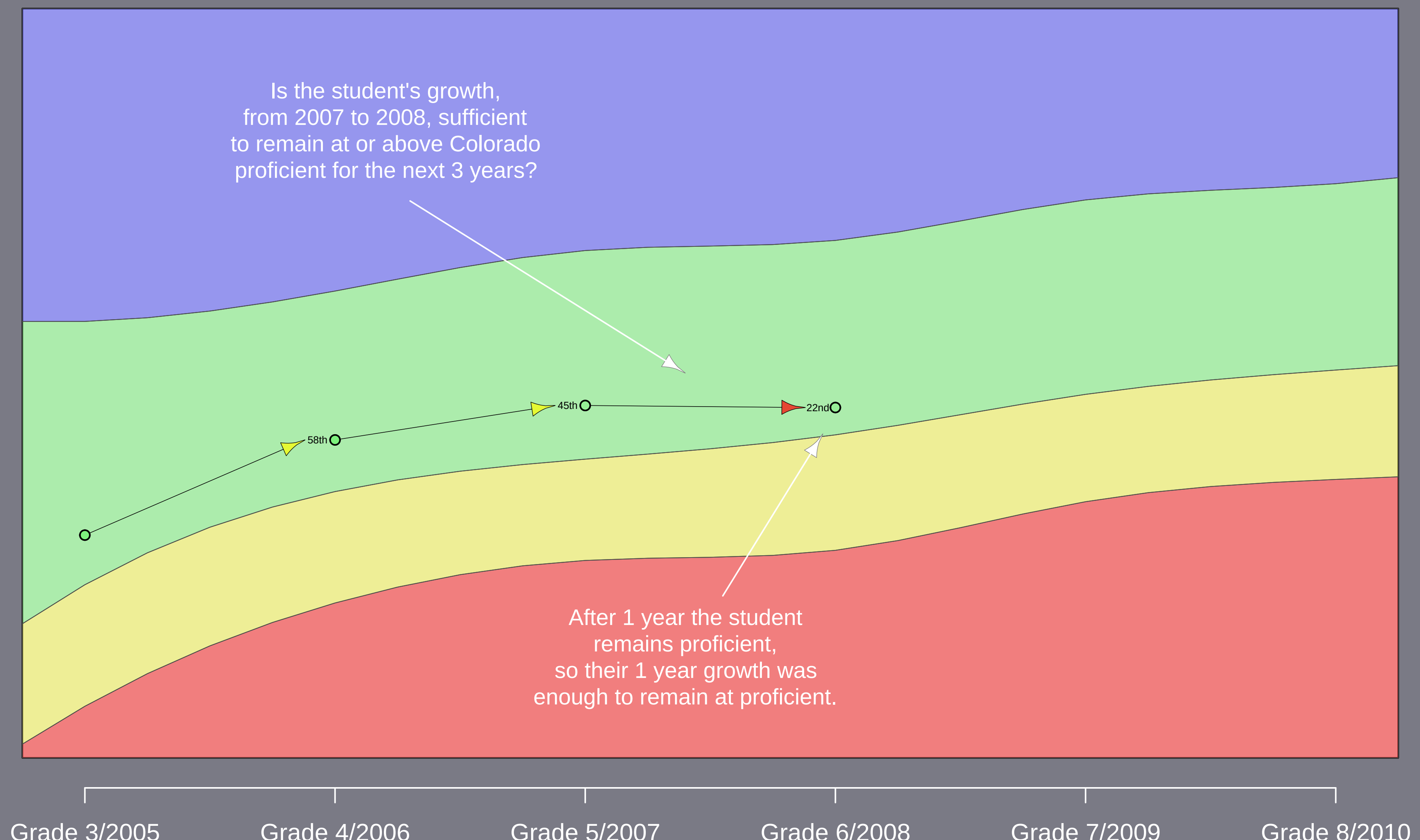
Is the student's growth, from 2007 to 2008, sufficient to remain at or above Colorado proficient for the next 3 years?
After 1 year the student remains proficient, so their 1 year growth was enough to remain at proficient.
66th
26th
63rd
18th
22nd
In 2008 CDE estimated that it would take, at a minimum, 18th percentile growth, consecutively for two years, to maintain at or above proficient. Their 63rd percentile growth puts them above that 2 year minimal target.
In 2008 CDE estimated that it would take, at a minimum, 22nd percentile growth, consecutively for three years, to maintain at or above proficient. Their 63rd percentile growth puts them above that 3 year minimal target.
Grade 5/2005
Grade 6/2006
Grade 7/2007
Grade 8/2008
Grade 9/2009
Grade 10/2010

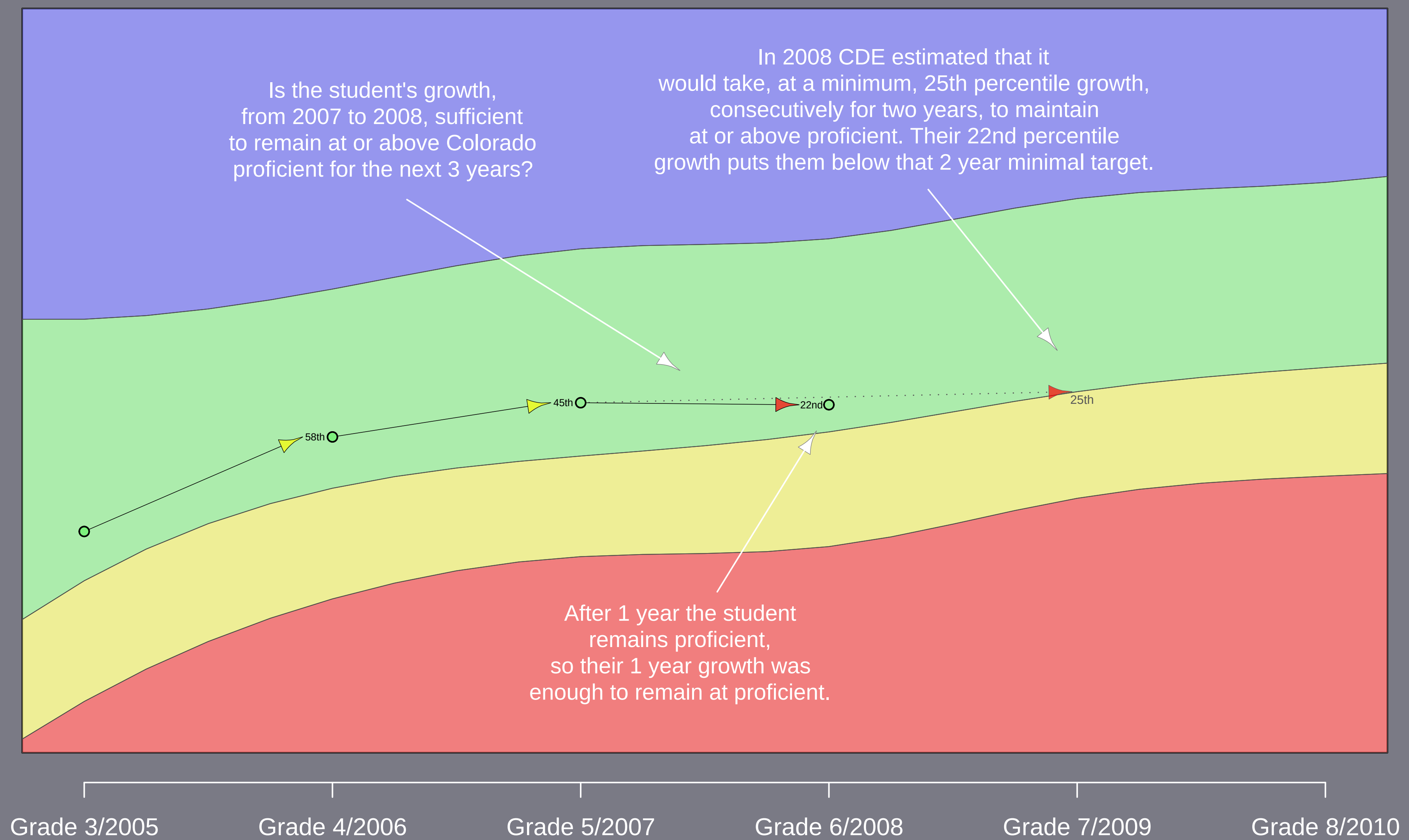

Is the student's growth, from 2007 to 2008, sufficient to remain at or above Colorado proficient for the next 3 years?
After 1 year the student remains proficient, so their 1 year growth was enough to remain at proficient.
Conclusion: Because the student was Colorado proficient in 2008 and their 2007−08 growth percentile of 63 was greater than both the two and three year minimum targets, the student's growth is considered to be sufficient to remain proficient during the next three years. In short, the student is on track to remain Colorado proficient and is "keeping up".
22nd
66th
26th
18th
In 2008 CDE estimated that it would take, at a minimum, 22nd percentile growth, consecutively for three years, to maintain at or above proficient. Their 63rd percentile growth puts them above that 3 year minimal target.
In 2008 CDE estimated that it would take, at a minimum, 18th percentile growth, consecutively for two years, to maintain at or above proficient. Their 63rd percentile growth puts them above that 2 year minimal target.
Grade 5/2005
Grade 6/2006
Grade 7/2007
Grade 8/2008
Grade 9/2009
Grade 10/2010

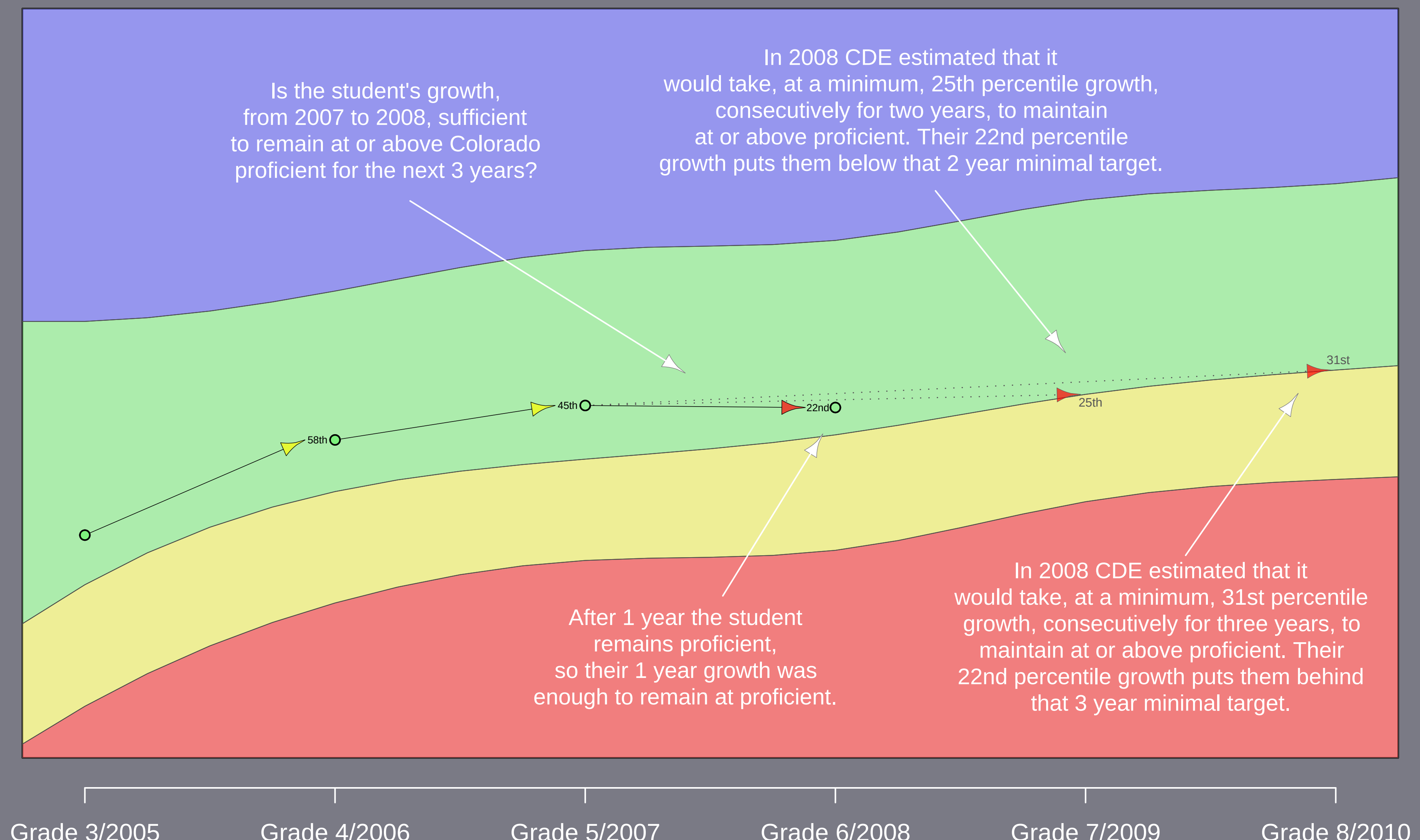


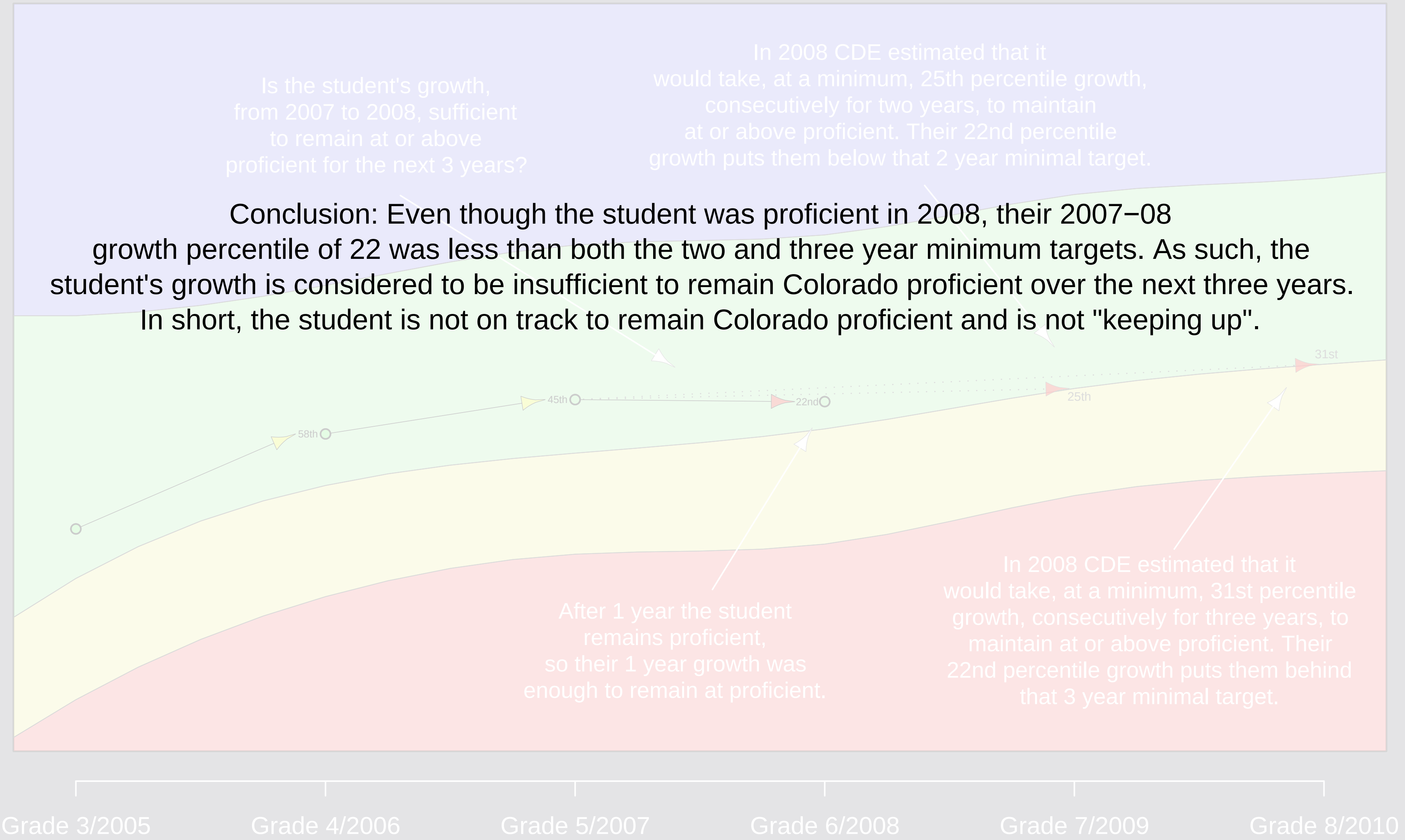
Not On Track to Remain Colorado Proficient − Not Keeping Up
58th
45th
22nd
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010


