TEXAS EDUCATION AGENCY
GROWTH MODEL PILOT APPLICATION
FOR ADEQUATE YEARLY PROGRESS DETERMINATIONS UNDER THE NO CHILD LEFT BEHIND ACT

Submitted to the U. S. Department of Education
October 15, 2008
(Revised November 12, 2008)

Abstract

Texas proposes to include a student prediction measure in calculations of Adequate Yearly Progress (AYP) in 2009. Students' 2009 Texas Assessment of Knowledge and Skills (TAKS), TAKS (Accommodated), and LAT scores in both reading/English language arts and mathematics and campus-level mean scores in these subjects will be used to predict their performance in the prediction grade, which is the next high-stakes testing grade (i.e., grade 5, 8, and 11), separately for reading/English language arts and for mathematics. Students who are predicted to be at or above proficiency in the prediction grade are defined as making adequate yearly growth in that subject. Starting in 2010, when academic achievement standards are set for the TAKS-Modified (TAKS-M) and TAKS-Alternate (TAKS-Alt) assessments, Texas proposes to expand the prediction method to the TAKS-M assessments and implement a transition table approach to growth for students participating in TAKS-Alt. Adding a growth measure to AYP calculations will result in one change to the Texas AYP determination process. To meet AYP in Texas under the current process, for all districts and campuses, all students and each student group (African American, Hispanic, white, economically disadvantaged, special education, and limited English proficient) meeting minimum size requirements must meet (1) either the performance standard for percent proficient or performance gains criteria, and (2) the standard for participation in the assessment program. The inclusion of the prediction measure would impact the way the performance standard for percent proficient is calculated. Students who are predicted to meet proficiency will be counted in the numerator of the AYP percent proficiency calculation along with students meeting the standard, and this new percent would be compared with the AYP targets to determine if the performance standard for percent proficient is met. For all students and each student group, AYP performance standard requirements would be met if the percent proficient ***or*** meeting adequate yearly growth, for grades 3-8 and 10 summed across grades by subject for reading/English language arts and mathematics, meets or exceeds the AYP targets. The inclusion of a growth measure will not change the way the performance gains criteria are applied.

Introduction

The primary purpose of the state student assessment program is to provide an accurate measure of student achievement and student progress in reading/English language arts, writing, mathematics, science, and social studies. Test performance results are also used as an indicator for district and school accountability.

To meet AYP in 2008-2009, each student group (African American, Hispanic, white, economically disadvantaged, special education, and limited English proficient) meeting minimum size requirements for a campus or district must meet (1) either the performance standard for percent proficient or performance gains criteria, and (2) the standard for participation in the assessment program. There are no minimum size requirements for All Students. Test results evaluated for calculation of AYP include reading/English language arts and mathematics performance on the following assessments:

- Texas Assessment of Knowledge and Skills (TAKS), English and Spanish versions
- Texas Assessment of Knowledge and Skills (Accommodated), English and Spanish versions, for students served by special education who meet the eligibility requirements for certain specific accommodations
- Texas Assessment of Knowledge and Skills - Modified (TAKS-M), an alternate assessment based on modified academic achievement standards designed for students receiving special education services who have a disability that significantly affects academic progress in the grade-level curriculum and precludes the achievement of grade-level proficiency within a school year. Proficient results from the TAKS-M assessments will be subject to the 2% federal cap limit on proficient scores used in AYP in accordance with the United States Department of Education (USDE) final regulations.
- Linguistically accommodated testing (LAT) for recent immigrant English language learners, administered for TAKS, TAKS (Accommodated), and TAKS–M
- Texas Assessment of Knowledge and Skills-Alternate (TAKS-Alt), the assessment for students receiving special education services who have the most significant cognitive disabilities and are unable to participate in the other statewide assessments even with substantial accommodations and/or modifications. Proficient results from the TAKS-Alt assessments will be subject to the 1% federal cap limit on proficient scores used in AYP in accordance with the USDE final regulations.
- Texas English Language Proficiency Assessment System (TELPAS) reading proficiency tests that measure growth in the state reading standards in a manner that takes second language learning into account. As allowed by federal regulation, the results of English language learners (ELLs) taking TELPAS reading and no other reading/English language arts assessment who are enrolled in their first school year in a United States school are included in participation rates, but their scores are not used for AYP performance calculations.

