On Defining a Year's Growth

Damian W. Betebenner
Center for Assessment
Dover, NH
DBetebenner@nciea.org

Yi Shang
ERME/Lynch School of Education
Boston, College
shangyi@gmail.com

Abstract

The November 2005 announcement of the Growth Model Pilot Program by Secretary of Education Spellings permitting states to use growth models as a means for compliance with NCLB AYP requirements greatly increased interest in measures of student growth. To ensure compliance with NCLB's 2014 universal proficiency mandate, the majority of growth models approved for the Pilot Program utilize projection-based methodologies requiring students to be on track to proficiency. Absent from growth-to-proficiency models are analyses demonstrating the attainability of these individually mandated proficiency goals. In essence, the growth-to-standard methodology establishes the criterion of "a year's growth" without consideration of the normative dimension of annual student growth. To overcome this shortcoming, we introduce student growth percentiles to provide a normative context for understanding student change. We discuss how growth percentiles can be used by any state to define a year's growth realistically for every student.

Background

Following the 2001 reauthorization of the *Elementary and Secondary Education Act of 1965*, commonly known as the *No Child Left Behind Act* (NCLB), states were required to implement large scale testing of all students to an extent never before seen in the United States (No Child Left Behind Act of 2001 (NCLB), 2002). Starting with the 2005-2006 school year, NCLB required states to test students in reading and mathematics from grades 3 through 8 and at least once in grades 10 through 12. In addition, beginning with the 2007-2008 school year, states must assess students in science at least once in elementary, middle and high school. As a consequence, states find themselves buried in assessment data with mandates from state and federal policy makers to utilize the data to improve the quality of education.

Accountability systems constructed according to federal adequate yearly progress (AYP) requirements currently rely upon annual measurement of student achievement to make judgments about school quality. Since their adoption, such *status measures* have been the focus of persistent criticism (Linn, 2003; Linn, Baker, & Betebenner, 2002). Status measures, though appropriate for making judgments about the achievement level of students at a school for a given year, are inappropriate for judgments about educational *effectiveness*. In this regard, status measures are blind to the possibility of low achieving students attending effective schools. It is this possibility that has led some critics of

NCLB to label its accountability provisions as unfair and misguided and to demand the use of growth analyses as a better means of auditing the quality of schools.

A fundamental premise associated with the use of student growth for school accountability is that "good" schools bring about student growth in excess of that found at "bad" schools. Students attending such schools—commonly referred to as highly effective/ineffective schools—tend to demonstrate extraordinary growth that is causally attributed to the school or teachers within the school. The inherent believability of this premise is at the heart of current enthusiasm to incorporate growth models into state accountability systems. It is not surprising that the November 2005 announcement by Secretary of Education Spellings for the Growth Model Pilot Program permitting states to use growth model results as a means for compliance with NCLB achievement mandates was met with great enthusiasm by states. (Spellings, 2005).

The primary thrust of growth analyses over the last decade has been to determine, using sophisticated statistical techniques, the amount of student progress/growth that can be justifiably attributed to the school or teacher—that is, to disentangle current *aggregate* level achievement from effectiveness (Braun, 2005; Rubin, Stuart, & Zanutto, 2004; Ballou, Sanders, & Wright, 2004; Raudenbush, 2004). Such analyses, often called *value-added* analyses, attempt to estimate the teacher/school contribution to student achievement. This contribution to student achievement, called the *school* or *teacher effect*, purports to quantify the impact on achievement that this school or teacher would have, on average, upon similar students assigned to them for instruction. Clearly, such analyses lend themselves to accountability systems that hold schools or teachers responsible for student achievement. Despite this utility, however, such analyses fail to address one of the fundamental questions stakeholders have concerning student growth: How much growth did a student make?

This paper/presentation returns to the task of quantifying individual student growth. To do so, we introduce a normative quantification of student change which we call *student growth percentiles*. These quantities, derived using quantile regression techniques, are easily interpretable descriptive statistics that permit growth comparisons between all students regardless of the scale students are measured on. In addition, growth percentiles can be criterion referenced *vis-á-vis* current growth-to-standard methodologies in order to establish qualifications of what is *enough* growth. The purpose of this paper/presentation is to introduce student growth percentiles and to demonstrate, using state assessment data, how current discussions of student growth lack a normative dimension necessary to set challenging yet attainable individual achievement goals. We assert that the establishment of a normative basis for student growth eliminates a number of the problems of incorporating growth into accountability systems by providing needed insight to various stakeholders by answering the "simple" question of how much a student progressed.

Student Growth

Measuring individual change is among the most important topics in educational measurement. Educators are ultimately concerned with individual learning, and the very notion of learning implies growth and change (Willett, 1988). Individual academic growth is the only intended outcome of building effective schools with competent teachers and sound curriculum. To come up with any evaluation of schools, teachers, or educational programs and policies, one has to start with measuring student growth.

With many achievement testing programs using ordinal achievement levels as a major component of their score reports, perhaps the most intuitive way of indicating student growth is to describe their change of performance levels over time. The current AYP and safe harbor provisions under NCLB are aggregated indicators of school status and growth that are based on individual students' change of achievement levels. This way of describing growth can be easily understood by all stakeholders given the convenient definitions of the labels such as non-proficient, proficient, and advanced. Performance levels, however, form a very coarse measurement of student academic status, and any measurement of growth based on it would be very imprecise. Students often make substantial progress while remaining in the same performance level. Such types of growth would be inexcusably neglected if we evaluate growth based only on performance levels. On the other hand, very small changes across cut scores of the performance levels would be captured and magnified. To measure change with more precision requires a more refined scale.

Another straightforward choice of assessing student growth is to use the difference or gain scores. They can be derived from the same test that is administered at different times or from vertically-equated tests that address different grade levels. In the former case, bias, reliability, and other relevant coefficients are estimated for difference scores as evidence for its inherent deficiency (see for example Lord, 1956 and Willett, 1988). Besides the psychometric properties, this type of difference score has a critical drawback: Using the same test at different time points makes it impossible for the test to target the current level of student knowledge and skills. Much is taught and learned during an academic year, and with the same test being used in the beginning and end, student learning would not be measured with appropriate precision. Over longer time periods, use of the difference score to measure student growth would be even more lacking.

With vertical scaling, difference scores can be produced from tests that are specifically designed for students' current learning levels. But this method of measuring growth is also very problematic due to inherent problems with vertical scaling. Of greatest concern is that difference scores from vertical scales are not interpretable or comparable. For a 20 point gain in vertically-equated math test scores, it is not possible to conclude exactly how much a student has grown in her mathematical ability or whether this growth is adequate, because these kinds of judgments depend heavily on whether the score scale of the test is completely interpretable (which they are usually not), where the student started on the scale, and how much did students who started at the same place normally grow.

Doran (2004) reviews the normal educational growth model which is based on Normal Curve Equivalent (NCE) scores. In this model, it is considered "adequate" for a student to maintain or exceed the same position in the distribution over time. In other words, a gain score of 0 computed from the NCE scores is considered as expected growth. Doran (2004) points out three major problems with this method. First, it requires different amount of growth to maintain one's position in the distribution depending on where one starts, therefore growth measured in this way is not comparable. Second, the gain score of

NCE is obtained using only two data points which does not provide enough information about growth trend over time. And third, growth measured in NCE gain scores provide no information as to whether students are growing toward an acceptable standard of academic performance.