Is the student's growth,
from 2007 to 2008, sufficient
to remain at or above Colorado
proficient for the next 3 years?
58th
45th
22nd
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010


Is the student's growth,
from 2007 to 2008, sufficient
to remain at or above Colorado
proficient for the next 3 years?
58th
45th
22nd
After 1 year the student
remains proficient,
so their 1 year growth was
enough to remain at proficient.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010


Is the student's growth, from 2007 to 2008, sufficient to remain at or above Colorado proficient for the next 3 years?
In 2008 CDE estimated that it would take, at a minimum, 25th percentile growth, consecutively for two years, to maintain at or above proficient. Their 22nd percentile growth puts them below that 2 year minimal target.
58th
45th
22nd
25th
After 1 year the student remains proficient, so their 1 year growth was enough to remain at proficient.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010


Is the student's growth, from 2007 to 2008, sufficient to remain at or above Colorado proficient for the next 3 years?
In 2008 CDE estimated that it would take, at a minimum, 25th percentile growth, consecutively for two years, to maintain at or above proficient. Their 22nd percentile growth puts them below that 2 year minimal target.
31st
58th
45th
22nd
25th
After 1 year the student remains proficient, so their 1 year growth was enough to remain at proficient.
In 2008 CDE estimated that it would take, at a minimum, 31st percentile growth, consecutively for three years, to maintain at or above proficient. Their 22nd percentile growth puts them behind that 3 year minimal target.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010



Is the student's growth, from 2007 to 2008, sufficient to remain at or above proficient for the next 3 years?
In 2008 CDE estimated that it would take, at a minimum, 25th percentile growth, consecutively for two years, to maintain at or above proficient. Their 22nd percentile growth puts them below that 2 year minimal target.
Conclusion: Even though the student was proficient in 2008, their 2007−08 growth percentile of 22 was less than both the two and three year minimum targets. As such, the student's growth is considered to be insufficient to remain Colorado proficient over the next three years. In short, the student is not on track to remain Colorado proficient and is not "keeping up".
31st
58th
45th
22nd
25th
After 1 year the student remains proficient, so their 1 year growth was enough to remain at proficient.
In 2008 CDE estimated that it would take, at a minimum, 31st percentile growth, consecutively for three years, to maintain at or above proficient. Their 22nd percentile growth puts them behind that 3 year minimal target.
Grade 3/2005
Grade 4/2006
Grade 5/2007
Grade 6/2008
Grade 7/2009
Grade 8/2010

Student Growth Percentile and Percentile Growth Trajectory Calculation

Damian W. Betebenner
National Center for the Improvement of Educational Assessment
Dover, NH
DBetebenner@nciea.org

October 12, 2008

Background

It is a common misconception that to measure student growth in education, the subject matter and grades over which growth is examined must be on the same scale—referred to as a vertical scale. Not only is a vertical scale not necessary, but its existence obscures fundamental concepts necessary to understand growth. Growth, fundamentally, requires change to be examined for a single construct, like math achievement, over time—*growth in what?* A single scale for the construct is only necessary to measure the *magnitude* of growth, but not growth in general (Betebenner, 2008; Yen, 2007).

Consider the familiar situation from pediatrics where the interest is on measuring the height and weight of children over time. The scales on which height and weight are measured possess properties that educational assessment scales aspire towards but can never meet.

> An infant male toddler is measured at 2 and 3 years of age and is shown to have grown 4 inches. The magnitude of increase—4 inches—is a well understood quantity that any parent can grasp and calculate at home using a simple yardstick. However, parents leaving their pediatrician's office knowing only how much their child has grown would likely be wanting for more information: Parents are not interested in an absolute magnitude of growth, but instead in a normative criterion locating that 4 inch increase alongside the height increases of similar children. Examining this height increase relative to the increases of similar children permits a diagnosis of how (ab)normal such an increase is.

Given this reality in the examination of change where scales of measurement are perfect, it is absurd to think that in education, where scales are, at best, quasi-interval, one can/should examine growth differently.

If scales did exist in education similar to height/weight scales that permitted the calculation of absolute measures of annual academic growth for students, a parent asking, "How much did my child progress?", would be answered with a number of scale score points. Such an answer would likely leave most parents confused wondering whether the number of points is good or bad. As in pediatrics, the search for a description regarding changes in achievement over time (i.e., growth) is best served by considering a normative quantification of student growth—*a student growth percentile*.

A student's growth percentile describes how (ab)normal a student's growth is by examining their current achievement relative to their *academic peers*—those students with identical prior achievement. That is, a student growth percentile examines the current achievement of a student relative to other students who have, in the past, "walked the same achievement path". Heuristically, if the state assessment data set were extremely large (in fact, infinite) in size, one could examine the data set and select out those students with the exact same prior scores and compare how the selected student's current year score compares to the current year scores of those students with the same prior year's scores—their academic peers. If the student's current year score exceeded the scores of most of their academic peers, in a normative sense they have done well. If the student's current year score was less than the scores of their academic peers, in a normative sense they have not done well.

The four panels of Figure 1 depict what a student growth percentile represents in a situation considering students having only two consecutive achievement test scores.

Upper Left Panel Considering all pairs of scores for all students in the state yields a bivariate (two variable) distribution.

Upper Right Panel Taking account of prior achievement (i.e., conditioning upon prior achievement) fixes a the value of the 2005 scale score (in this case at 600) and is represented by the red slice taken out of the bivariate distribution.

Lower Left Panel Conditioning upon prior achievement defines a *conditional distribution* which represents the distribution of outcomes on the 2006 test assuming a 2005 score of 600. This distribution is indicating as the solid red curve.

Lower Right Panel The conditional distribution provides the context within which a student's 2006 achievement can be understood normatively. Students with achievement in the upper tail of the conditional distribution have demonstrated high rates of growth relative to their academic peers whereas those students with achievement in the lower tail of the distribution have demonstrated low rates of growth. Students with current achievement in the middle of the distribution could be described as demonstrating "average" or "typical" growth. In the figure provided the student scores approximately 650 on the 2006 test. Within the conditional distribution, the value of 650 lies at approximately the 70th percentile. Thus the student's growth from 600 in 2005 to 650 in 2006 met or exceeded that of approximately 70 percent of students starting from the same place. This 50 point increase is above average. It is important to note that qualifying a student growth percentile as "adequate", "good", or "enough" is a standard setting procedure requiring stakeholders to examine a student's growth relative to external criteria such as performance standards/levels.

Figure 1 also serves to illustrate the relationship between a vertical scale and student growth percentiles. Using the vertical scale implied by Figure 1, the student grew 50 points (from 600 to 650) between 2005 and 2006. This 50 points represents the absolute magnitude of change. Quantifying the magnitude of change is scale dependent. For example, different vertical achievement scales in 2005 and 2006 would yield different annual scale score increases: A scale score increase of 50 could be changed to a scale score increase of 10 using a simple transformation of the vertical scale on which all the students are measured. However, relative to other students, their growth has not changed—their growth percentile is invariant to scale transformations common in educational assessment. Student growth percentiles normatively situate achievement change bypassing questions associated with the magnitude of change, and directing attention toward relative standing which is what stakeholders are most interested in.

The percentile of a student's current score within their corresponding conditional distribution translates to a probability statement of a student obtaining that score taking account of prior achievement. That is:[1]

$$\textsf{Student Growth Percentile} \equiv \Pr(\mathit{Current\ Achievement} | \mathit{Past\ Achievement}) \cdot 100.$$

Whereas unconditional percentiles normatively quantify achievement, conditional percentiles normatively quantify growth. Because past scores are used solely for conditioning purposes, one of the major advantages of using growth percentiles to measure change is that estimation does not require a vertical scale.

In contrast to the majority of longitudinal analysis techniques using student assessment data that seek to *explain* the variability of student scores *vis-á-vis* teacher or school effects, the primary concern with calculating student growth percentiles is to *describe* this variability and, following similar descriptions used in pediatrics, give stakeholders a sense of what the current range of growth in student achievement is.[2] And with a common descriptive vocabulary in place, these descriptions form the basis of discussions of student growth regarding *what is?*, *what should be?*, and *what could be?*

[1]Technically, the expression denotes a student growth quantile since $\Pr(\mathit{Current\ Achievement}|\mathit{Past\ Achievement}) \cdot 100$ is not always an integer between 0 and 100. To simplify, the result is rounded and termed a percentile.

[2]See http://www.nutropin.com/patient/3_5_4_growth_velocity.jsp for an online implementation of pediatric growth percentiles associated with height.

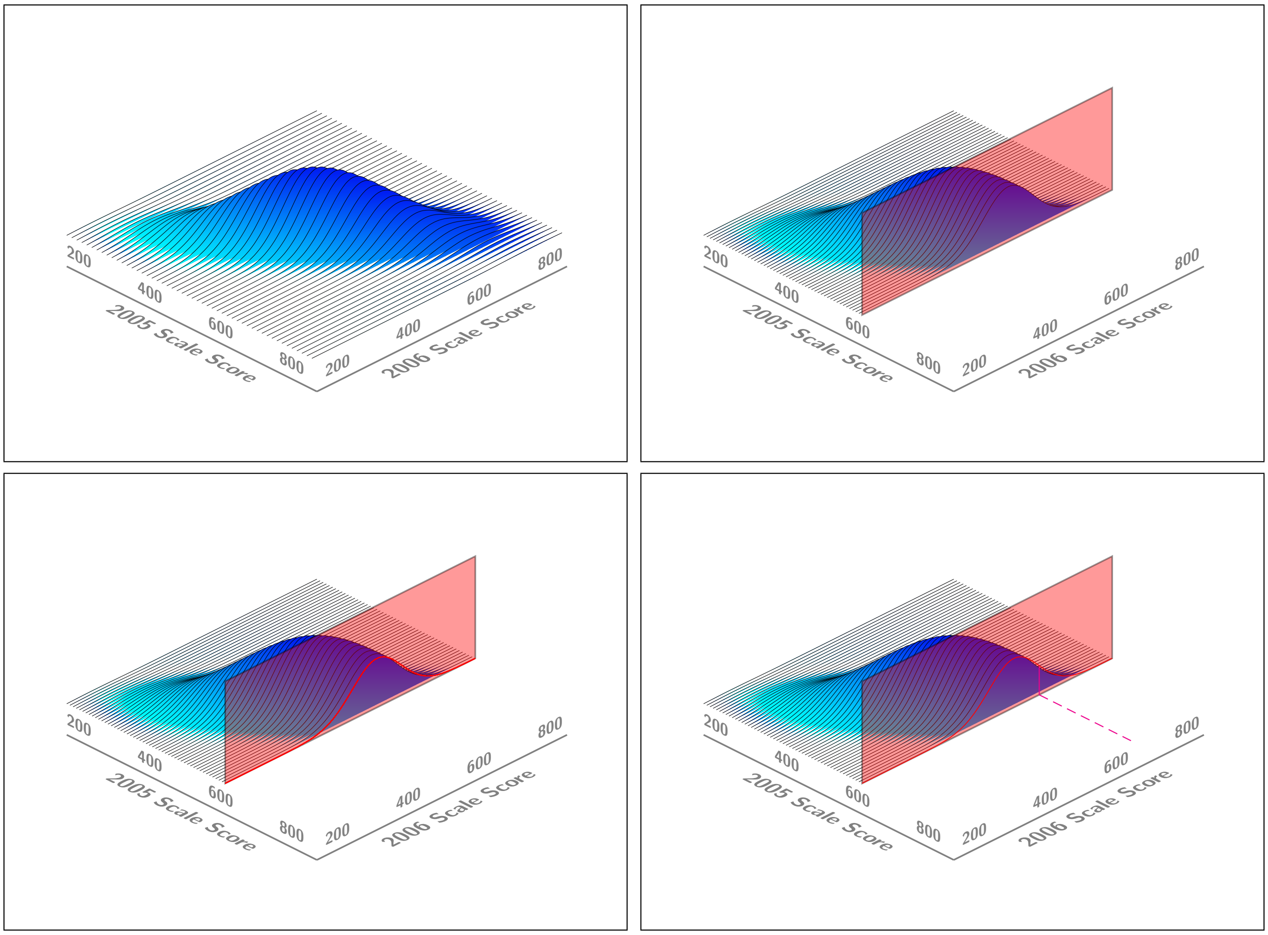


Figure 1: Figures depicting the distribution associated with 2005 and 2006 student scale scores together with the conditional distribution and associated growth percentile

Student Growth Percentile Estimation

Calculation of a student's growth percentile is based upon the estimation of the conditional density associated with a student's score at time t using the student's prior scores at times $1, 2, \ldots, t-1$ as the conditioning variables. Given the conditional density for the student's score at time t, the student's growth percentile is defined as the percentile of the score within the time t conditional density. By examining a student's current achievement with regard to the conditional density, the student's growth percentile normatively situates the student's outcome at time t taking account of past student performance. The percentile result reflects the likelihood of such an outcome given the student's prior achievement. In the sense that the student growth percentile translates to the probability of such an outcome occurring (i.e., rarity), it is possible to compare the progress of individuals not beginning at the same starting point. However, it is important to note that occurrences being equally rare does not necessarily imply that they are equally "good". Qualifying student growth percentiles as "(in)adequate", "good", or as satisfying "a year's growth" is a standard setting procedure requiring external criteria (e.g., growth relative to state performance standards).