Texas proposes to use student data from the assessments listed above for calculating growth in the 2008-2009 school year. The assessments that will be used in the growth calculations in 2008-2009 are those that have received USDE approval based on the USDE Standards and Assessment peer review process.

History of Testing and Measuring Student Growth

In 1979 the state of Texas instituted a statewide testing program that, through periodic changes in legislation and policy, has grown in size, scope, and rigor. From 1981 to 1990, as required by state statute, Texas assessed minimum skills in reading, mathematics, and writing with the Texas Assessment of Basic Skills (TABS) tests (1981–1984) and then with the Texas Educational Assessment of Minimum Skills (TEAMS) tests (1985–1990). In 1990 the implementation of another criterion-referenced testing program, the Texas Assessment of Academic Skills (TAAS), shifted the focus of assessment from minimum skills to academic skills.

In the 2002–2003 school year, the Texas Assessment of Knowledge and Skills (TAKS) replaced TAAS as the primary statewide assessment program. TAKS is designed by legislative mandate to be more comprehensive than its predecessors and measures more of the state-mandated curriculum, the Texas Essential Knowledge and Skills (TEKS), at more grade levels than did TAAS. Since 2003, TAKS has been administered in English in grades 3-11 mathematics; grades 3-9 reading; grades 10-11 English language arts; grade 4 and 7 writing; grades 5, 10, and 11 science; and grades 8, 10, and 11 social studies. TAKS grade 8 science was added in 2006. Spanish TAKS has been administered since 2003 and includes grades 3-6 reading and mathematics, grade 4 writing, and grade 5 science. The high school assessments, administered at grades 9, 10, and 11, are aligned to the high school curriculum. By law, students for whom TAKS is the graduation testing requirement must pass exit level tests in four content areas—English language arts, mathematics, science, and social studies—in order to graduate from a Texas public high school.

With experience in testing, Texas has refined its ability to generate reliable test scores and promote evidence-based test-score interpretations. Regarding reliability, Texas produces an annual technical digest that contains estimates of internal consistency, standard errors of measurement (both classical and conditional), and classification accuracy (for estimates, see document at http://www.tea.state.tx.us/student.assessment/resources/techdig07/Appendices/AppendixC_MeanP-Values.pdf). To facilitate correct uses of this information, Texas dedicates a chapter in the annual digest to describing these estimates and recommending ways in which these estimates should be used and interpreted (see document at http://www.tea.state.tx.us/student.assessment/resources/techdig07/Chapters/Chapter16-Reliability.pdf). Regarding validity evidence, Texas publishes annual interpretive guides that provide examples of standard and optional assessment reports along with an explanation of appropriate uses of the scores (see document at http://www.tea.state.tx.us/student.assessment/resources/guides/interpretive/index.html). By showing sample reports and defining terms and numbers on the reports, these guides assist school personnel in understanding and interpreting student performance data as required by Section 39.030(b) of the Texas Education Code. Furthermore, Texas publishes a chapter in the technical digest describing evidence supporting the uses of the test scores from the current school year as described in the interpretive guides. The chapter includes evidence based on test content, relations to other variables, response processes, and consequences of testing (see chapter at http://www.tea.state.tx.us/student.assessment/resources/techdig07/Chapters/Chapter17-Validity.pdf). After a review of the reliability and validity evidence for the TAKS and LAT by the external peer reviewers, USDE noted in the October 27, 2006 assessment letter, “we have determined that both the English and Spanish versions of the Texas Assessment of Knowledge and Skills (TAKS) and the Linguistically Accommodated Test (LAT) meet the standards and assessment requirements under the ESEA for grades 3-8 and high school.”