Doran (2004), uses the terms of "adequate growth", "normal growth", and "expected growth" interchangeably. In the following, to facilitate discussions, we refer these terms separately. Adequate growth, as Doran (2004) argues, is a concept that implies adequacy toward some externally defined standard. That is, adequate growth, like adequate achievement (i.e., proficiency), refers to an underlying criterion. Following current growth-to-standard approaches, it could be defined as the rate of growth necessary for a student to reach proficiency in three years. By contrast, normal growth relies more upon norm-referencing than an external criterion. Whether a student's growth is judged normal or typical is dependent on how much other students have grown, especially those students sharing similar backgrounds. We suggest that, in quantifying normal growth, the conditional distribution of current scores given past scores be used instead of the unconditional distributions of scores from the whole population, because the former forms a more relevant reference group to be compared with. If, among those who started from the same place, a student falls within a reasonable interval around the median in the conditional distribution of current scores, we would consider this growth normal. By contrast, expected growth is an expectation that takes into account both normal growth and adequate growth. It is defined for specific individuals according to the magnitude of growth of peers the historical patterns of growth of the student herself, and how far this student is from the goal of proficiency. As Linn (2003) argues, "current levels of performance and past gains provide a context for judging future gains".

In this paper, we propose using the percentiles of the conditional distribution to quantify student growth. Figure 1 depicts 4 conditional distributions of grade 7 scores conditioned upon grade 6 reading scores along with the unconditional distribution of grade 7 reading scores. The percentile of a student's current score within their corresponding conditional distribution, which we call the *student's growth percentile*, gives the probability of a student obtaining their current score taking account of prior achievement. That is:

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

Whereas unconditional percentiles normatively quantify achievement, conditional percentiles normatively quantify growth. In this case, because past scores are used solely for conditioning purposes using quantile regression techniques, the assessments do not have to be vertically linked. One of the major advantages of using growth percentiles to measure change over most other methods with the same purpose—the validity and interpretation of this quantity does not require any kind of vertical scaling.[1]

Student growth percentiles are comparable across all students because they represent probability statements explicating how (extra)ordinary a students current status is taking

[1] See Betebenner & Shang (2007) for details associated student growth percentile estimation and the quantile regression techniques employed to construct the conditional distributions.

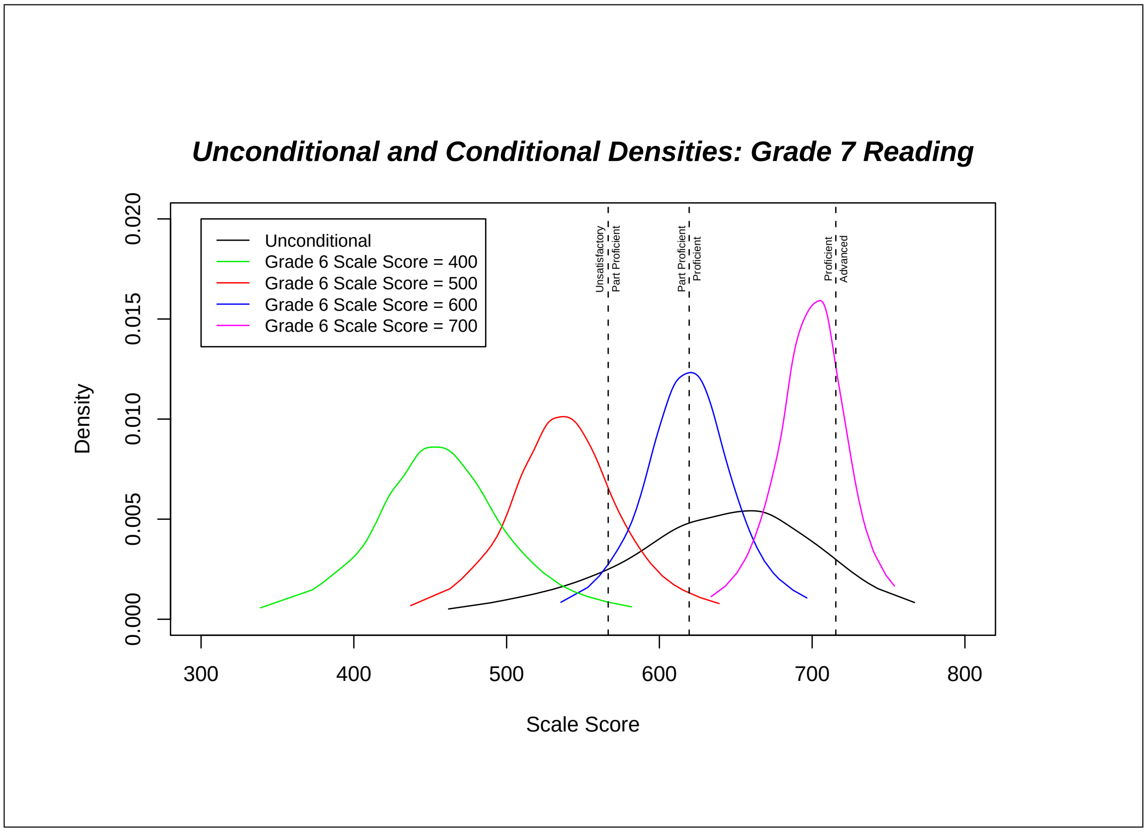

Figure 1: Conditional and unconditional distributions of Grade 7 reading scores conditioning upon Grade 6 reading data.

account of their prior status. For example, an "advanced" student with a growth percentile of 50 is, in terms of achievement, well above a non-proficient student demonstrating 95th percentile growth. However, given how much more exemplary the low achieving student's growth is, the achievement gap between them is almost certainly closing. We can also compare a student with herself. Suppose a student is 50th percentile in year 1 conditioning on year 0, and is 20th percentile in year 2 conditioning on year 1. Even if her scale scores show that she is still growing, we know she is actually slowing down and might need some additional attention.

Lastly and most importantly, the important concepts of adequate growth, normal growth and expected growth can be conveniently and properly quantified in terms of student growth percentiles. All three terms have different quantitative meanings for those with different backgrounds. Put differently, the same amount of growth measured by scale scores reflects differing amounts of effort, motivation, and advancement of learning for students who started high and those starting low on the scale. By putting students into different reference groups, growth percentiles make it possible to understand and compare student progress.

How much Growth is Enough?

We have briefly described how to normatively quantify growth in terms of student growth percentiles. We now turn our attention and demonstrate how student growth percentiles can form the basis of discussions/standard setting regarding what is adequate' growth. This is the process of going from a normative standard to a criterion referenced standard of growth and is, in many ways, analogous to current normative and criterion referenced understandings of achievement. In order to do so, we begin by unpacking the imbroglio of terminology currently being used to "inform" student growth.

In a discussion of growth, NCLB, and vertical scaling, Yen (2007) provides a set of questions parents, teachers, and administrators ask in terms of quantifying student growth:

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 student 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 in the previous section, student growth percentiles provide a normative and, we would argue, more informative way to address these questions than an absolute metric on which to interpret progress.