Estimation of the conditional density is performed using quantile regression (Koenker, 2005). Whereas linear regression methods model the conditional mean of a response variable Y , quantile regression is more generally concerned with the estimation of the collection of conditional quantiles of Y. Quantile regression provides a more complete picture of both the conditional distribution associated with the response variable(s). The techniques are ideally suited for estimation of the family of conditional quantile functions (i.e., reference percentile curves). Using quantile regression, the conditional density associated with each student's prior scores is derived and used to situate the student's most recent score. Position of the student's most recent score within this density can then be used to qualify deficient/sufficient/excellent growth. Though many state assessments possess a vertical scale, such a scale is not necessary to produce student growth percentiles.

In analogous fashion to the least squares regression line representing the solution to a minimization problem involving squared deviations, quantile regression functions represent the solution to the optimization of a loss function (Koenker, 2005, p. 5). Formally, given a class of suitably smooth functions, $\mathcal{G}$, one wishes to solve

$$\arg\min_{g \in \mathcal{G}} \sum_{i=1}^{n} \rho_\tau(Y(t_i) - g(t_i)), \tag{1}$$

where t_i indexes time, Y are the time dependent measurements, and ρ_τ denotes the piecewise linear loss function defined by

$$\rho_\tau(u) = u \cdot (\tau - I(u < 0)) = \begin{cases} u \cdot \tau & u \geq 0 \\ u \cdot (\tau - 1) & u < 0. \end{cases}$$

The elegance of the quantile regression Expression 1 can be seen by considering the more familiar least squares estimators. For example, calculation of $\arg\min \sum_{i=1}^{n} (Y_i - \mu)^2$ over $\mu \in \mathbb{R}$ yields the sample mean. Similarly, if $\mu(x) = x'\beta$ is the conditional mean represented as a linear combination of the components of x, calculation of $\arg\min \sum_{i=1}^{n} (Y_i - x_i'\beta)^2$ over $\beta \in \mathbb{R}^p$ gives the familiar least squares regression line. Analogously, when the class of candidate functions $\mathcal{G}$ consists solely of constant functions, the estimation of Expression 1 gives the τth sample quantile associated with Y. By conditioning on a covariate x, the τth conditional quantile function, $Q_y(\tau|x)$, is given by

$$Q_y(\tau|x) = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(y_i - x_i'\beta).$$

In particular, if $\tau = 0.5$, then the estimated conditional quantile line is the median regression line.[3]

Following Wei & He (2006), we parameterize the conditional quantile functions as a linear combination of B-spline cubic basis functions.. B-splines are employed to accommodate non-linearity, heteroscedasticity and skewness of the conditional density associated with values of the dependent variable. B-splines are attractive both theoretically and computationally in that they provide excellent data fit, seldom lead to estimation problems (Harrell, 2001, p. 20), and are simple to implement in available software.

Figure 2 gives a bivariate representation of linear and B-splines parameterization of decile growth curves. The assumption of linearity imposes conditions upon the heteroscedasticity of the conditional densities. Close examination of the linear deciles indicates slightly greater variability for higher grade 5 scale scores than for lower scores. By contrast, the B-spline based decile functions better capture the greater variability at both ends of the scale score range together with a slight, non-linear trend to the data.

Currently, calculation of student growth percentiles is performed using R, a language/environment for statistical computing, using the `SGP` package (R Development Core Team, 2008). Other possible software (untested with regard to student growth percentiles) with quantile regression capability include SAS and Stata. Estimation of student growth percentiles is conducted using all available prior data, subject to certain suitability conditions. Given assessment scores for t occasions, ($t \geq 2$), the τ-th conditional quantile for Y_t based upon $Y_{t-1}, Y_{t-2}, \ldots, Y_1$ is given by

$$Q_{Y_t}(\tau|Y_{t-1}, \ldots, Y_1) = \sum_{j=1}^{t-1} \sum_{i=1}^{7} \phi_{ij}(Y_j)\beta_{ij}(\tau), \tag{2}$$

where $\phi_{i,j}$, $i = 1, 2, \ldots, 7$ and $j = 1, \ldots, t-1$ denote the B-spline basis functions. Currently, bases consisting of 7 cubic polynomials are used to "smooth" irregularities found in the multivariate assessment data. A bivariate rendering of this is found is Figure 2 where linear and B-spline conditional deciles are presented. The cubic polynomial B-spline basis functions model the heteroscedasticity and non-linearity of the data to a greater extent than is possible using a linear parameterization. Estimation of student growth percentiles utilizes all available student achievement data subject to some suitability requirements.

[3] For a detailed treatment of the procedures involved in solving the optimization problem associated with Expression 1, see Koenker (2005), particularly Chapter 6.

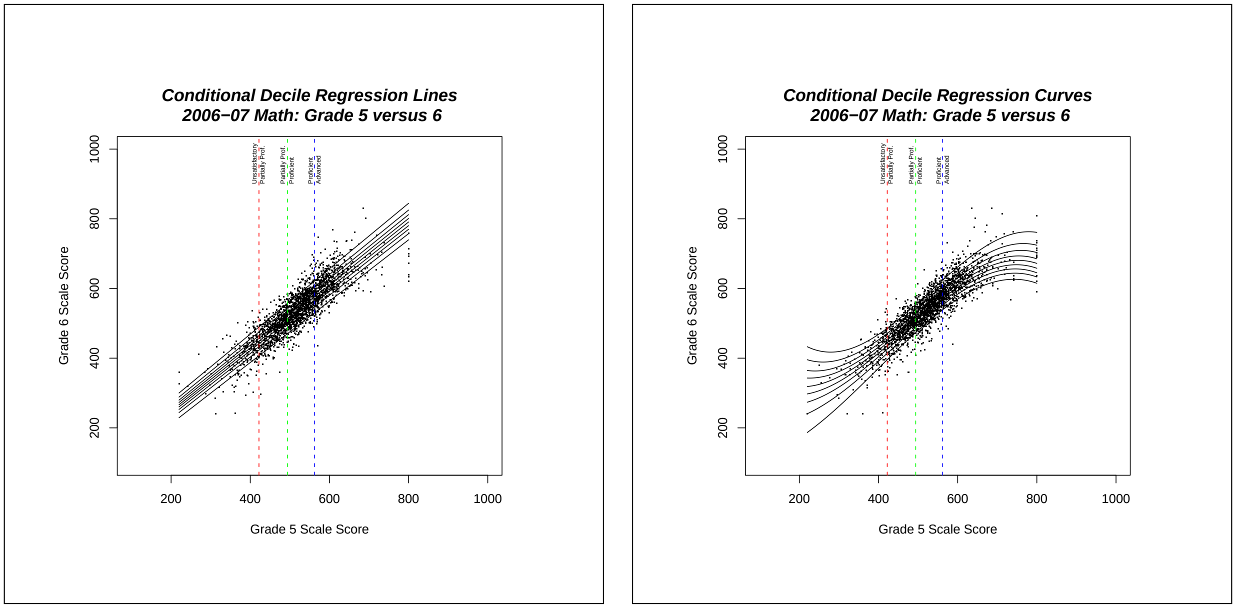


Figure 2: Linear and B-spline conditional deciles based upon bivariate math data, grades 5 and 6

Data Requirements

Calculation of a student growth percentile corresponding to a student's current score requires conditioning the current score for students on all available consecutive prior scores. For the purposes of state wide calculation of student growth percentiles, growth percentile analyses are performed for each subject and each grade separately. Each grade by subject cohort's data are analyzed so that each student receives a growth percentile in each subject they were tested. As with any analysis of large scale observational data, missing or incomplete data is an issue. 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 to condition upon. Because the conditional distribution is derived based upon the observations of thousands of other students in the state population, operational decisions are made that define which student scores are part of the norming population and which are not:

- Students must have at a minimum two scores coming from consecutive years and grades (e.g., grade 4 in 2007 and grade 5 in 2008).
- Only data from consecutive grades is used to estimate the conditional density for a student. Students repeating/skipping grades associated with the most recent score (i.e., the dependent variable of 2) are not included in the analyses at this time.[4]
- Only data from consecutive years is used to estimate the conditional density for a student. For example, given a student possessing 2005, 2007, and 2008 CSAP reading data, only scale scores from 2008 and 2007 are used to calculate their student growth percentile, not their 2005 score. 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

[4] In theory, the vertical scale allows for treating a scale score on a CSAP examination as "equivalent" to the scale score a student would receive on another grade's exam. However, initial analysis suggests that dramtically different results can occur depending upon the cohort with which the repeated students are compared to.

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 all students, 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; student with exactly 3 consecutive CSAP scores are compared with all other students with *at least* 3 scores; and so on. Student growth percentiles are normed relative to all students possessing scores on the same test as the student.

The above definitions lead to the estimation of student growth percentiles for the overwhelming majority of the students in a state assessment program with 2 or more scale scores.

Student Growth Percentile Properties

Student growth percentiles possess a number of attractive properties from both a theoretical as well as a practical perspective. Foremost among practical considerations is that the percentile descriptions are familiar and easily communicated to teachers and other non-technical stakeholders. Furthermore, implicit within the percentile quantification of student growth is a statement of probability. Questions of "how much growth is enough?" or "how much is a year's growth?" ask stakeholders to establish growth percentile thresholds deemed adequate. These thresholds establish growth standards that translate to probability statements. In this manner, percentile based growth forms a basis for discussion of rigorous yet attainable growth standards for all children supplying a normative context for Linn's (2003) existence proof with regard to student level growth.

In addition to practical utility, student growth percentiles possess a number of technical attributes well suited for use with assessment scores. The more important theoretical properties of growth percentiles include:

Robustness to outliers Estimation of student growth percentiles are more robust to outliers than is traditionally the case with conditional mean estimation. Analogous to the property of the median being less influenced by outliers than is the median, conditional quantiles are robust to extreme observations. This is due to the fact that influence of a point on the τ-th conditional quantile function is not proportional (as is the case with the mean) to the distance of the point from the quantile function but only to its position above or below the function (Koenker, 2005, p. 44).

Uncorrelated with prior achievement Analogous to least squares derived residuals being uncorrelated with independent variables, student growth percentiles are not correlated with prior achievement. This property runs counter to current multilevel approaches to measuring growth with testing occassion nested within students (Singer & Willett, 2003). These models, requiring a vertical scale, fit lines with distinct slopes and intercepts to each student. The slopes of these lines represent an average rate of increase, usually measured in scale score points per year, for the student. Whereas a steeper slope represents more learning, it is important to understand that using a normative quantification of growth, one cannot necessarily infer that a low achieving student with a growth percentile of 60 "learned as much" as a high achieving student with the same growth percentile. Growth percentiles bypass questions associated with magnitude of learning and focus on normatively quantifying changes in achievement.

Equivariance to monotone transformation of scale An important attribute of the quantile regression methodology used to calculate student growth percentiles is their invariance to monotone transformations of scale. This property, denoted by Koenker (2005) as *equivariance to monotone transformations* is particularly helpful in educational assessment where a variety of scales are present for analysis, most of which are related by some monotone transformation. For example, it is a common misconception that one needs a vertical scale in order to calculate growth. Because vertical and non-vertical scales are related via a monotone transformation, the student growth percentiles do not change given such alterations in the underlying scale. This result obviates much of the discussion concerning the need for a vertical scale in measuring growth.[5]

Formally, given a monotone transformation h of a random variable Y,

[5] As already noted with regard to pediatrics, the existence of nice "vertical" scales for measuring height and weight still leads to observed changes being normed.

$$Q_{h(Y)|X}(\tau|X) = h(Q_{Y|X}(\tau|X)).$$

This result follows from the fact that $\Pr(T < t|X) = \Pr(h(T) < h(t)|X)$ for monotone h. It is important to note that equivariance to monotone transformation does not, in general, hold with regard to least squares estimation of the conditional mean. That is, except for affine transformations h, $E(h(Y)|X) \neq h(E(Y|X))$. Thus, analyses built upon mean based regression methods are, to an extent, scale dependent.

Student growth percentiles derived using quantile regression procedures possess a number of attractive properties that make them ideal candidates as normative descriptors of student growth. An obvious criticism of growth percentiles, as well as any other normatively derived quantity, is that it is purely descriptive and inappropriate for NCLB like determinations requiring adequacy judgments. The next section confronts this criticism and shows that growth percentiles and the methodology underlying them are ideally suited for tasks involving criterion reference growth.

Given multivariate normal assessment data, calculation of student growth percentiles is approximately equivalent to conversion of linear regression based least squares residuals to percentiles using a parametric assumption about the distribution of residuals to convert residual to percentile. However, empirical analysis of percentiles derived from least squares residuals demonstrate large bias in a large number of cases. This issue will be addressed more fully in the goodness-of-fit section on Page 7.

Accuracy and Precision of Student Growth Percentiles

As with all statistical procedures, the accuracy and precision of student growth percentiles is dependent upon a number of factors including models assumptions, sparcity of data at the scale extremes, and measurement error. The importance of precision is directly related to the inferences one wishes to draw from the growth percentiles. As with achievement comparison using scale scores, fine grained comparisons using growth percentiles cannot be sustained. Definitive statements to the effect that a student with a growth percentile 5 points higher than another student's demonstrates "superior growth" are not possible. However, like with pediatrics, it's possible to use the normative data to define regions of growth percentiles representing, for example, problem, average, and superior growth. Judgments concerning thresholds between the regions can be made based upon statistical as well as practical information, especially information about what growth percentiles are necessary to reach standards based achievement targets.

To investigate precision, a set of analyses was conducted that compared scale score differences related to 75th-20th percent growth and 75th-50th percentile growth to understand both the magnitudes of such differences across the range of prior achievement and its relation to the conditional standard error of measurement for the exam. Figure 3 uses growth percentiles for 2006 grade 10 math to examine scale score difference across the range of prior grade 9 math achievement in 2005.

The results of Figure 3 suggest that at the scale score extremes, error in measurement exceed scale score differences associated with 75th-50th percentile growth but does not exceed 75th-20th percentile scale score differences.[6] Placing growth on the percentile metric appears to supply great enough precision for coarse judgments at the individual level. For example, a deficient/average/superior trichotomy appears defensible. Of course, aggregation of student growth percentiles, as with any statistic, provides greater precision and the possibility for finer distinctions to be made.