Texas also has a history of measuring student growth. With the implementation of consecutive grade testing at the same time of year in the TAAS program in 1994, Texas introduced the Texas Learning Index, or TLI. The TLI made it possible to compare student performance across years within a given subject area. The TLI was provided for both the TAAS reading and mathematics tests at Grades 3 through 8 and at the exit level. The TLI was a normative growth measure such that a student with the same TLI in fourth and fifth grade mathematics demonstrated one year's typical progress and his or her performance was in about the same position in fifth grade, relative to other fifth graders, as the student's performance was in fourth grade, relative to other fourth graders. With this system all students, regardless of where they were on the scale, were able to demonstrate progress toward ultimately passing the exit level test.

When TAKS was implemented, a new growth measure, the Texas Growth Index (TGI), was introduced. The TGI provides an estimate of a student's academic growth on the TAKS tests, over two consecutive years and in two consecutive grades. This growth index is used at the campus or district level in the state accountability system. The TGI is a linear equating growth measure, such that equating methods set the mean and standard deviations of the distributions of consecutive years equal. A student’s growth is defined as the student’s score in Year 2 minus the student’s predicted score for Year 2. A student’s predicted score for Year 2 is the score in the distribution at Year 2 that corresponds to the student’s Year 1 score. The linear equating methods result in a function for each grade and subject that is applied year to year. If the student’s score is above the expected score, the student is considered to have grown. If the student's score is below the expected (predicted) score, the student is considered to have regressed. Expected growth is defined as maintaining location in the distribution year to year.

Though Texas currently has a growth measure in place, the TGI does not meet the requirements for growth-based accountability for the USDE growth pilot. Therefore, when the USDE pilot growth program was announced and student growth legislation in Texas was passed, Texas initiated three steps that led to the submission of this growth pilot proposal. First, Texas evaluated growth models used in the state and determined that none of them would meet the alignment and foundational elements of NCLB. Second, Texas compared and contrasted several growth models used in other states on practical and psychometric features to identify models that both matched well with Texas data conditions and were also likely to meet state and federal requirements. Third, Texas conducted a growth study in which two types of growth models previously approved by USDE (a growth to proficiency model and a regression-based model) in the growth pilot program were compared on policy, psychometric, and practical features. In the study, data on approximately 2.4 million students for whom TAKS reading/English language arts, mathematics, science, and social studies in English and Spanish from 2004-2007 were available were evaluated using the two model types. The model proposed in this application resulted from the Texas growth study and is a variation of the regression-based model.

Texas Data System

The history of developing Texas' data system parallels the history of testing in Texas. Since 1979, Texas has been developing and refining the data systems and processes needed to track student progress over time, across campuses and districts, and in reporting groups. The accurate tracking of student data and the archiving of data over years provides the necessary foundation for Texas to report growth and incorporate growth into its state and federal accountability systems. Three elements of the Texas data system that facilitate growth reporting include the careful tracking of current and prior-year testing data in current year data sets, a unique student tracking field, and the sophisticated data quality control procedures that Texas implements annually.

The Public Education Information Management System (PEIMS) enrollment record submitted by a district for each student enrolled on the fall snapshot date includes as data elements the district unique identification number and the unique identification number of the campus on which the student is enrolled or on which the student receives the majority of her or his instruction. Current year test answer documents submitted for each student enrolled in the grades tested on the test date also includes the district unique identification number and the campus unique identification number. Score data for each student from prior years is included in current year data sets. When current year test score data are merged with current year enrollment information, prior-year test scores are merged as well. Texas has carefully tracked prior-year student scores since the beginning of TAKS administrations.

The second element of the Texas data system that facilitates growth reporting is the unique student tracking field. Texas tracks students with a field created from combining four pieces of student information. The variable, a combination of student PEIMS number, last name, first initial, and date of birth, was used starting in 1999 with the TAAS, after analyses showed that combining student identifying information into one field provided the capability to accurately identify students over time and across campuses and districts. It is this combined field that is used to match students across time and locations.