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. In this situation, parents are not interested in an absolute measure of growth, but instead in a normative description locating that 4 inch increase alongside the height increases of similar children. Examining this height increase relative to the increases of similar children permits one to diagnose how (ab)normal such an increase is.

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

Supposing scales did exist in education similar to height/weight scales that permitted the calculation of absolute measures of annual academic growth for students, the answers parents would receive to questions such as, “How much did my child progress?”, would come as a number of scale score points—an answer that would 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*. In the next section we use growth percentiles as the basis for determining whether the growth demonstrated by a child qualifies as *a year’s growth* and, more generally, whether the amount of growth is (in)adequate.

Methodology for Defining “A Year’s Growth”

The concept of a year’s growth appears rudimentary. Its simplicity, however, belies a subtle complexity. To adequately address the notion of a year’s growth, *aspirational* growth must be distinguished from the *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: This is 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 in the Growth Model Pilot Program defines adequate growth as growth leading toward proficiency. Using the conditional densities of Figure 1, it is straightforward to calculate the growth percentile necessary for each student to reach proficiency. This threshold could then be used to distinguish adequate from inadequate growth. A benefit of using the growth percentile scale is that the threshold has a normative context that can be used to set criterion referenced aspirational goals that are reasonable.

Returning to the concept of a year's growth, we have found in discussions with numerous stakeholders *no* consensus over what the term actually refers to. In particular, some people use the term to denote *typical* or *average* growth—a child has demonstrated at least a year's growth if their growth is average or better. Others use the term in an aspirational sense implying that if a student makes a year's growth they are on track to reach a desirable achievement outcome. These different understandings represent a difference between a normative and a criterion referenced understanding of the term a year's growth.

Norm referenced definitions

Perhaps the simplest way to define a year's growth or, more generally, adequate growth using growth percentiles is to stipulate a fixed growth percentile threshold that each student is required to meet or exceed. For example, a 50th percentile threshold (i.e., current typical growth) could be used to distinguish adequate from inadequate growth. Given present circumstances, 50% of students would be expected to demonstrate adequate and 50% inadequate growth.

There are advantages to establishing year's growth thresholds in a normative fashion. If the growth threshold (i.e., target growth) is defined uniformly for each student (e.g., establishing target growth at the 50th growth percentile from baseline data), then there is probabilistic equivalence in terms of the difficulty of elevating each student to this growth target. If percentages of students at a school achieving target growth is reported, then the goal for each school is to get students to grow at a rate exceeding the baseline 50th percentile growth. This establishes a uniform goal for all students and schools. If students were randomly distributed to schools, then in the baseline year, 50% of the students in any given school would be expected to demonstrate target growth or better. As time passes, one wishes to see schools having higher and higher percentages of their students in this baseline category.

A disadvantage to setting target growth normatively is that it doesn't equalize chances for individuals to reach proficiency (or other achievement outcomes associated with state defined performance levels). Student growth based upon quantile regression can be used to establish growth standards in terms of performance levels. To do so requires investigating what growth percentiles are necessary for students to reach the different achievement/performance level outcomes. These growth percentiles goals can then be used to define growth adequacy thresholds.

Criterion referenced definitions

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. Equivalently, the lower one's current achievement the lower their chances must be for them to reach the desired target. Specifically, if an individual must demonstrate 90th percentile growth to reach a desired achievement target (e.g., proficiency) in the coming year, then their chances of reaching such an outcome are 0.1 (i.e., 10 percent).

Establishing criterion referenced growth thresholds requires achievement "forecasting" to be done. Instead of inferring that prior student growth is indicative of future student growth (e.g., linearly projecting student achievement into the future), predictions of future student achievement is contingent upon initial student status (where the student starts) and the rate of growth (i.e., growth percentile) the student maintains in the following years. That is, if student X, starting in grade G at achievement level A consistently grows at growth percentile GP, then based upon the most recent growth patterns of the state's children, student X's projected achievement in t years is PA. Looking at future student achievement this way, instead of saying, "Student X is projected to be not proficient in three years" one says that "Given that Student X starts at this point and grows over the coming three years at this rate, Student X is projected to be not proficient. However, if Student X grows at a rate exceeding GP, then they are projected to be proficient. This focuses conversations on "what it will take" instead of "where will the student be".

This is more easily understood with the assistance of a picture. Figures 2 to 10 (Pages 13 to 21) depict three growth scenarios each in math, reading, and writing for students beginning in grade three at each of the 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 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 previous 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 the growth standard setting procedure.

Consider Figure 2, the student represented begins in the 3rd grade at the threshold between unsatisfactory and partially proficient. Based upon the achievement percentiles, approximately 7 percent of the population of 3rd graders rate as unsatisfactory. This percentage increases dramtically up to grade 10 where nearly 35 percent of students perform at the unsatisfactory level. The black lines in the figure represent six different growth scenarios for the student based upon consecutive growth at a given growth percentile. 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. With the green region representing proficient, a student would need to demonstrate growth percentile in excess of 75 to reach proficiency showing how unlikely such a event currently is. Considering NCLB universal proficiency mandates, the growth necessary for non-proficient students to reach proficiency is likely unattainable for a large percentage of non-proficient students given current levels of growth. Of course, the reality of the present is not a blueprint for eternity. However, without a radical restructuring of math education for those non-proficient students, it seems highly unlikely that the achievement targets of NCLB will be realized.

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 in terms of growth percentiles, 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 discussions. Achievement mandates are stipulated based upon the moral imperative of high standards for all children. Given current progress of students, it is unlikely that sustained levels of growth necessary for these universal high 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 of doing it. Growth percentiles bring Kant's dictum to the fore when considering criterion referenced growth standards.

Conclusions

This paper/presentation addresses the task of defining what a year's growth is for each student. In doing so, we distinguish between actual growth and aspiration growth. To understand actual growth, we introduce quantile regression based growth percentiles as a means of understanding the normative dimension of student growth. Given this normative foundation, we examine properties of the growth percentiles and demonstrate that current criterion referenced growth-to-standard approaches to modeling growth impose unrealistic growth expectations for a large number of low achieving students. However, using growth percentiles in conjunction with state performance standards, it is possible to define aspirational growth for each student in a manner that sets reasonable and challenging achievement goals for all students.

References

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment for teachers. *Journal of Educational and Behavioral Statistics*, *29*(1), 37–65.

Betebenner, D. W., & Shang, Y. (2007, April). *Reference growth charts for educational outcomes.* (Paper presented at the 2007 AERA Annual Conference, Chicago, IL. Retrieved May 12th, 2007 from http://monticello.bc.edu/Betebenner)

Braun, H. I. (2005). *Using student progres to evaluate teachers: A primer on value-added models* (Tech. Rep.). Princeton, New Jersey: Educational Testing Service.

Doran, H. C. (2004, June). *Value-added models and adequate yearly progress: Combining growth and adequacy in a standards-based environment.* (Paper Presented at the 2004 Annual CCSSO Large-Scale Assessment Conference)

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

Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher*, *31*(6), 3–16.

Lord, F. M. (1956). The measurement of growth. *Educational and psychological measurement*, *16*, 421–437.

No Child Left Behind Act of 2001, Public Law No. 107-110, 115 Stat. 1425.

Raudenbush, S. W. (2004). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics*, *29*(1), 121–129.

Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and Behavioral Statistics*, *29*(1), 103–116.