Goodness-of-Fit

Examination of goodness-of-fit was conducted by examining the estimated conditional density against the theoretical density. Despite the use of B-splines to accommodate heteroscedasticity and skewness of the conditional density, assumptions are made concerning the number and position of spline knots that impact the percentile curves that are fit. With an infinite CSAP population, at each prior scale score, with perfect model fit, the expectation is to have 10 percent of the estimated growth percentiles between 0 and 9, 10 and 19, 20 and 29, . . . , and 90 and 99. Deviations from 10 percent would be indicative of lack of model fit. To motivate the extent to which lack of fit might contaminate student growth percentiles, student growth percentiles are also calculated using a linear parameterization in addition to B-splines. Though the maximum number of

[6] Results for other subjects and other grades demonstrate similar overall patterns.

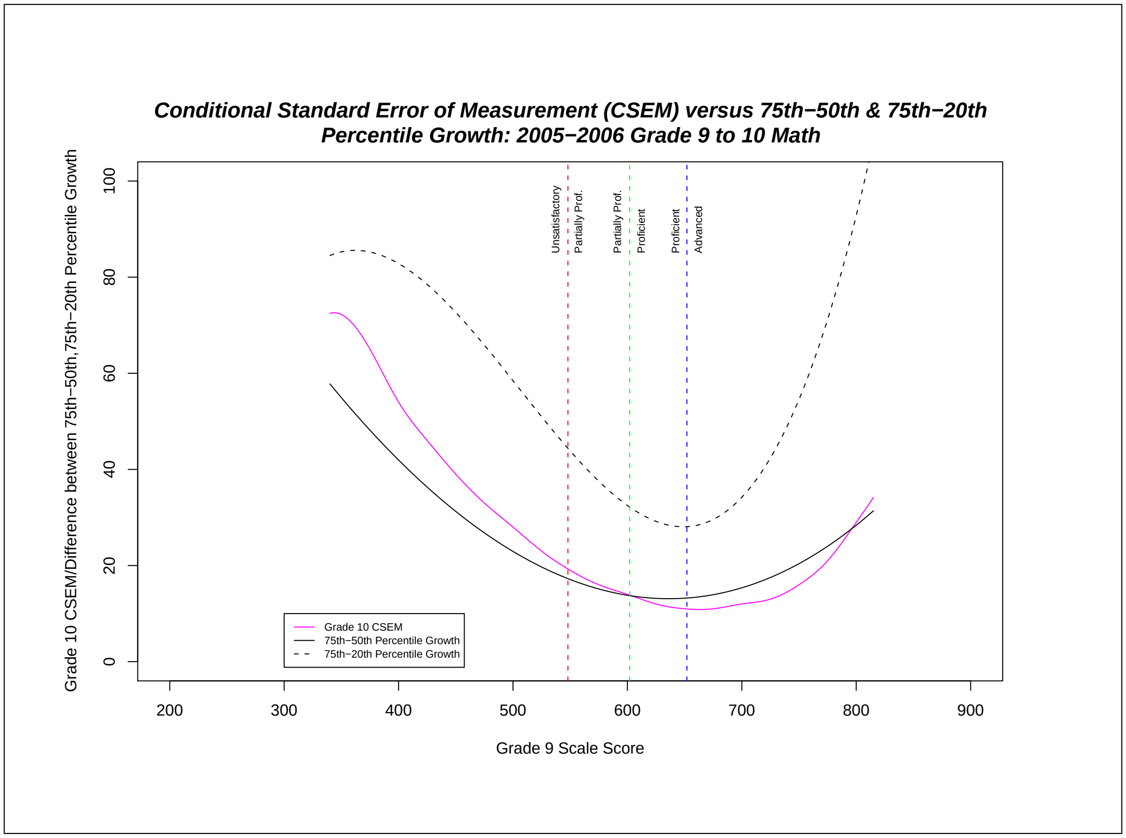


Figure 3: Scale score difference associated with 20th, 50th, and 75th percentiles relative to the CSEM for 2006 grade 10 math

prior were used to calculate the student growth percentiles, the result of linear versus B-spline estimation is roughly what is depicted in the bivariate scatterplots of Figure 2

Using 2007 math, reading, and writing scores as the dependent variable, estimation of student growth percentiles was conducted exactly as outlined previously using the maximum number of consecutive year by grade CSAP scores for each student. Percentages of student growth percentiles at the 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, and 90th percentiles were calculated based upon the decile of the prior year's scale score (each decile consists of approximately 5,000 students). Results for the B-spline and linear parameterizations for math, reading and writing are given in Tables 3 to 2 (pages 16 to 15).

Beginning with the growth percentiles derived using linear percentile curves (Tables 1 to 2), the results indicate, not surprisingly, rather poor fit in the highest and lowest prior scale score deciles. For example, in 2007 grade 4 math (Table 1), 20.46% of the estimated student growth percentiles were less than 10 for the *lowest* 2006 achievement decile of students. The result indicates model bias against low achieving students in grade 4 math. Similarly, 22.21% of the estimated student growth percentiles were less than 10 for the *highest* 2006 achievement decile of students. As with the low achievers in grade 4 math, the linear parameterization yields biased growth percentiles for high achieving students.

Contrasting the linear results with the B-spline results shows the impact that smoothing has on goodness-of-fit. Using the same 2007 grade 4 math data, the B-spline derived student growth percentiles yielded 9.87% and 10.11% of growth percentiles less than 10 for the lowest and highest 2006 achievers, respectively. Though slightly biased, the results are appreciably better than those derived from the linear model. Across all 2006 achievement deciles, grades, and subjects, the B-spline derived percentiles performed very well. Absolute deviations from expectation are generally quite small (less than 1.5 percent) with the largest absolute deviation just less than 4. By contrast, the results based upon the linear model can differ markedly from expected and would not be suitable for reporting purposes. Dependending upon the opinion of the technical review and the TAP, additional smoothing can be investigated to try and improve fit even further.

Percentile Growth Trajectories

Having described how to normatively quantify growth in terms of student growth percentiles, we now demonstrate how student growth percentiles can form the basis of growth-to-standard discussions regarding what is adequate growth. This is the process of going from a norm-referenced to a criterion-referenced standard of growth and is, in many ways, analogous to current normative and criterion referenced understandings of achievement. We begin by unpacking the imbroglio of terminology currently associated with discussion of student growth.

In a discussion of growth, NCLB, and vertical scaling, Yen (2007), provides a list of questions regarding growth taken from a survey of parents, teachers, and administrators:

Parent Questions:

- Did my child make a year's worth of progress in a year?
- Is my child growing appropriately toward meeting state standards?
- Is my child growing as much in Math as Reading?
- Did my child grow as much this year as last year?

Teacher Questions:

- Did my students make a year's worth of progress in a year?
- Did my students grow appropriately toward meeting state standards?
- How close are my students to becoming Proficient?
- Are there students with unusually low growth who need special attention?

Administrator Questions:

- Did the students in our district/school make a year's worth of progress in all content areas?
- Are our students growing appropriately toward meeting state standards?
- Does this school/program show as much growth as that one?
- Can I measure student growth even for students who do not change proficiency categories?
- Can I pool together results from different grades to draw summary conclusions?

As Yen concludes, all these questions rest upon a desire to understand whether observed student progress is "reasonable or appropriate" (Yen, 2007, p. 281). Moreover, the questions admit two paths to their resolution: the absolute and the normative. As discussed previously, student growth percentiles provide a normative and, we more informative way to address these questions than an absolute metric on which to interpret growth.

Methodology for Defining Adequate Growth

To adequately address the notion of defining enough, adequate, or a year's growth, *aspirational* growth must be distinguished from *actual* growth:

Actual What *is* a current year's growth?

Aspirational What *should* a current year's growth be?

Answering the second question establishes a threshold distinguishing adequate from inadequate growth. To make such a distinction requires answering the first question which defines a norm: What is the range of growth currently observed? Aspirational growth for each student should be possible—again, Linn's existence proof applied at the individual level (Linn, 2003).

Student growth percentiles provide an elegant means of answering the first question: What is a current year's growth? Answering the second question requires a qualification distinguishing adequate growth from inadequate growth. For example, the current growth-to-standard criterion utilized by most states defines adequate growth as growth leading toward proficiency. Using the quantile regression estimates derived by solving Equation 2, the relationship between prior scores and the range of outcomes based upon growth percentiles is articulated and can be exploited. It is straightforward to calculate the growth percentile necessary for each student to reach that level of achievement. This threshold could then be used to distinguish adequate from inadequate growth. A benefit of using the percentile scale for growth is that the threshold has a normative context that can be used to set criterion referenced aspirational goals that are reasonable.

To establish growth percentile targets (i.e., define what growth *should* be for each student) in terms of performance levels, it is necessary to investigate what growth percentile is necessary to reach the desired performance level threshold based upon the student's achievement history. Intuitively, the lower one's scale score, the higher their growth percentile must be in order for them to reach the desired target.

Establishing criterion referenced growth thresholds requires consideration of multiple future achievement scenarios. Instead of inferring ((as is the case with linear trajectory models like those based on HLM growth analyses) that prior student growth is indicative of future student growth (e.g., linearly projecting student achievement into the future based upon past rates of change), predictions of future student achievement are contingent upon initial student status (where the student starts) and subsequent rates of growth (the rate at which the student grows). Thus, instead of statements such as, "Student X is projected to be (not) proficient in three years", discussions that hypothesize different rates of growth and their consequences with regard to future achievement are considered: "Given that Student X has this achievement history and grows over the coming three years at rate G, we anticipate Student X will (not) be proficient."

Following analysis of state data to calculate student growth percentiles by estimating coefficients in Equation 2, the state has used it entire database of assessment data for students currently enrolled to quantify rates of growth over the most recent year. The results of these analyses (i.e., coefficient matrices of quantile regression coefficients) connect prior student achievement data with current achievement. For example, in 2008, growth percentile analyses for students in grade 4 created quantile regression coefficient matrices based upon grade 3 student scores in 2007. These coefficient matrices relate achievement across the entire scale score dimension 2007 and what different percentile growth rates (between 1 and 99) lead to in terms of 2008 achievement. Thus, these coefficient matrices can be used to estimate what growth percentile it will take for a student in grade 3 in 2008 to reach a desired grade 4 achievement outcome in 2009. This one year percentile growth projection can be used as a target, following the 2009 achievement test, to judge whether the student has made "adequate growth".

Similarly, to form a two year percentile growth projection/trajectory for this 3rd grade student, Colorado uses the 1 year projection together with the coefficient matrices that relate 5th grade performance with prior 3rd and 4th grade performance. Using the student's actual 3rd grade achievement and the 99 different growth percentile outcomes available for 4th grade achievement, Colorado examine what *consecutive* growth percentiles at a given level will yield for the student in grade 5. Thus, Colorado can estimate based upon the most recently available data what consecutive 75th percentile growth will yield for a given student. The calculation of three year percentile growth projections/trajectories follows similarly for this this student with the one actual achievement score being combined along with 99 two year growth scenarios to estimate what consecutive growth across three years at the different growth levels will lead to.

Colorado uses these 1, 2, and 3 year percentile growth projections/trajectories for each student to answer questions regarding what level of growth (quantified as an annual growth percentile) it will take for a student to reach proficient.[7] Presently, each student in the state of Colorado with a score in the current year receives 9 *percentile growth trajectories* indicating what level of annual growth (quantified as a growth percentile) is necessary for that student to reach each of the 3 cutpoints between the 4 performance levels in 1, 2, and 3 years (or by grade 10, whichever comes first).[8] These projections, based upon the most recent progress of students in the state, represent Colorado's best estimate of what it will take for each student reach these 3 achievement levels. These growth percentile projections/trajectories are used by the Colorado Growth Model as criteria to judge whether or not the students level of growth is adequate.

[7] Any achievement target X can be used to answer, "What level of growth is necessary to reach X within the next 3 years?" or "What level of growth is necessary to maintain above X for the next 3 years?".

[8] Note that though is is necessary two have at least two scores to determine growth, only 1 score (i.e., a starting point) is necessary to calculate what it will take for a student to reach future performance targets. For example, it is possible to provide third graders who possess a single score in each subject with estimates of what growth it will take for them to reach 4th, 5th, and 6th grade achievement targets.

Percentile growth projections/trajectories are more easily understood with the assistance pictures. Figures 4 to 9 (Pages 19 to 24) depict three growth scenarios each in math and reading for students beginning in third grade at each of three performance level cutpoints (i.e., between unsatisfactory/partially proficient, partially proficient/proficient, and proficient/advanced). The figures depict the four state performance levels across grades 3 to 10 in color together with the 2007 achievement percentiles superimposed in white. Beginning at grade 3 at the given cutpoint, a grade 4 achievement projection is made based upon the the growth percentile derived using prior 3rd to 4th grade student progress. Next, using this projected 4th grade score combined with the 3rd grade score, a 5th grade achievement projection is made using prior student progress from 3rd and 4th to 5th. The process repeats to plot out different "growth percentile trajectories". The figures allow stakeholders to consider what 10th, 25th, 40th, 50th, 60th, 75th, and 90th percentile growth (sustained year-over-year) yields for students with three hypothetical starting points in the 3rd grade. Like all forecasting, these projections are not exact, especially as the timeframe extends. However, the charts do allow for a "bird's eye view" that can aid stakeholders in growth standard setting.

Consider Figure 4, math growth trajectories for a student beginning at the unsatisfactory/partially proficient threshold. Based upon the achievement percentiles, approximately 7 percent of the population of 3rd graders rate as unsatisfactory. Moving toward grade 10, the percentage of unsatisfactory students increases dramtically to near 35 percent. The black lines in the figure represent six different growth scenarios for the student based upon consecutive growth at a given growth percentile, denoted by the right axis. At the lower end, for example, consecutive 25th percentile growth leaves the student, unsurprisingly, mired in the unsatisfactory category. Consecutive 40th, 50th and 60th percentile growth also leave the child in the unsatisfactory category. This demonstrates how difficult (based upon current rates of progress) it is for students to move up in performance level in math. With the green region representing proficient, a student would need to demonstrate growth percentiles in excess of 75 to reach proficiency showing how unlikely such a event currently is.

If the goal of an accountability system is universal proficiency, then the growth percentile targets can be set accordingly. One of the strengths of quantifying student growth normatively is that growth percentile targets quickly translate into the likelihood of such an event occurring. This dimension of student improvement as it relates to accountability is absent from most growth-to-standard discussions. Today, achievement mandates are stipulated based upon the moral imperative of high standards for all children. Given current progress of students, it is unlikely that the sustained levels of growth necessary to reach these standards will ever occur. A fundamental dictum of moral philosophy ascribed to Kant is that "ought implies can": If someone *ought* to do something, they can do it, in the sense that they have the possibility/capacity to do it (Betebenner & Howe, 2007). Growth percentiles bring Kant's dictum to the fore when considering criterion referenced growth standards.