Spellings, M. (2005, Nov). *Secretary Spellings announces growth model pilot* [Press Release]. U.S. Department of Education. (Retrieved August 7, 2006 from http://www.ed.gov/news/pressreleases/2005/11/1182005.html)

Willett, J. B. (1988). Questions and answers in the measurement of change. *Review of Reserach in Education*, *15*, 345–422.

Yen, W. M. (2007). Vertical scaling and no child left behind. In N. J. Dorans & P. W. Holland (Eds.), *Linking and aligning scores and scales* (pp. 273–283). New York: Springer.

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 meant 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.

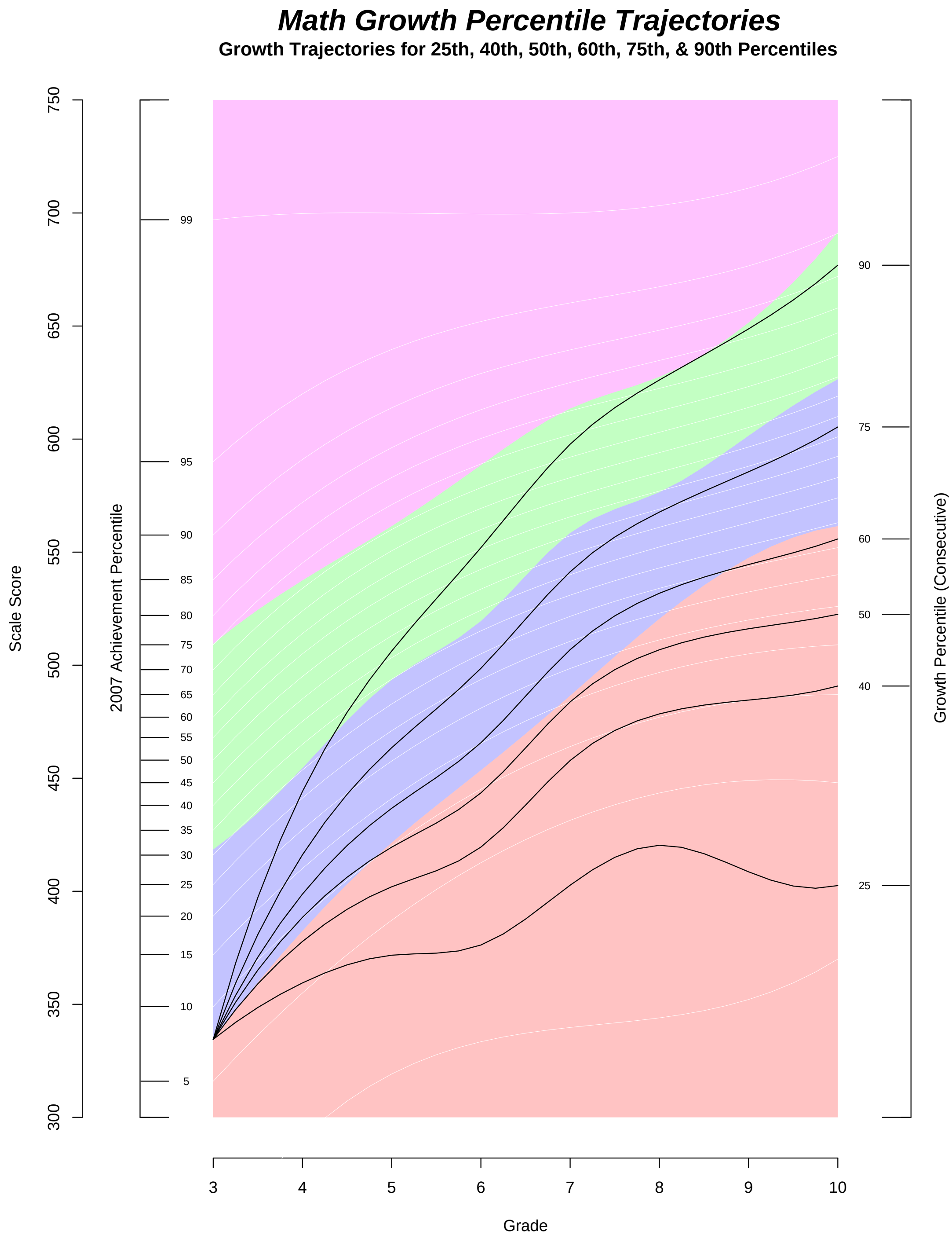

Figure 2: Growth chart depicting future math achievement conditional upon consecutive 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning in the third grade at the unsatisfactory/partially proficient cutpoint