Discussion

This paper has introduced student growth percentiles and the quantile regression analysis techniques used for their calculation. Currently, the recently adopted Colorado Growth Model uses student growth percentiles and the associated methodology as the means of fulfilling recently enacted legislation (HB 07-1048) directing the Department to develop and implement a growth model. The purpose of developing these percentiles is to provide educational stakeholders at various levels (students, parents, teachers, principals, administrators, and policy makers) a simple yet rigorous means by which to understand student progress given the vast amounts of longitudinal data currently available.

Performance of the quantities with regard to goodness-of-fit analyses is excellent, indicating that the student growth percentiles are accurate descriptions of student progress. Descriptive results support the contention that the growth percentiles are not biased. Results for student growth percentiles are used to calculate percent growth projections/trajectories that are used to define adequate growth for a student. If the student's growth exceeds percentile rates determined to put them on track to proficient within three years, then the student is deemed to be on track to proficient. For example, if the student's growth exceeds percentiles rates determined to put them on track to remain proficient for three years, then the student is deemed to be on track to remain proficient. Percentages of student's reaching their growth targets can then be used as a summary measure of school performance.

Student growth percentiles provide a descriptive measure of *what is*, that is, a quantification of how much a student grew. Questions of *what should be* coincide with decisions about whether growth is "enough" or "adequate". These qualifications rest outside of statistics and require reasoned judgment on the part of stakeholders to set such standards. The student growth percentile metric serves to inform the standard setting procedure by immediately relaying *what is possible*. Only by considering, what is, what should be, and what is possible simultaneously, can growth standard setting and accountability systems built upon such standards be equitable, just, and truly informed.

References

Betebenner, D. W. (2008). Toward a normative understanding of student growth. In L. A. Shepard & K. E. Ryan (Eds.), *Festschrift in honor of the life and work of Robert L. Linn.* New York: Routledge.

Betebenner, D. W., & Howe, K. R. (2007). *On NCLB, equal educational opportunity and closing achievement gaps.* (Unpublished. To appear in Educational Researcher)

Harrell, F. E. (2001). *Regression modeling strategies.* New York: Springer.

Koenker, R. (2005). *Quantile regression.* Cambridge: Cambridge University Press.

Linn, R. L. (2003, July). *Accountability: Responsibility and reasonable expectations* (Tech. Rep.). Los Angeles, CA: Center for the Study of Evaluation, CRESST.

R Development Core Team. (2008). *R: A language and environment for statistical computing.* Vienna, Austria. (3-900051-07-0)

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis.* New York: Oxford University Press.

Wei, Y., & He, X. (2006). Conditional growth charts. *The Annals of Statistics*, *34*(5), 2069–2097.

Yen, W. M. (2007). Vertical scaling and No Child Left Behind. In N. J. Dorans, M. Pommerich, & P. W. Holland (Eds.), *Linking and aligning scores and scales* (pp. 273–283). New York: Springer.

Goodness Of Fit

The following tables present percentages of 2007 student growth percentiles at or below 10, 20, 30, 40, 50, 60, 70, 80, and 90 based upon the 2006 scale score decile of the student. Given adequate model fit, one would expect a uniform distribution of growth percentiles across the prior score distribution and for there to be $n\%$ of the observed percentiles to lie at or below n. Deviations suggest a model parameterization that doesn't adequately fit the data. Such deviations can be seen in the results associated with a linear parameterization given in Tables 2 through 2. The B-spline parameterization results in much better fit (Tables 3 to 4).

2006	2006	Observed Percentages of SGP in Given Percentile Range									
Grade	SS Decile	0 to 9	10 to 19	20 to 29	30 to 39	40 to 49	50 to 59	60 to 69	70 to 79	80 to 89	90 to 99
4	0	20.46	14.93	9.87	9.47	7.76	7.89	7.10	6.96	6.43	9.13
4	1	11.99	13.05	9.15	10.82	10.61	9.43	9.17	8.92	8.50	8.37
4	2	8.71	10.72	9.49	9.99	10.23	10.38	10.79	10.07	10.66	8.97
4	3	7.21	10.09	9.04	10.62	10.58	10.16	10.73	11.34	10.79	9.44
4	4	6.32	9.04	8.27	9.50	10.64	10.66	11.26	11.63	12.42	10.27
4	5	5.48	8.15	8.30	9.48	10.88	12.15	11.03	11.85	11.41	11.26
4	6	5.40	8.76	8.16	9.18	10.16	11.01	12.16	11.69	11.91	11.57
4	7	5.82	8.58	8.00	9.81	10.22	10.32	11.24	11.61	11.86	12.53
4	8	6.42	10.48	9.64	11.24	10.84	10.29	10.06	9.57	10.29	11.18
4	9	22.21	16.25	10.09	10.01	8.06	7.51	6.52	6.35	5.72	7.30
5	0	16.57	11.59	10.37	9.23	8.94	7.89	7.53	7.61	8.66	11.59
5	1	11.80	10.98	10.94	10.15	10.92	8.99	9.25	9.52	9.05	8.40
5	2	10.33	9.79	9.52	9.79	11.25	10.62	10.70	10.43	9.67	7.90
5	3	8.56	10.10	9.50	10.94	11.06	10.19	11.27	10.06	11.02	7.30
5	4	7.34	9.09	9.96	9.83	11.71	10.78	10.98	10.74	11.47	8.09
5	5	6.93	8.19	9.35	9.83	12.14	11.61	10.86	11.74	10.30	9.05
5	6	6.56	8.92	9.78	9.82	11.30	10.83	10.95	11.46	11.05	9.34
5	7	6.83	8.65	9.72	10.07	11.72	10.67	10.24	10.50	11.62	9.98
5	8	7.77	10.43	9.53	10.15	11.11	9.77	10.79	9.96	10.21	10.28
5	9	18.07	12.47	11.24	9.80	9.01	7.95	7.51	7.26	7.06	9.64
6	0	17.19	10.61	8.88	7.48	7.92	8.34	7.71	8.27	10.31	13.29
6	1	12.02	10.04	9.85	10.07	9.86	10.67	10.63	9.54	9.12	8.20
6	2	9.21	10.37	9.85	10.71	10.67	11.72	10.35	10.77	9.25	7.09
6	3	8.20	10.36	11.41	10.93	10.92	11.89	10.34	10.17	9.39	6.39
6	4	8.59	10.25	10.98	10.63	10.89	11.67	10.91	9.88	8.92	7.28
6	5	7.74	9.96	10.43	10.72	10.45	11.44	10.57	10.89	9.92	7.89
6	6	7.56	9.24	9.92	10.80	9.84	12.19	10.18	10.60	10.52	9.14
6	7	7.12	8.07	9.42	9.61	10.39	11.15	10.94	10.91	12.05	10.35
6	8	7.25	9.00	9.15	9.06	9.97	11.48	10.22	10.80	11.62	11.44
6	9	16.46	12.30	9.83	9.31	8.51	9.06	7.95	7.76	8.37	10.44
7	0	14.66	9.31	7.98	7.67	8.05	7.81	8.67	9.21	11.26	15.39
7	1	10.87	10.07	11.60	9.82	9.30	9.53	10.54	10.09	9.71	8.47
7	2	10.09	10.51	11.23	11.22	11.41	10.76	10.26	9.74	8.71	6.08
7	3	9.06	10.16	12.50	11.55	10.89	10.52	10.05	10.07	9.03	6.18
7	4	8.67	10.34	12.28	11.25	10.76	11.07	10.50	9.99	9.10	6.04
7	5	8.69	10.97	12.01	10.73	10.42	10.37	10.55	10.72	9.34	6.21
7	6	7.43	9.85	12.01	10.21	10.47	11.49	10.57	10.15	10.17	7.64
7	7	7.74	9.66	10.82	9.68	10.36	10.09	10.53	10.15	11.40	9.57
7	8	7.41	8.48	9.78	8.93	9.53	9.95	10.68	11.02	11.51	12.73
7	9	15.87	10.54	9.21	8.58	8.19	7.96	8.16	8.50	9.54	13.45
8	0	14.24	9.91	9.04	7.81	7.86	8.12	8.27	9.67	10.81	14.27
8	1	12.31	10.01	11.60	9.88	10.33	9.27	10.33	8.86	9.34	8.06
8	2	9.99	10.04	11.85	10.47	10.10	10.23	10.45	10.45	10.06	6.36
8	3	9.31	10.21	11.57	11.25	10.71	11.11	9.56	10.00	9.79	6.49
8	4	8.44	10.13	11.48	10.52	11.01	11.43	10.81	9.73	9.02	7.42
8	5	7.45	9.58	11.80	11.07	11.32	10.65	10.58	9.74	9.79	8.02
8	6	7.27	9.69	10.81	10.62	10.99	10.66	10.68	10.48	10.18	8.62
8	7	7.16	9.67	10.90	10.04	9.67	9.98	10.50	11.59	10.04	10.44
8	8	7.78	10.04	9.98	9.10	9.05	9.21	10.26	10.22	12.35	12.00
8	9	17.11	10.00	9.79	8.63	8.27	8.48	8.46	8.59	9.38	11.28
9	0	13.70	9.31	8.43	8.70	7.76	7.88	8.76	9.37	11.05	15.04
9	1	12.62	10.45	10.11	9.56	10.33	9.80	10.53	9.09	9.39	8.11
9	2	11.45	10.31	10.16	10.81	10.45	10.13	11.07	9.79	9.32	6.52
9	3	10.99	11.32	10.57	10.63	10.19	10.65	10.71	9.22	8.81	6.90
9	4	9.21	10.18	10.55	10.72	10.64	10.64	12.08	10.66	9.17	6.16
9	5	8.11	9.37	10.10	10.38	11.19	11.62	12.00	10.77	9.68	6.78
9	6	6.71	9.67	10.54	10.56	10.93	11.16	11.55	10.91	10.29	7.68
9	7	6.44	9.78	9.67	10.00	10.42	10.41	12.17	10.64	11.51	8.96
9	8	7.64	9.30	9.87	10.31	10.02	9.76	11.06	10.23	10.95	10.85
9	9	15.91	10.19	9.08	8.13	7.78	7.94	8.58	8.26	9.38	14.75
10	0	10.93	9.64	8.41	7.70	7.96	9.07	9.50	11.21	9.84	15.75
10	1	12.69	11.49	9.76	10.32	9.86	9.47	10.52	9.72	8.75	7.41
10	2	12.58	12.01	10.30	10.75	10.06	10.40	9.67	9.81	7.72	6.70
10	3	10.93	11.35	11.03	10.28	10.58	9.76	9.09	11.11	8.91	6.96
10	4	9.20	11.83	10.31	10.02	10.25	9.96	10.71	11.78	8.46	7.48
10	5	8.56	10.90	10.26	10.62	9.82	10.30	10.84	11.52	9.22	7.98
10	6	8.02	9.84	9.86	10.09	10.07	10.88	10.55	12.30	9.68	8.69
10	7	6.93	9.61	9.69	10.72	11.07	10.08	10.62	12.10	9.77	9.41
10	8	7.71	11.07	9.92	9.12	10.02	9.92	10.80	10.68	9.77	10.98
10	9	13.15	10.65	8.90	8.62	8.74	8.56	8.54	9.77	9.24	13.83

Table 1: Decile by grade examination of goodness-of-fit for 2007 CSAP math student growth percentiles using a linear parameterization

2006 Grade	2006 SS Decile	Observed Percentages of SGP in Given Percentile Range									
		0 to 9	10 to 19	20 to 29	30 to 39	40 to 49	50 to 59	60 to 69	70 to 79	80 to 89	90 to 99
4	0	19.55	9.30	9.06	8.47	8.84	7.53	7.83	8.94	9.00	11.47
4	1	13.94	10.39	11.87	10.76	9.83	10.49	9.57	9.28	8.60	5.26
4	2	10.34	9.78	11.20	10.98	10.61	11.00	10.90	10.07	9.20	5.92
4	3	8.78	9.87	9.85	10.49	10.60	11.08	11.48	10.76	9.85	7.23
4	4	6.86	8.36	10.12	11.51	11.41	11.51	11.41	10.10	10.51	8.21
4	5	6.75	7.63	9.25	9.80	11.03	10.63	11.32	11.36	11.72	10.50
4	6	6.16	7.38	9.33	10.04	10.27	10.41	10.92	11.55	12.08	11.85
4	7	5.98	8.30	8.79	9.72	10.33	10.07	9.92	11.00	12.12	13.77
4	8	8.03	9.00	9.69	9.91	9.37	10.01	9.53	9.87	10.14	14.46
4	9	23.60	10.27	10.47	8.36	7.69	7.09	7.01	7.09	6.83	11.59
5	0	15.02	9.65	9.65	7.73	8.14	8.06	8.12	9.01	10.65	13.98
5	1	11.78	10.90	12.26	10.41	10.10	10.39	10.31	9.69	9.03	5.14
5	2	10.87	10.68	11.37	11.26	10.79	10.61	11.26	10.27	8.30	4.59
5	3	8.19	9.84	12.52	11.51	11.62	10.67	10.63	10.34	9.35	5.32
5	4	8.01	10.01	11.21	10.94	11.05	10.80	10.73	10.76	10.32	6.17
5	5	7.13	9.37	10.64	10.46	10.57	11.31	10.27	11.10	10.99	8.16
5	6	6.93	9.35	10.84	9.77	10.31	10.53	9.95	11.14	11.10	10.07
5	7	6.86	9.24	10.94	9.37	10.17	9.97	10.46	10.79	11.20	11.00
5	8	8.63	9.60	10.14	9.56	9.15	9.71	9.69	9.69	11.06	12.77
5	9	17.62	11.14	9.82	8.25	7.64	7.50	7.72	7.56	8.56	14.19
6	0	18.54	10.72	8.87	8.97	8.03	7.61	7.09	8.07	9.16	12.95
6	1	11.11	10.71	10.47	10.11	10.28	9.60	10.84	10.41	9.90	6.56
6	2	9.09	10.27	11.09	10.81	10.35	10.68	10.80	10.89	9.31	6.71
6	3	8.74	10.48	10.68	10.33	10.96	10.76	11.18	10.83	9.74	6.28
6	4	7.13	8.68	10.18	10.88	11.26	11.50	11.33	10.35	10.92	7.76
6	5	7.83	9.60	9.28	10.87	11.24	11.20	10.69	10.44	10.87	7.97
6	6	8.20	8.68	10.21	10.13	10.21	10.98	10.59	10.96	10.60	9.43
6	7	9.17	8.94	9.52	9.89	10.25	9.86	9.95	9.66	11.36	11.40
6	8	11.33	9.60	9.56	9.12	8.82	9.04	9.16	10.00	10.51	12.86
6	9	19.08	11.65	9.64	8.83	8.50	7.59	7.77	8.04	8.24	10.66
7	0	16.85	9.83	9.67	7.74	8.35	7.72	8.29	8.25	10.02	13.28
7	1	10.55	11.09	10.88	9.82	10.27	10.18	10.53	9.90	9.40	7.39
7	2	9.47	10.62	11.27	10.74	10.23	10.92	10.64	10.11	9.84	6.17
7	3	8.33	9.41	11.69	10.91	10.53	10.83	9.88	11.21	10.40	6.81
7	4	8.08	9.72	10.70	10.92	11.27	10.46	10.85	10.63	9.79	7.60
7	5	6.68	9.21	11.56	11.06	10.42	10.39	11.17	9.94	11.15	8.41
7	6	7.75	9.33	11.22	9.37	10.18	10.72	11.03	10.39	10.95	9.06
7	7	7.17	9.81	10.34	10.13	10.02	10.23	10.30	11.23	10.85	9.92
7	8	9.11	9.89	10.56	9.87	9.72	9.49	9.72	10.16	10.38	11.11
7	9	18.14	10.80	10.51	8.04	8.35	8.04	7.69	7.89	8.08	12.47
8	0	11.45	9.76	7.91	7.82	7.70	8.53	8.98	10.36	11.77	15.71
8	1	10.77	11.72	11.19	10.86	10.28	10.39	10.26	9.74	8.47	6.31
8	2	10.25	11.87	11.14	10.90	10.99	10.40	9.76	9.84	8.87	5.99
8	3	10.41	12.12	10.83	11.06	10.49	10.45	10.03	9.08	9.60	5.93
8	4	9.72	11.44	9.94	10.32	11.15	10.20	10.58	9.61	9.96	7.08
8	5	9.29	10.60	10.35	10.18	10.28	10.45	10.26	10.50	10.39	7.70
8	6	9.10	10.47	9.45	10.02	10.55	10.56	10.58	10.28	10.19	8.80
8	7	7.77	8.95	9.97	9.47	9.74	10.80	10.24	10.66	11.06	11.33
8	8	7.92	9.96	9.45	9.21	8.86	9.23	10.28	11.11	11.24	12.75
8	9	14.47	11.83	8.92	9.06	8.76	8.84	8.57	8.45	9.13	11.96
9	0	12.44	8.08	7.63	8.10	8.64	9.95	8.91	10.20	11.56	14.50
9	1	10.36	10.80	11.19	11.13	10.29	10.83	9.64	10.29	9.29	6.19
9	2	9.78	11.18	11.39	10.90	10.80	10.82	10.14	10.39	8.29	6.32
9	3	10.86	11.87	10.16	10.43	10.92	12.25	9.50	8.78	9.07	6.17
9	4	9.91	10.27	11.42	10.53	10.14	11.69	10.20	9.48	9.29	7.08
9	5	8.89	10.94	10.21	10.18	9.88	10.51	11.04	10.46	10.08	7.82
9	6	9.55	10.35	10.02	9.22	10.02	11.02	10.53	9.81	10.78	8.69
9	7	9.38	9.11	9.47	9.56	10.01	11.60	10.37	10.32	10.86	9.32
9	8	9.15	8.63	9.16	10.26	10.17	10.36	9.81	9.87	10.47	12.12
9	9	12.95	9.21	8.52	8.76	8.32	9.65	8.38	9.42	10.63	14.17
10	0	14.61	10.74	8.82	8.52	8.01	7.55	7.73	8.21	11.06	14.75
10	1	11.55	11.05	9.91	10.27	10.07	9.81	10.25	9.46	10.69	6.93
10	2	11.04	10.23	10.15	10.57	10.67	10.00	9.58	11.12	10.67	5.97
10	3	9.05	10.24	10.49	11.33	10.51	10.96	11.31	9.85	10.30	5.95
10	4	8.18	9.60	10.15	11.06	10.43	11.92	10.78	9.99	11.06	6.83
10	5	7.61	9.00	10.41	9.99	10.96	11.42	10.82	11.34	11.04	7.41
10	6	7.95	9.22	10.02	10.84	11.09	10.66	10.51	10.06	10.60	9.06
10	7	8.06	9.69	10.36	10.26	10.59	9.65	10.30	9.75	11.49	9.84
10	8	9.58	10.86	10.34	9.02	8.68	9.02	9.82	9.87	11.13	11.69
10	9	17.44	10.82	8.59	8.51	7.72	8.24	7.32	8.47	10.26	12.62

Table 2: Decile by grade examination of goodness-of-fit for 2007 CSAP read student growth percentiles using a linear parameterization

2006	2006	Observed Percentages of SGP in Given Percentile Range									
Grade	SS Decile	0 to 9	10 to 19	20 to 29	30 to 39	40 to 49	50 to 59	60 to 69	70 to 79	80 to 89	90 to 99
4	0	9.87	11.08	10.30	10.17	8.99	9.49	10.25	9.98	10.11	9.75
4	1	9.99	11.06	9.41	10.28	8.80	10.74	9.82	9.84	9.91	10.15
4	2	9.93	10.30	9.99	9.59	8.99	10.12	10.72	9.72	10.09	10.55
4	3	9.97	11.92	10.33	10.14	8.71	9.70	9.57	10.49	9.38	9.78
4	4	10.75	10.59	10.05	9.58	9.12	9.69	9.98	9.91	10.64	9.70
4	5	9.14	11.33	10.09	10.42	9.77	10.17	9.68	9.73	9.83	9.83
4	6	10.16	10.99	10.01	9.80	9.10	10.10	10.42	9.65	9.74	10.03
4	7	10.10	10.68	9.64	9.97	8.33	10.01	10.43	10.16	10.45	10.22
4	8	10.04	11.28	10.31	10.00	9.36	9.91	9.72	9.98	9.70	9.70
4	9	10.11	10.74	9.96	10.01	8.80	10.03	10.07	9.90	10.20	10.18
5	0	10.21	9.74	10.12	9.84	10.55	9.67	10.08	10.06	10.17	9.57
5	1	9.67	10.27	10.48	9.69	11.25	9.93	9.13	10.10	10.05	9.43
5	2	10.50	9.52	9.52	9.58	11.25	10.31	10.16	9.94	9.64	9.58
5	3	10.35	10.29	9.89	10.82	10.60	9.42	10.31	9.46	10.27	8.58
5	4	9.79	10.19	10.27	9.84	10.55	10.06	10.23	9.61	10.59	8.87
5	5	9.83	9.48	10.19	9.68	11.46	10.47	9.70	10.77	9.25	9.16
5	6	10.15	10.59	9.68	10.33	10.33	9.68	10.23	9.97	9.97	9.06
5	7	10.13	9.47	10.73	9.72	11.51	9.77	9.73	9.66	10.03	9.25
5	8	10.24	10.26	9.24	10.02	10.81	10.41	10.13	9.81	10.17	8.90
5	9	9.60	9.76	10.43	9.97	10.87	9.97	10.16	10.26	9.93	9.05
6	0	10.14	9.93	9.86	9.71	9.88	10.85	9.91	10.25	10.27	9.20
6	1	10.79	9.67	10.12	10.19	9.74	10.68	10.21	9.83	9.42	9.35
6	2	9.75	10.16	9.53	10.10	10.10	10.76	9.88	9.87	10.31	9.53
6	3	9.41	10.40	10.42	9.94	9.81	11.62	9.41	9.86	10.53	8.60
6	4	10.59	9.82	10.07	9.71	9.80	10.65	10.66	10.01	9.43	9.26
6	5	10.15	9.92	9.60	9.92	9.94	10.72	10.36	10.38	10.00	9.03
6	6	10.20	10.54	9.48	10.45	10.01	11.27	9.44	9.76	9.63	9.21
6	7	9.63	9.36	10.55	9.89	10.49	10.28	10.40	9.87	10.51	9.01
6	8	10.09	10.51	9.43	9.95	9.78	11.97	9.76	9.87	10.07	8.59
6	9	10.16	10.08	10.42	9.93	10.10	10.58	9.94	9.87	9.66	9.27
7	0	10.04	9.78	10.68	10.10	9.93	9.55	10.25	10.23	10.36	9.10
7	1	10.43	10.17	11.12	9.98	9.46	9.61	10.47	9.51	9.79	9.46
7	2	10.80	10.21	10.21	9.98	10.34	11.10	9.41	10.27	9.19	8.48
7	3	9.68	9.81	11.47	10.42	9.94	9.47	10.08	9.64	10.15	9.35
7	4	9.78	10.00	11.03	9.93	9.85	10.06	10.04	9.81	10.45	9.05
7	5	10.08	10.20	10.72	9.75	9.83	9.58	10.26	10.60	10.16	8.82
7	6	9.27	9.61	11.32	9.98	10.14	10.45	10.21	9.97	9.95	9.10
7	7	11.00	9.82	10.71	9.80	10.46	9.90	9.45	9.57	10.05	9.26
7	8	9.81	9.83	10.60	10.49	9.68	10.04	10.11	10.49	9.81	9.12
7	9	9.83	10.17	11.13	9.71	9.94	9.96	10.15	9.96	10.04	9.12
8	0	10.30	10.21	10.83	9.63	9.88	9.54	9.76	10.01	10.32	9.52
8	1	10.63	9.60	10.82	9.66	10.09	10.20	10.29	9.38	10.05	9.27
8	2	9.85	9.76	10.96	10.23	9.84	9.93	9.97	10.13	10.43	8.90
8	3	9.93	10.09	11.28	10.09	10.18	9.87	9.71	9.64	9.87	9.33
8	4	9.97	10.16	10.77	9.62	9.64	10.58	9.93	10.48	9.24	9.60
8	5	9.80	9.31	11.03	10.11	9.96	9.57	10.40	9.53	10.78	9.51
8	6	10.31	9.89	10.59	9.91	10.08	10.33	9.63	10.35	10.31	8.59
8	7	10.28	10.22	11.16	10.15	9.50	10.02	10.15	10.04	9.43	9.06
8	8	10.12	10.29	11.07	9.37	10.62	9.54	9.81	9.79	10.31	9.08
8	9	9.62	9.83	10.80	10.31	9.78	10.19	9.93	10.10	10.16	9.28
9	0	11.79	9.25	10.10	10.05	9.86	9.97	9.41	11.19	9.17	9.21
9	1	10.34	10.81	10.12	9.51	10.01	10.46	10.03	10.49	9.45	8.78
9	2	10.57	9.57	9.33	10.42	10.19	9.50	10.14	11.17	10.02	9.08
9	3	10.85	10.10	10.23	10.21	9.18	10.14	9.31	10.96	9.65	9.37
9	4	9.78	9.81	10.67	10.19	9.70	9.57	10.17	11.03	10.34	8.73
9	5	10.26	9.81	9.34	9.67	10.85	10.24	10.08	10.66	10.05	9.04
9	6	9.47	10.11	10.31	9.90	9.94	10.19	9.61	11.24	10.00	9.24
9	7	10.15	9.78	9.53	9.94	9.93	9.98	9.98	10.97	10.06	9.68
9	8	10.23	9.74	10.04	10.41	10.27	9.70	10.02	11.00	9.74	8.85
9	9	10.00	9.75	10.03	9.73	10.02	9.96	10.09	10.82	10.40	9.21
10	0	11.18	10.34	10.33	9.89	9.63	9.95	9.55	9.83	10.01	9.29
10	1	9.96	10.23	10.35	9.53	9.80	9.30	10.73	10.33	10.07	9.70
10	2	11.12	9.74	10.99	10.20	9.90	10.10	10.24	9.70	9.46	8.54
10	3	10.20	9.70	11.22	9.58	10.32	10.12	9.08	9.64	10.72	9.46
10	4	9.27	10.19	10.49	9.95	9.87	9.79	10.78	10.40	9.70	9.57
10	5	9.62	9.72	10.59	10.10	9.39	9.88	10.49	9.49	10.81	9.92
10	6	10.09	9.39	10.52	9.87	9.85	10.28	9.79	10.44	10.18	9.59
10	7	9.58	9.42	11.60	10.38	10.27	9.46	10.23	10.01	9.80	9.25
10	8	10.84	10.74	10.88	9.45	9.86	9.61	10.32	9.61	9.64	9.07
10	9	9.56	9.64	10.69	10.12	9.95	9.77	9.81	10.61	10.53	9.33

Table 3: Decile by grade examination of goodness-of-fit for 2007 CSAP math student growth percentiles using a B-spline parameterization with 4 knots