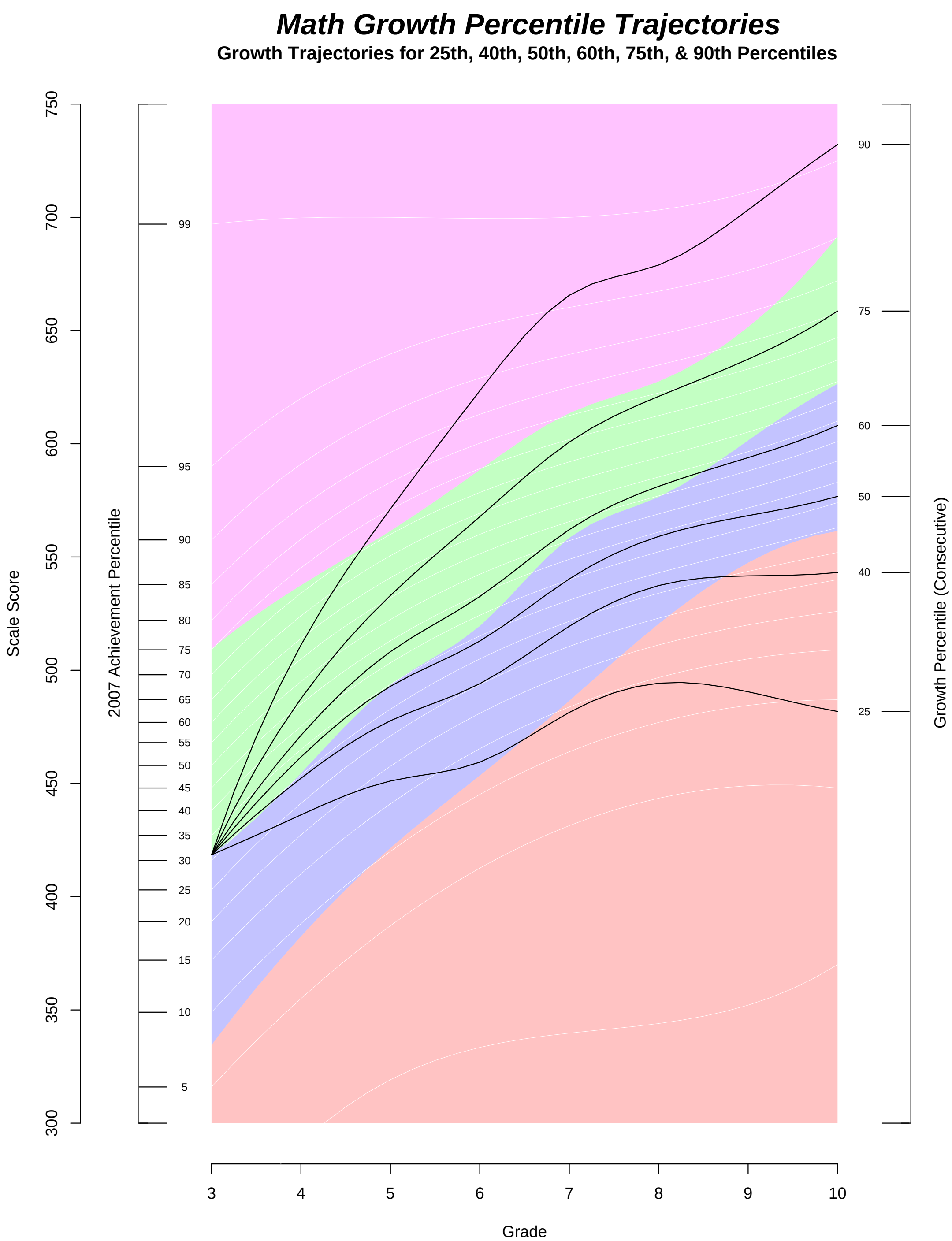

Figure 3: Growth chart depicting future math achievement conditional upon consecutive 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning at the third grade at the partially proficient/proficient cutpoint

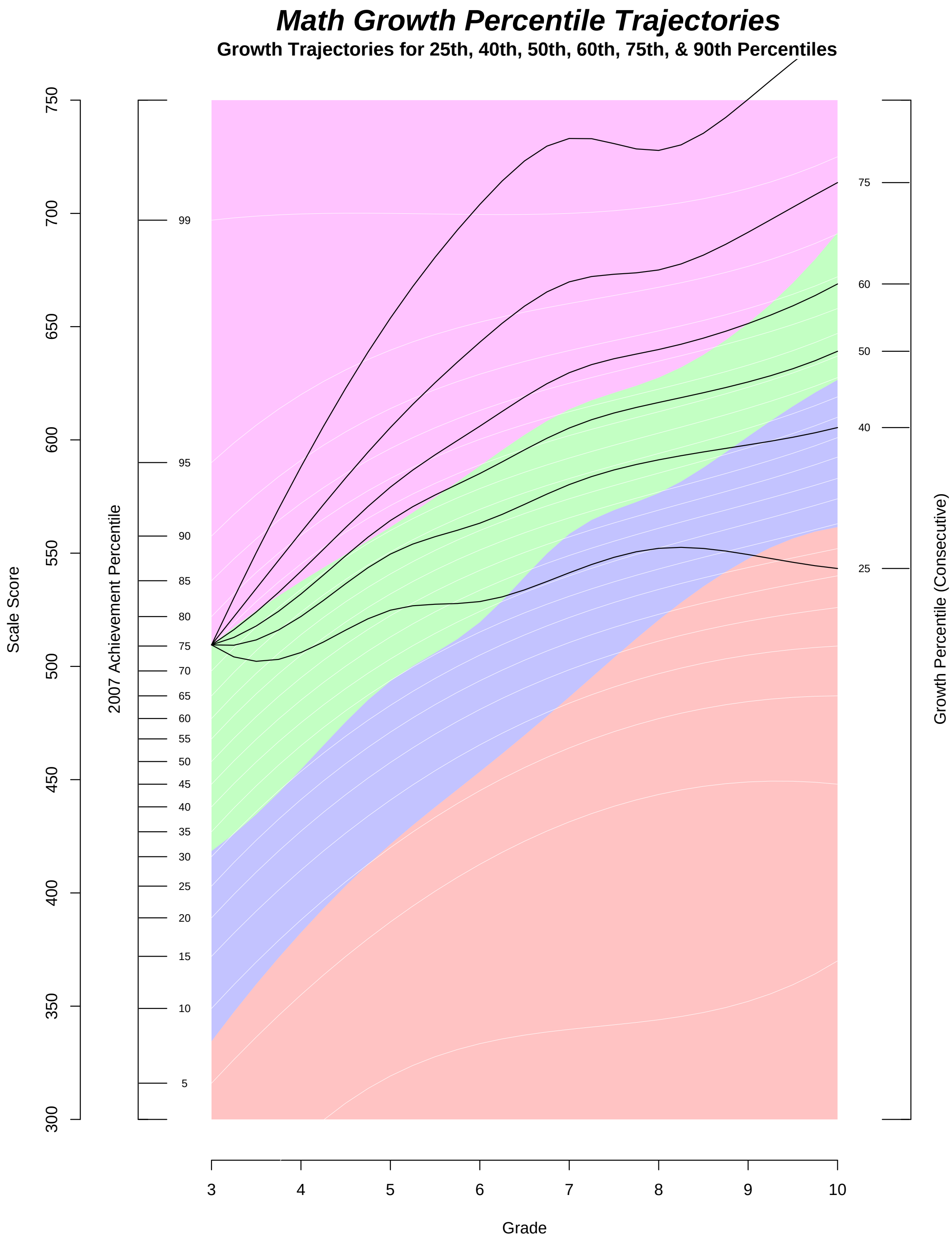

Figure 4: Growth chart depicting future math achievement conditional upon consecutive 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning at the third grade at the proficient/advanced cutpoint