2006	2006	Observed Percentages of SGP in Given Percentile Range									
Grade	SS Decile	0 to 9	10 to 19	20 to 29	30 to 39	40 to 49	50 to 59	60 to 69	70 to 79	80 to 89	90 to 99
4	0	10.04	11.41	9.02	9.90	10.10	10.08	10.55	9.20	9.92	9.77
4	1	9.73	10.86	9.61	10.29	10.12	9.73	11.33	8.56	10.27	9.50
4	2	9.82	11.04	8.75	9.88	10.09	9.78	11.18	9.43	9.86	10.19
4	3	10.56	11.16	8.36	9.57	9.42	10.19	11.04	9.24	9.83	10.62
4	4	10.00	10.29	9.35	10.60	9.95	10.74	11.16	8.13	9.93	9.85
4	5	10.50	10.26	8.58	9.71	10.30	9.78	10.48	9.43	10.36	10.61
4	6	10.12	10.79	9.27	9.86	9.94	9.80	11.30	9.07	9.86	10.00
4	7	9.54	11.61	9.13	10.19	10.13	9.80	10.90	8.95	10.25	9.50
4	8	9.85	11.84	8.92	9.99	10.12	10.22	11.01	8.94	9.61	9.49
4	9	10.04	10.59	9.14	10.00	9.80	9.78	10.97	9.10	10.31	10.27
5	0	10.12	10.35	10.88	10.01	10.22	9.56	9.52	9.77	9.95	9.62
5	1	10.09	9.23	11.75	9.74	9.38	10.26	10.22	9.59	10.56	9.18
5	2	10.59	9.95	10.04	9.77	10.14	10.14	10.76	10.01	9.83	8.76
5	3	9.16	9.64	11.92	10.13	10.77	9.74	9.57	10.38	9.96	8.73
5	4	10.29	10.09	10.97	9.83	10.01	9.68	9.81	10.42	10.26	8.65
5	5	9.77	10.47	10.16	10.31	9.75	10.41	10.06	9.94	9.69	9.44
5	6	10.39	10.11	10.80	9.79	9.55	10.25	9.99	9.69	9.91	9.51
5	7	9.94	10.07	11.21	9.94	10.28	9.47	10.05	10.13	10.11	8.79
5	8	10.59	9.39	10.89	9.97	10.18	9.97	9.93	10.08	10.08	8.91
5	9	9.70	10.41	10.67	10.16	9.80	10.18	9.92	9.98	10.02	9.15
6	0	10.68	10.02	10.60	9.87	10.47	9.92	9.62	9.66	9.64	9.51
6	1	11.12	9.66	9.18	9.60	10.18	9.52	10.68	10.37	10.95	8.75
6	2	12.11	9.50	9.75	10.14	9.32	9.60	10.16	9.89	10.00	9.54
6	3	11.53	10.83	10.07	9.48	9.93	9.76	9.56	10.30	9.85	8.69
6	4	9.67	9.33	10.54	10.29	10.58	10.09	10.03	9.14	10.67	9.67
6	5	10.75	10.30	9.85	10.28	10.46	10.18	9.62	9.85	9.71	8.99
6	6	10.62	10.00	10.40	9.98	9.68	10.08	10.17	10.66	9.47	8.94
6	7	11.20	9.80	9.41	10.15	10.13	9.86	9.91	10.11	10.32	9.11
6	8	11.58	9.67	9.92	9.71	9.56	9.83	10.32	9.48	10.68	9.23
6	9	10.39	10.17	9.92	10.05	10.57	10.03	10.15	10.07	9.62	9.03
7	0	9.63	10.32	11.09	10.06	9.48	9.97	10.27	9.89	9.93	9.35
7	1	10.19	10.16	9.98	9.26	10.21	10.00	10.12	10.23	10.44	9.39
7	2	10.79	10.13	10.19	10.26	9.72	10.24	10.06	9.65	10.26	8.70
7	3	9.96	9.59	11.61	9.92	9.43	10.22	9.55	10.38	10.14	9.20
7	4	9.83	10.41	10.55	10.50	10.41	9.46	10.32	9.79	9.48	9.26
7	5	9.04	10.29	11.61	9.93	10.04	9.52	10.54	9.39	10.27	9.37
7	6	11.18	9.49	10.79	9.35	9.66	10.91	9.49	10.12	9.97	9.04
7	7	10.06	10.15	10.49	10.09	9.85	10.07	9.92	10.72	10.00	8.65
7	8	10.04	9.97	11.19	10.06	10.21	9.79	10.18	9.78	9.97	8.82
7	9	9.42	10.08	11.26	9.48	10.20	10.23	9.98	9.90	10.06	9.38
8	0	9.87	10.25	10.25	10.46	9.79	9.87	9.70	10.25	10.48	9.07
8	1	9.97	9.23	9.44	11.45	9.55	10.65	10.48	10.10	9.84	9.29
8	2	9.60	9.97	9.84	10.77	10.25	9.93	10.01	10.12	10.17	9.35
8	3	10.19	10.35	10.44	11.04	9.56	10.28	9.70	9.83	9.52	9.07
8	4	9.79	9.83	9.77	10.50	10.12	9.81	9.95	10.16	10.24	9.81
8	5	10.49	9.48	10.10	10.86	9.84	9.91	9.98	9.96	10.60	8.78
8	6	10.54	10.54	9.79	10.87	10.47	9.58	10.43	9.58	9.52	8.69
8	7	10.30	9.22	10.37	11.22	9.86	10.37	9.57	9.76	9.91	9.42
8	8	10.22	10.08	10.00	10.83	9.36	9.95	10.22	9.95	10.35	9.04
8	9	10.22	10.08	9.70	11.31	10.20	10.14	9.93	10.22	9.54	8.67
9	0	10.56	10.14	9.99	9.89	10.25	9.87	10.62	10.29	9.61	8.79
9	1	9.87	10.30	9.99	10.24	9.27	9.78	10.83	10.24	10.28	9.19
9	2	10.63	9.81	9.93	9.23	10.48	9.74	10.61	10.16	10.11	9.31
9	3	10.85	10.23	10.05	9.21	10.24	10.89	10.78	9.19	9.61	8.93
9	4	10.30	9.40	10.57	10.24	10.01	10.10	10.49	10.30	9.64	8.96
9	5	9.32	10.38	9.71	10.36	9.77	9.15	11.16	10.40	10.66	9.10
9	6	10.30	10.24	10.30	9.26	9.67	9.91	11.24	9.46	10.16	9.46
9	7	10.67	9.80	9.76	9.87	10.20	10.18	11.15	9.89	9.65	8.85
9	8	10.16	9.72	10.35	10.59	10.20	9.40	10.67	9.98	9.90	9.03
9	9	10.16	10.32	9.99	9.42	9.72	10.22	10.91	9.86	10.63	8.76
10	0	10.15	10.41	9.70	10.13	10.29	10.07	10.19	9.95	9.82	9.30
10	1	9.91	10.24	9.89	10.12	10.08	10.55	9.47	9.98	9.65	10.10
10	2	11.47	9.99	9.61	9.89	10.01	9.49	9.47	9.79	10.16	10.12
10	3	10.37	10.53	10.14	10.39	9.89	9.49	10.37	9.39	9.91	9.53
10	4	9.76	10.34	10.00	9.68	10.06	10.11	9.68	9.91	9.96	10.51
10	5	10.30	10.08	9.48	9.63	9.91	10.45	10.10	10.82	9.67	9.56
10	6	10.76	9.86	10.03	10.31	10.48	10.06	10.12	9.10	9.21	10.06
10	7	10.38	10.34	10.15	10.36	10.18	9.63	9.86	9.59	9.95	9.55
10	8	10.91	9.94	10.70	9.51	9.13	9.70	10.00	9.86	10.28	9.96
10	9	9.99	10.51	9.66	9.87	10.30	10.51	9.87	9.99	9.66	9.66

Table 4: Decile by grade examination of goodness-of-fit for 2007 CSAP read student growth percentiles using a B-spline parameterization with 4 knots

Growth Projection Figures

The growth projection figures that follow present 9 different growth scenarios (3 in reading and 3 in math) for students with 3rd grade state assessment scores at the performance level cutpoints. The growth scenarios depict consecutive (i.e., year-over-year) growth quantified in terms of student growth percentiles. That is, the figures present what, for example, 50th percentile growth leads to versus 60th percentile growth. The figures are intended to aid stakeholders in better understanding the range of student growth and what different growth rates lead to in terms of student achievement relative to state designated performance performance levels.

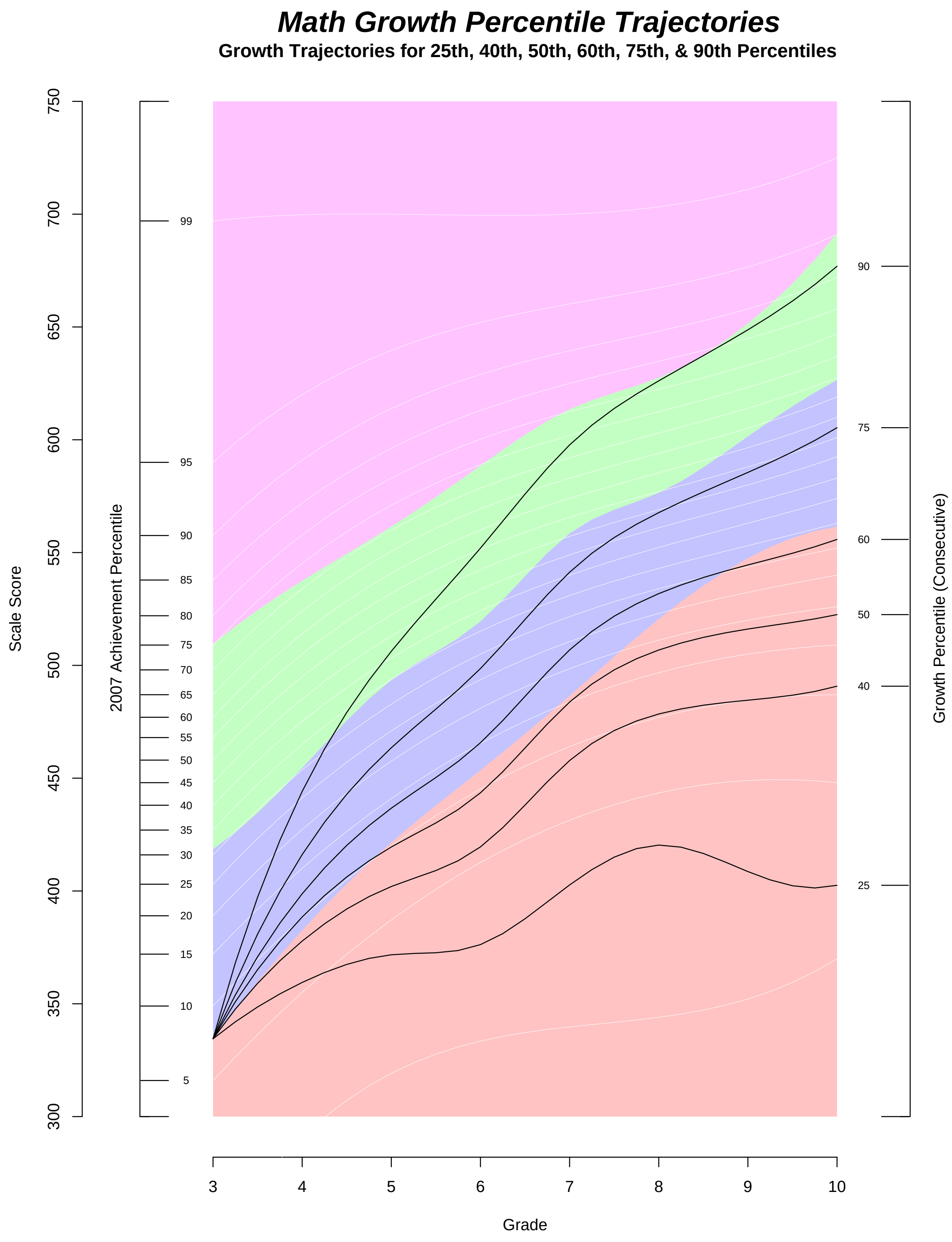


Figure 4: Growth chart depicting future math achievement conditional upon consecutive 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning the third grade at the unsatisfactory/partially proficient cutpoint

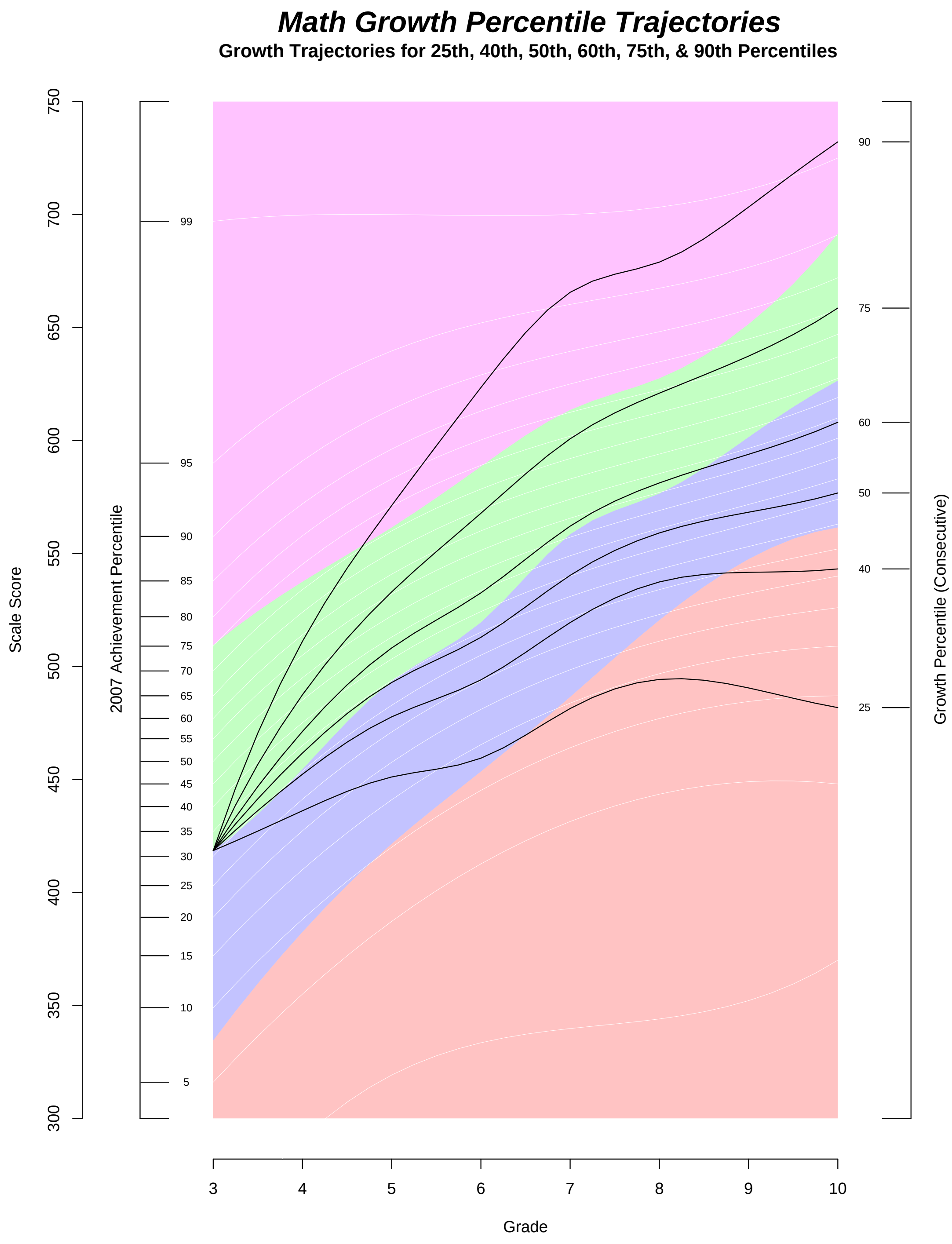


Figure 5: Growth chart depicting future math achievement conditional upon consecutive 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning the third grade at the partially proficient/proficient cutpoint