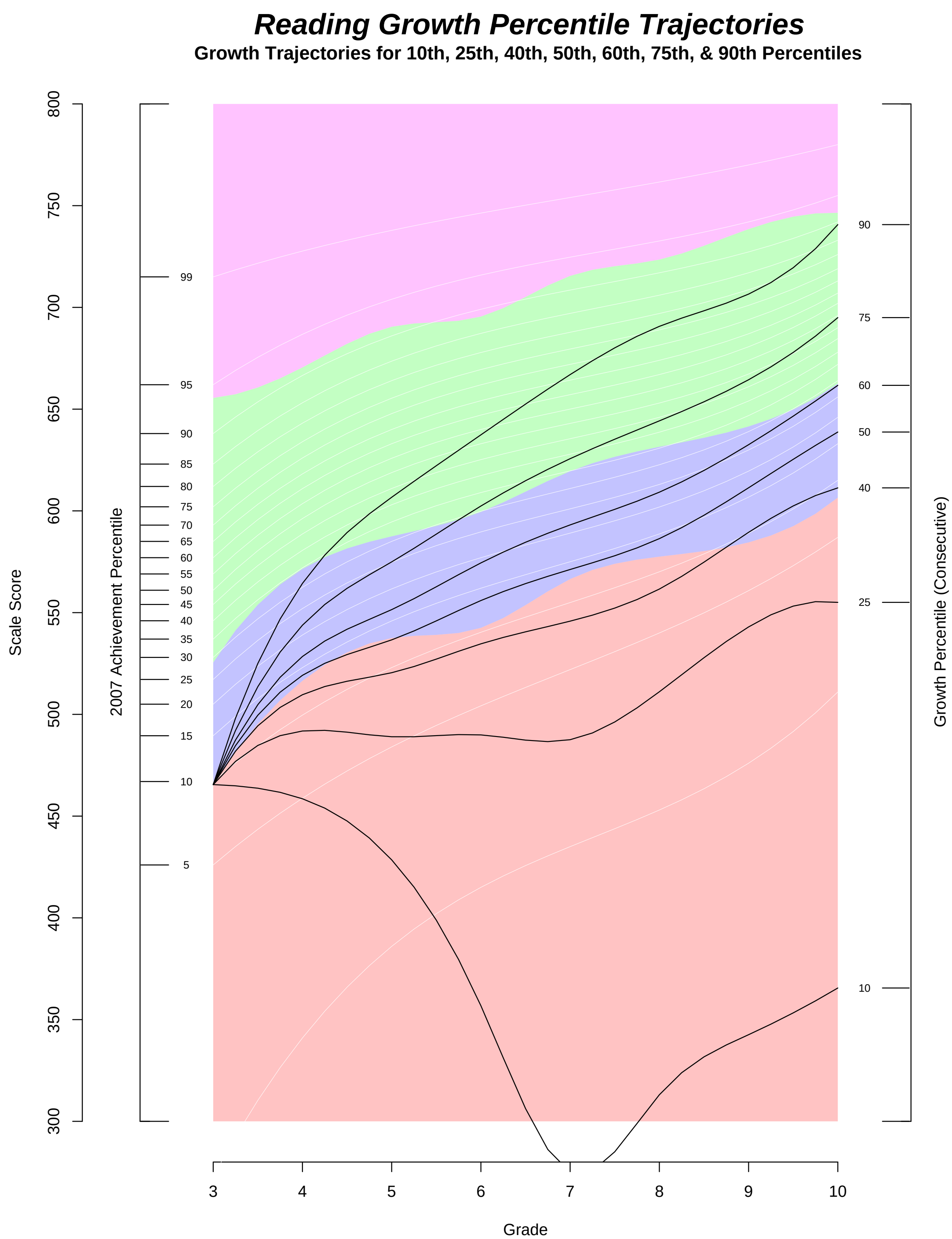

Figure 5: Growth chart depicting future reading achievement conditional upon consecutive 10th, 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning in the third grade at the unsatisfactory/partially proficient cutpoint

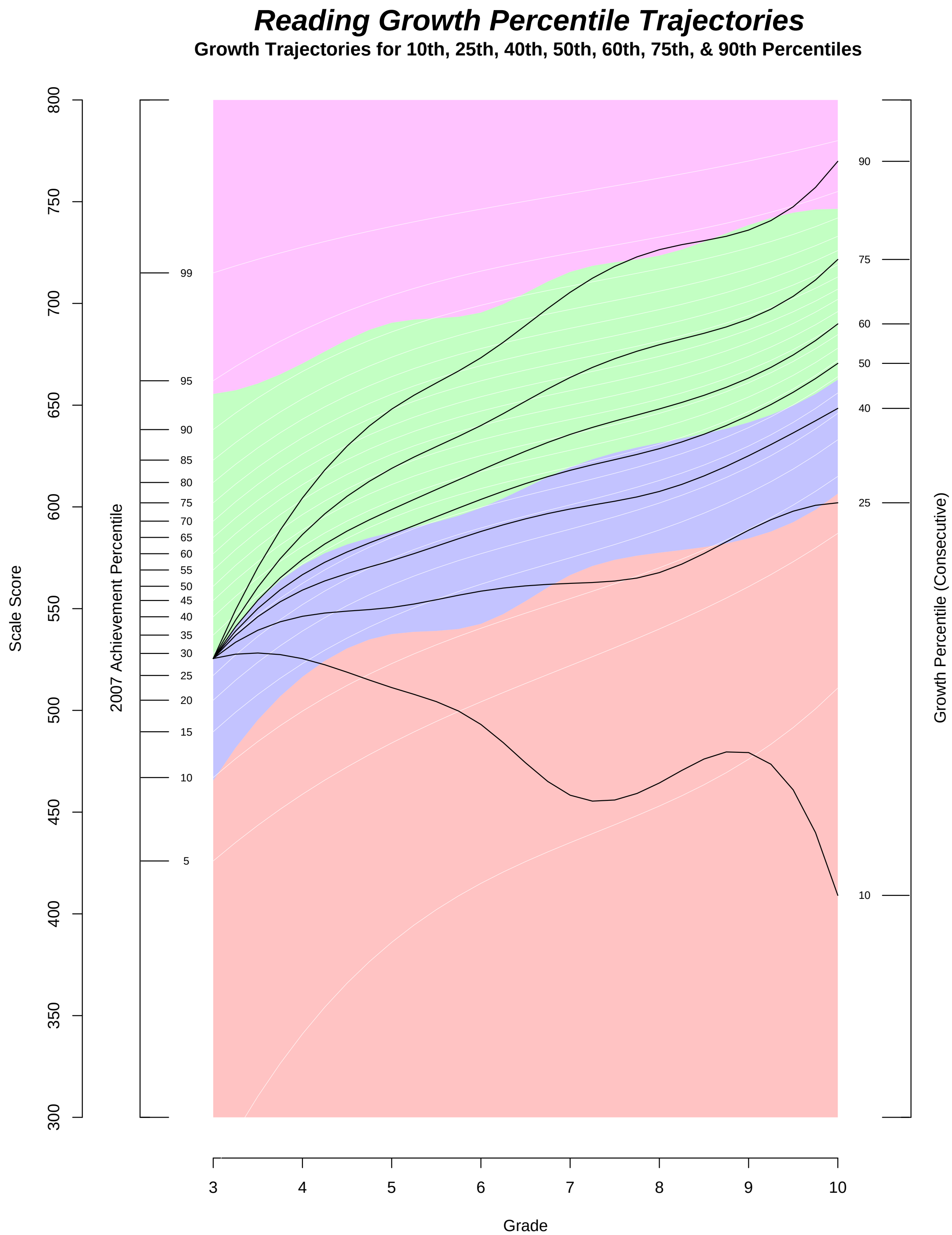

Figure 6: Growth chart depicting future reading achievement conditional upon consecutive 10th, 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning at the third grade at the partially proficient/proficient cutpoint