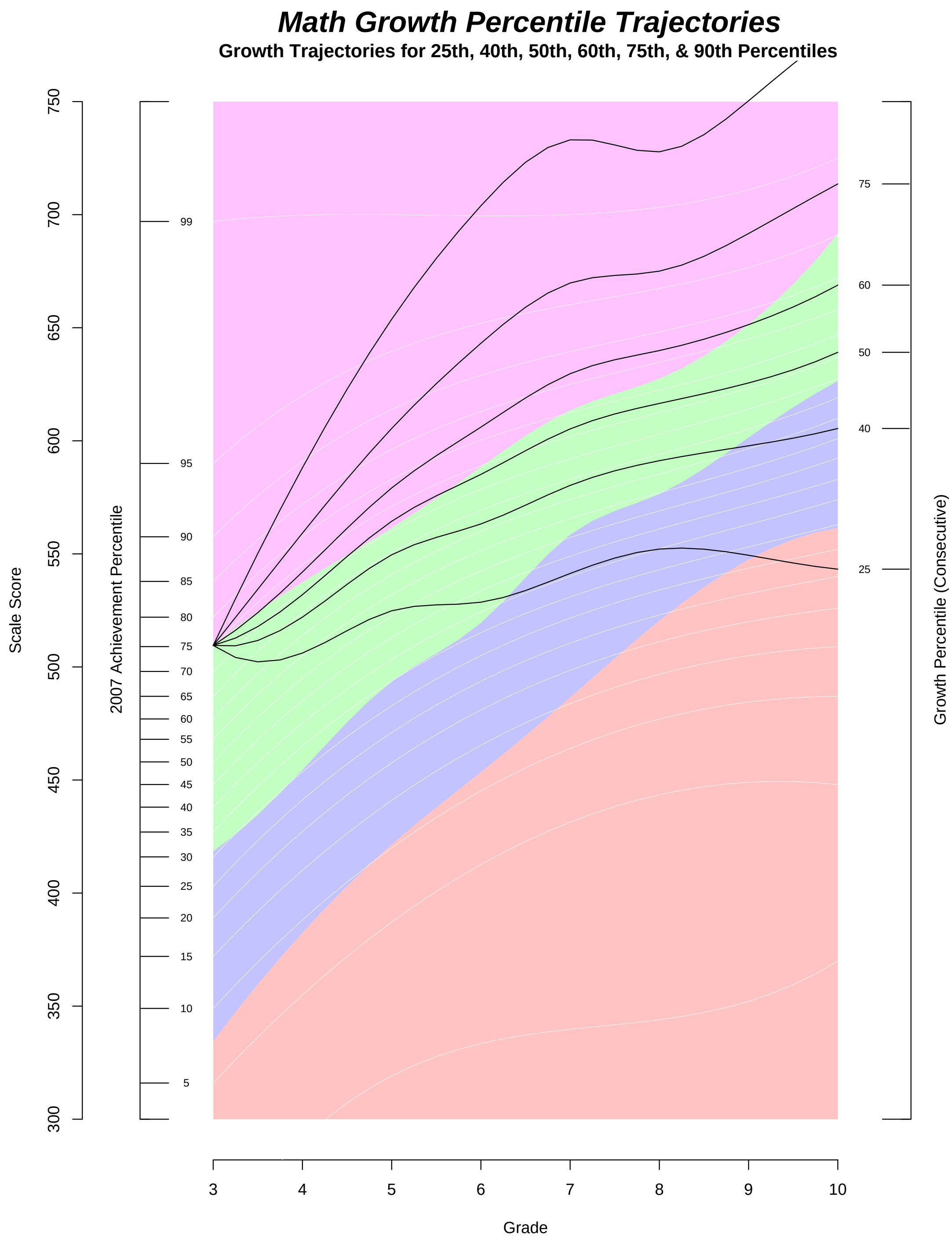


Figure 6: Growth chart depicting future math achievement conditional upon consecutive 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning the third grade at the proficient/advanced cutpoint

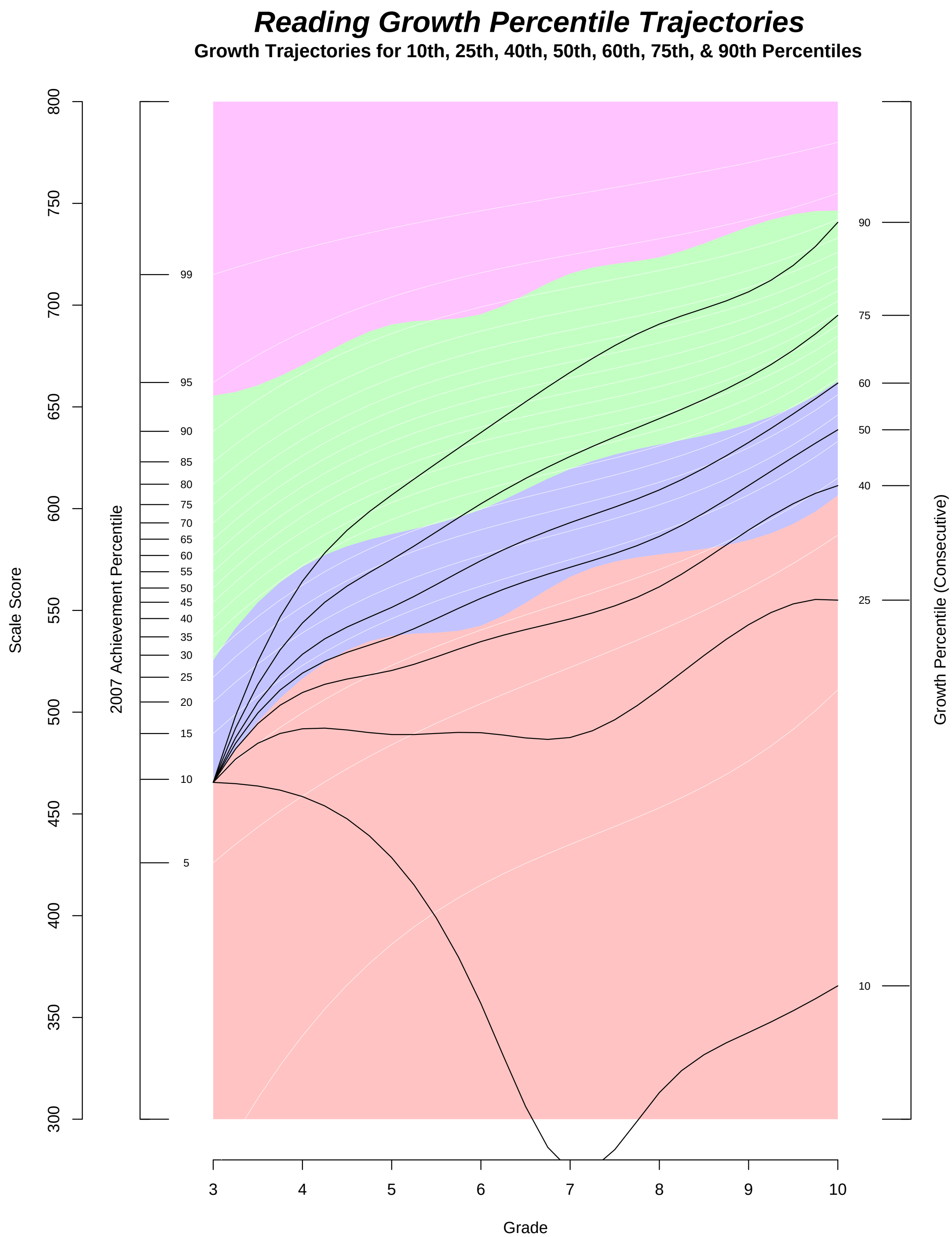


Figure 7: Growth chart depicting future reading achievement conditional upon consecutive 10th, 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning the third grade at the unsatisfactory/-partially proficient cutpoint

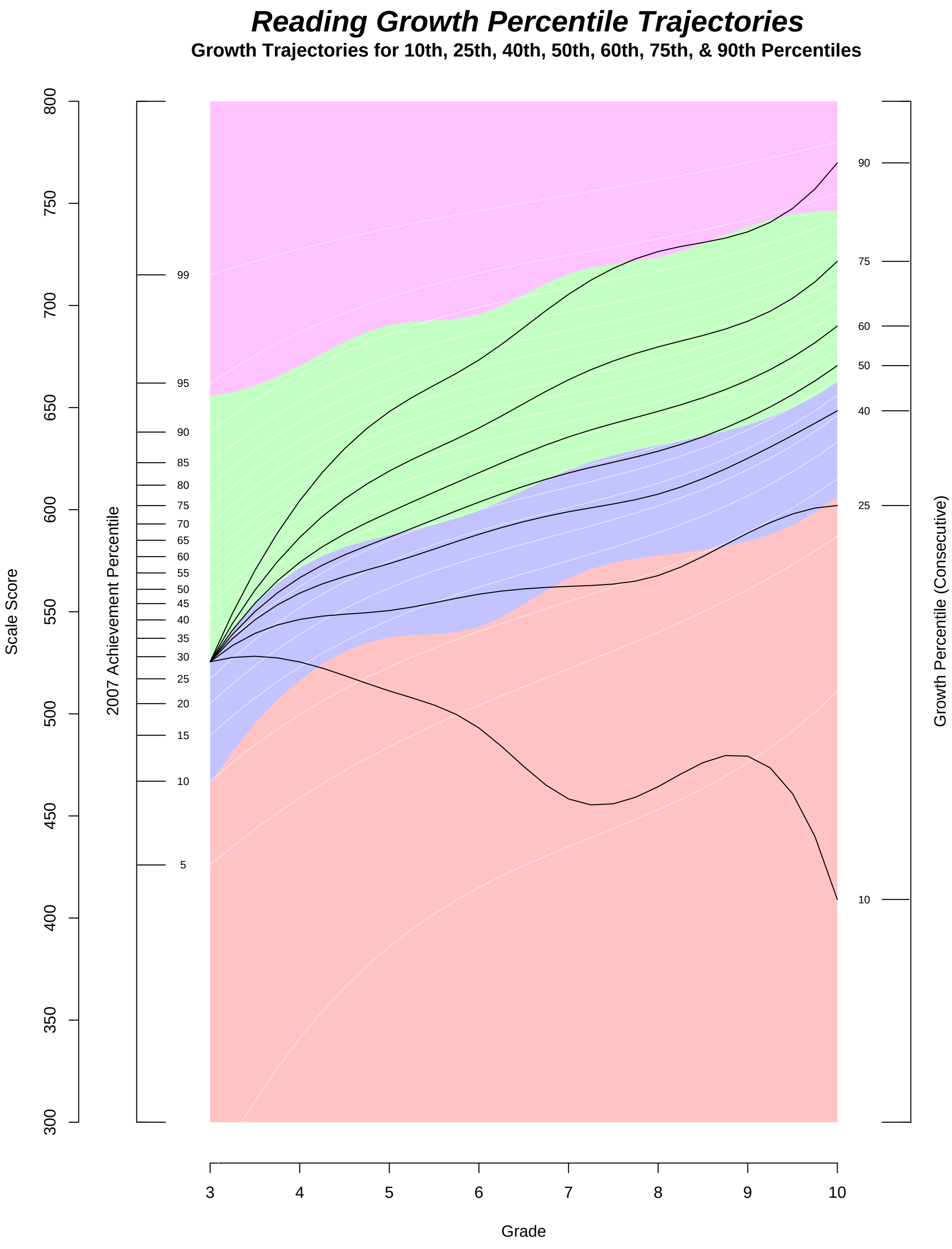


Figure 8: Growth chart depicting future reading achievement conditional upon consecutive 10th, 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning the third grade at the partially proficient/proficient cutpoint

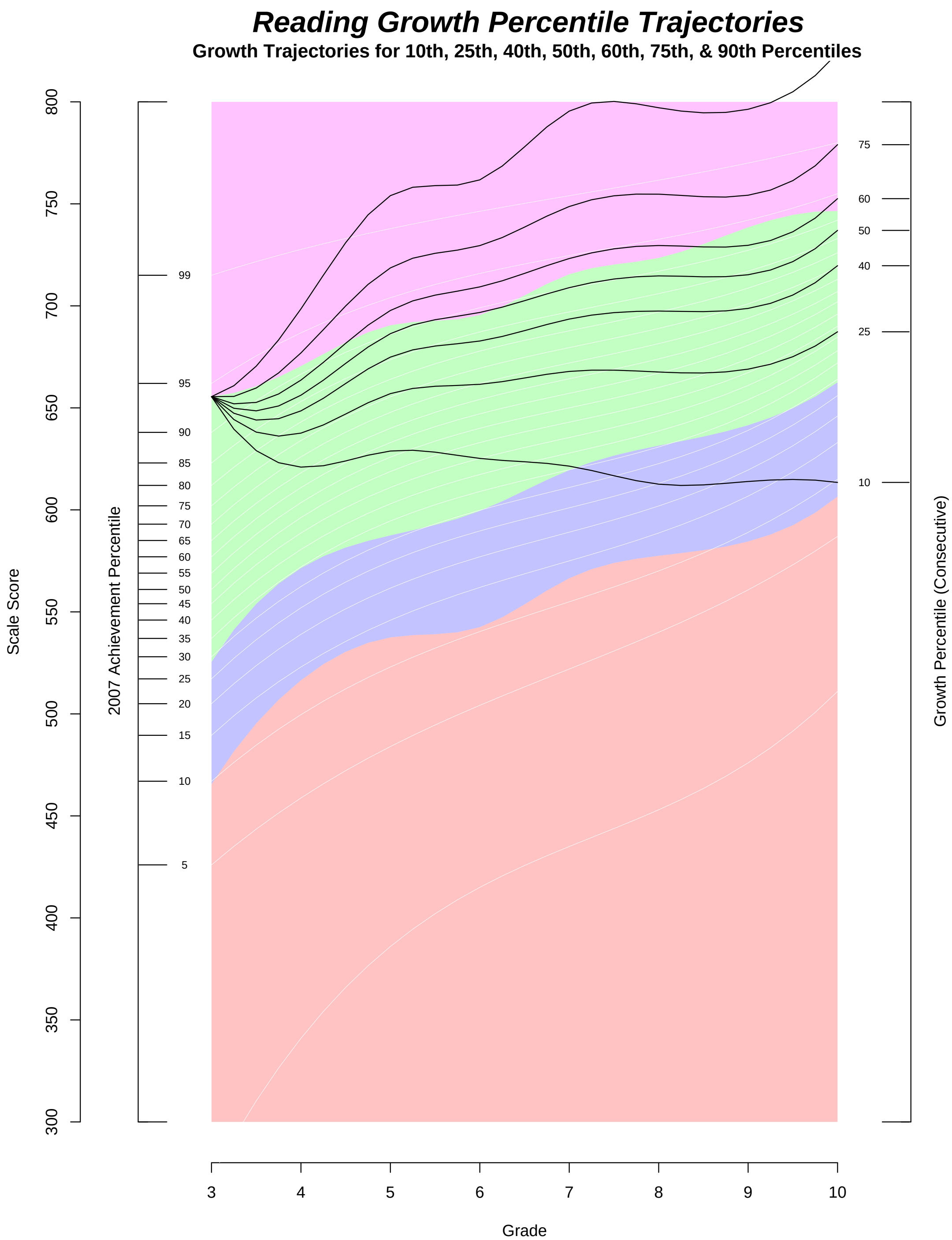


Figure 9: Growth chart depicting future reading achievement conditional upon consecutive 10th, 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning the third grade at the proficient/advanced cutpoint