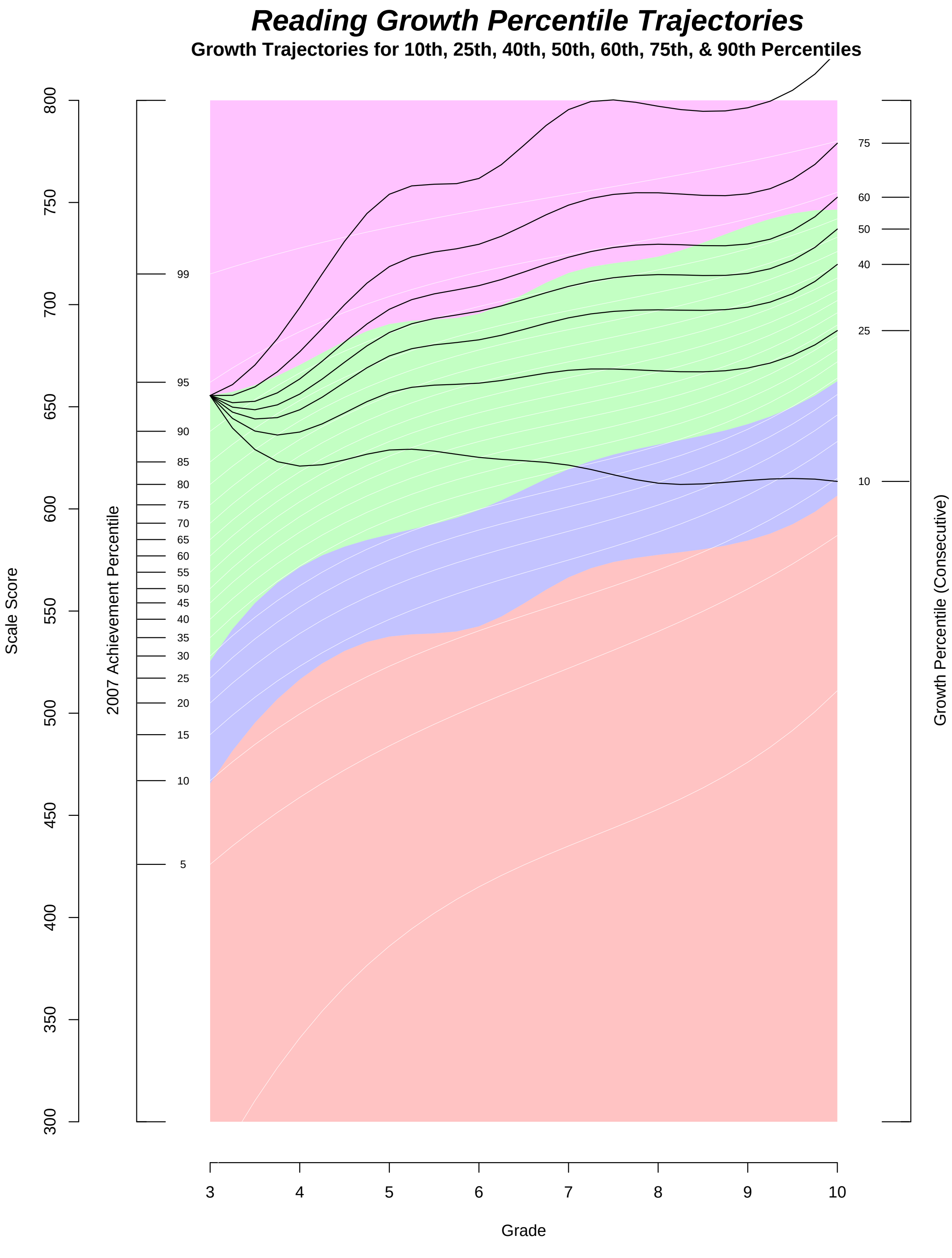

Figure 7: Growth chart depicting future reading achievement conditional upon consecutive 10th, 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning at the third grade at the proficient/advanced cutpoint

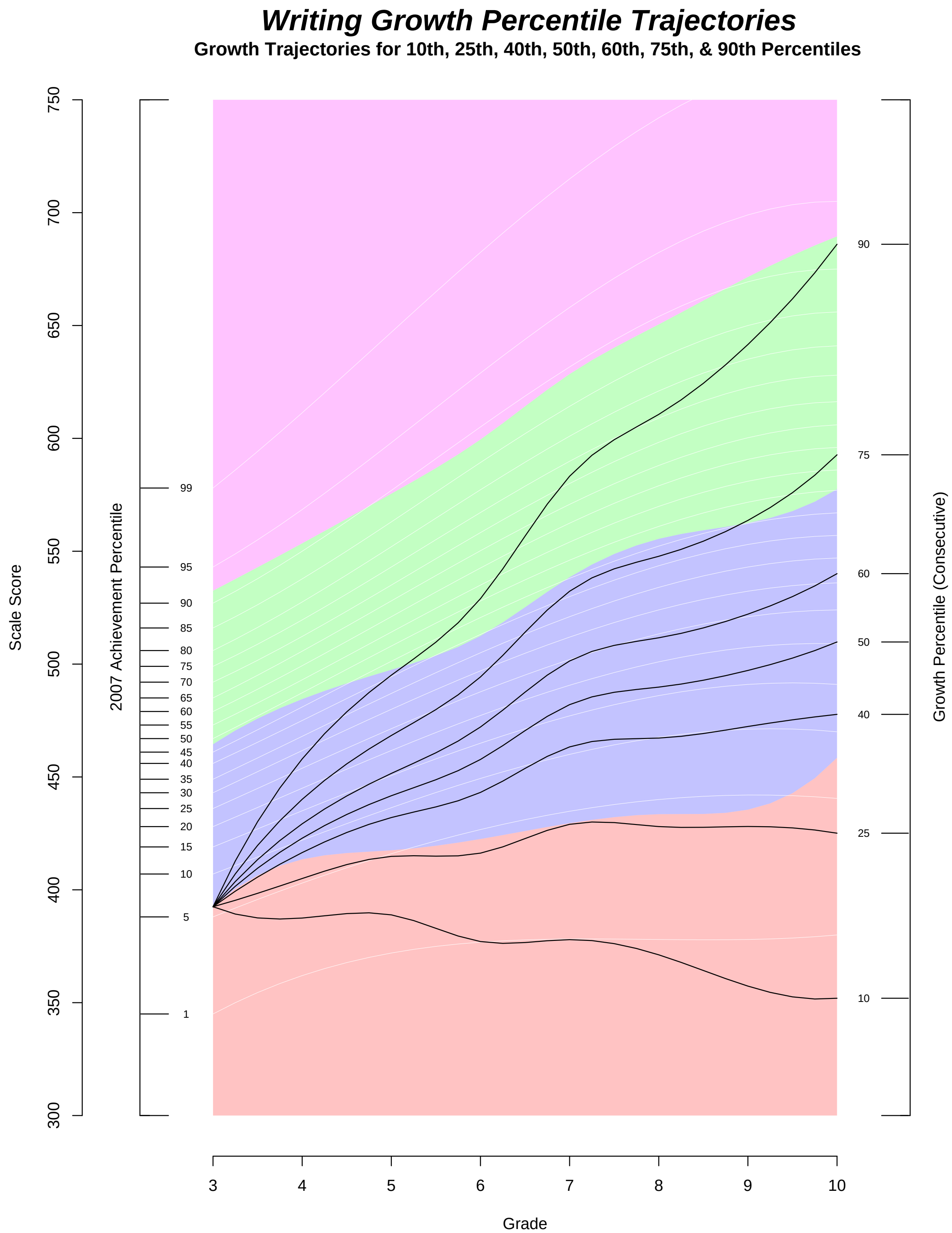

Figure 8: Growth chart depicting future writing achievement conditional upon consecutive 10th, 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning in the third grade at the unsatisfactory/partially proficient cutpoint

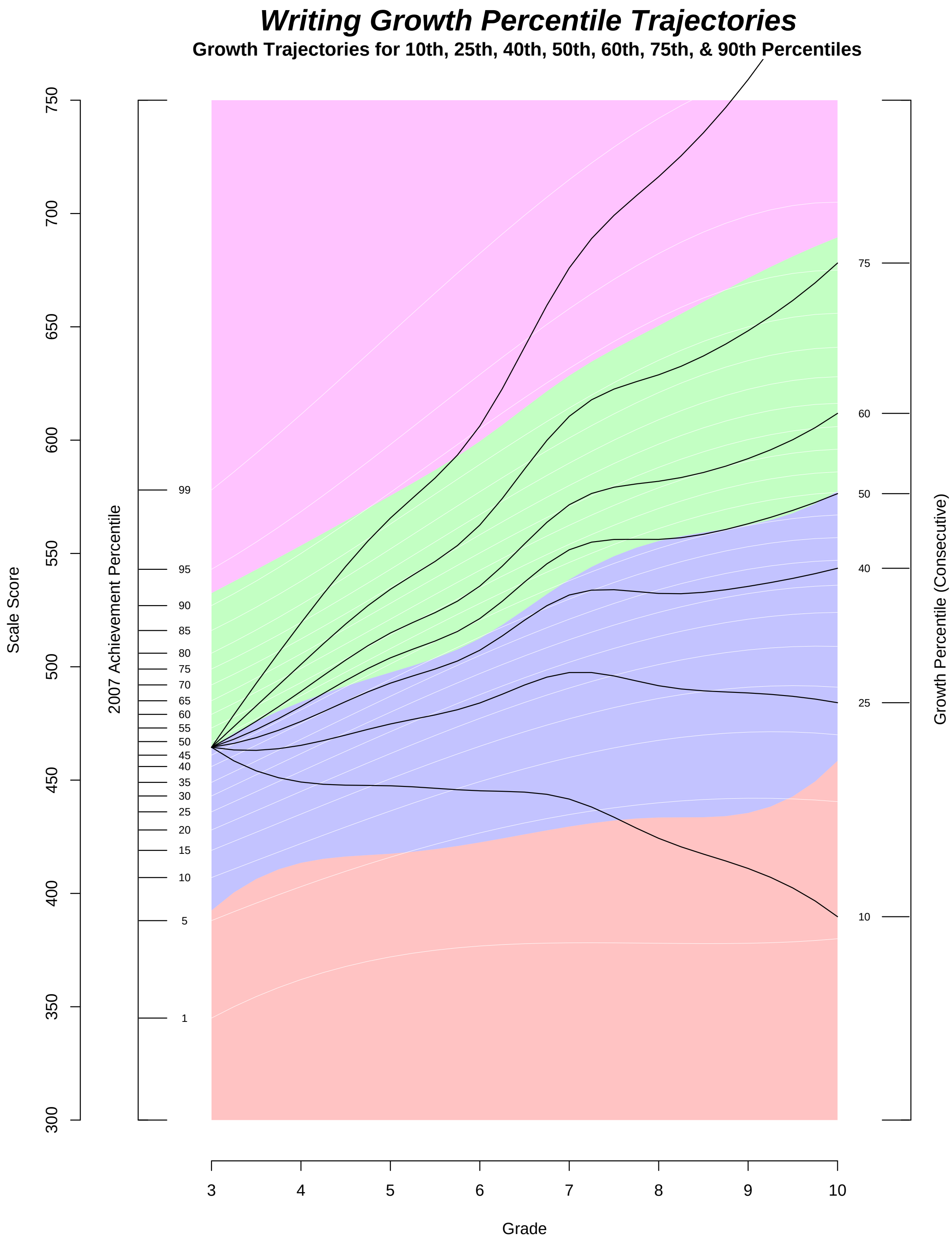

Figure 9: Growth chart depicting future writing achievement conditional upon consecutive 10th, 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning at the third grade at the partially proficient/proficient cutpoint

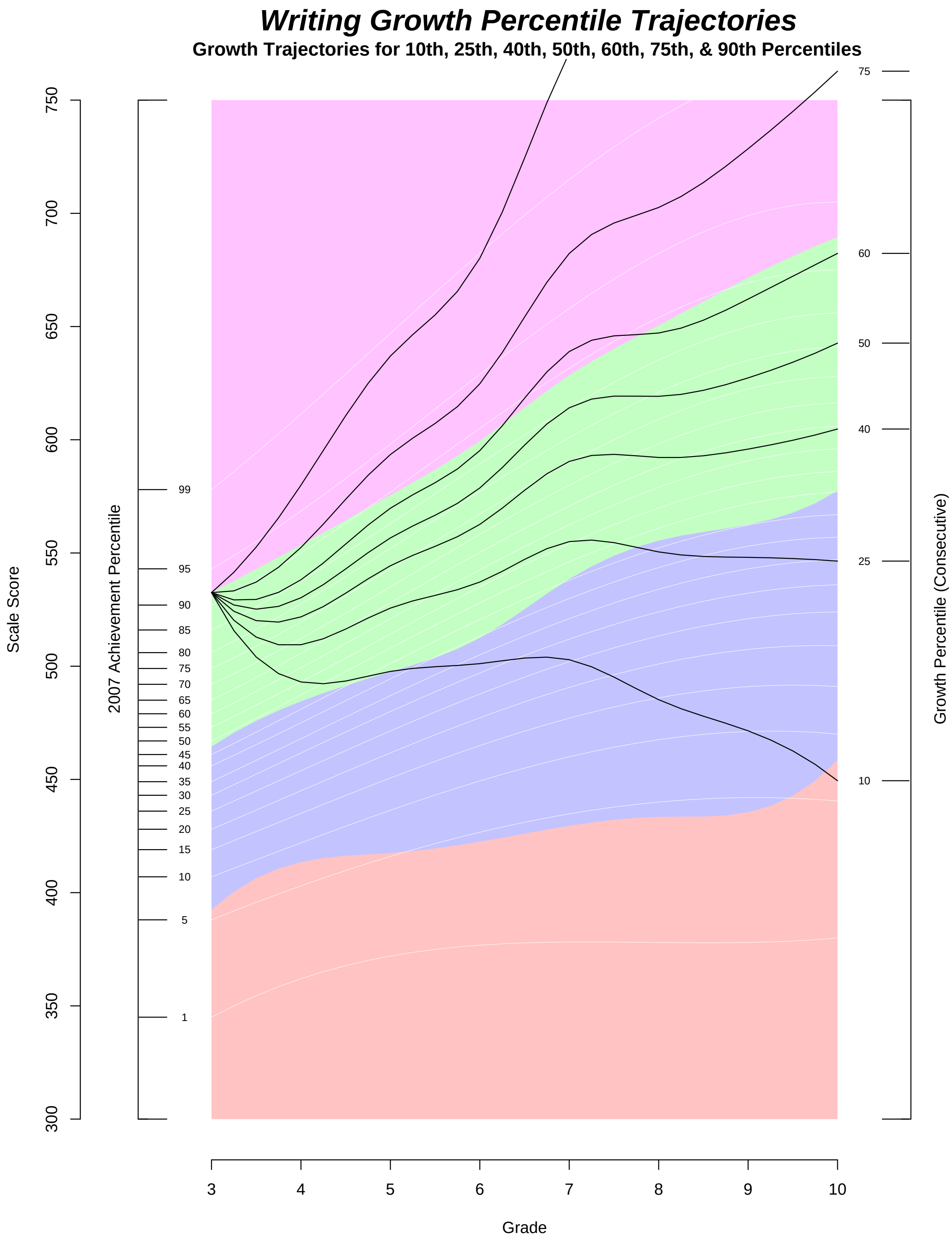

Figure 10: Growth chart depicting future writing achievement conditional upon consecutive 10th, 25th, 40th, 50th, 60th, 75th, and 90th percentile growth for a student beginning at the third grade at the proficient/advanced cutpoint