The third element of the Texas data system that facilitates growth reporting is the sophisticated quality control procedures that Texas implements annually. The Texas

Education Agency (TEA) verifies the accuracy of the data produced by the testing contractor, Pearson, through a comprehensive verification system. In addition, Pearson has its own internal quality control system to verify the accuracy of the reports it produces for Texas school districts.

TEA's quality control system includes a number of steps. For each test administration,

1. TEA and Pearson prepare answer documents for hundreds of fictitious students who are assigned to a campus in one of three fictitious districts. Pearson grids these students' answer documents (marks the answer choices and student identification information) using detailed instructions provided by TEA. The answer documents represent real-world scenarios of the numerous correct and incorrect ways answer documents are completed by students and districts.
2. Pearson then processes, scores, and prepares reports for these fictitious students using answer keys, editing rules, and formats approved previously by TEA.
3. TEA simultaneously processes the same student-level information and produces its own reports.
4. When TEA receives Pearson's reports for the fictitious students and districts, it compares Pearson's reports with its own reports. In addition to scores, calculations, and other numerical data printed on the reports, all text, formats, and customized messages are verified.

The goal of this part of the quality control process is to verify that changes to the test documents are made properly when the scanner encounters missing or invalid data. Reports are not sent to districts until all discrepancies in the comparative data for the fictitious districts are resolved and the reports generated by TEA and Pearson agree. In addition, the verification system allows TEA to monitor the distribution of all test materials, reports, and information letters.

Rationale for Using a Growth Model in Texas

The purpose of including growth in Texas' state and federal accountability systems is to offer alternative approaches to demonstrating achievement that meet both state and NCLB goals. The intention is not to lower student performance expectations but to hold all students, student groups, public schools, local education agencies, and the state to the same high expectations. Using growth in Texas' federal accountability system in 2009 is supported by other policy decisions made concerning the Texas assessment program.

The first rationale for use of a growth measure is that Texas must report student growth in 2008-2009 to meet requirements of two Texas legislative bills, House Bill 1 and Senate Bill 1031. According to House Bill 1, the Commissioner is required to determine a method for measuring annual improvement in student achievement. This requirement is tied to preparation to pass exit-level graduation tests. TEA is required to provide reports to districts on student annual improvement, and districts are required to report this information to teachers and parents. Senate Bill 1031 (SB 1031) also requires the measurement of student growth. The committee on public school accountability created by SB 1031 will review methods available to monitor each public school student, with emphasis on methods that identify demonstrable growth in academic achievement. New end-of-course assessments currently being constructed must measure annual improvement. Furthermore, TEA may consider using an existing instrument to satisfy requirements around developing criterion-referenced or end-of-course assessments only if that existing instrument allows for the measure of annual improvement. In SB 1031, Texas is required to implement a vertical scale for TAKS in grades 3-8 in reading/English language arts and mathematics in 2009. Since the vertical scale will facilitate interpretations of score changes across years, this aspect of SB 1031 indirectly relates to measuring student growth.

A second reason supporting the inclusion of growth in the Texas federal accountability system is Texas' history and expertise measuring and reporting growth. As explained earlier, Texas has been reporting student growth using the TLI and TGI since 1994.

A third reason is the investment Texas has made in selecting the growth model best suited to the state's data structures and to meeting state growth requirements and federal requirements for including growth in AYP calculations. The study took place over a year and included input from numerous advisory groups (e.g., the Select Committee on Public School Accountability, the Growth Advisory Meeting, the Student Assessment Advisory Committee, Texas Technical Advisory Committee, the Accountability Focus Group, the District Advisory Committee). The evaluation indicated the percent of students for whom sufficient data were available in the pilot study for calculating growth for each method and the percent of students who met growth expectations under each method in 2007. Texas would prefer to implement the same growth measure model for both state and federal purposes. Given the state requirement to report growth in 2009, federal approval of a growth model for use with mathematics and reading in 2009 AYP would allow this dual use of a growth measure to occur.

Match Rates and Sufficient Growth Data

Texas is able to match student data across subjects, years, and locations because of the three features of the data system described above: the inclusion of prior-year scores in current-year data sets, the unique student identification field, and the quality control system used to verify the accuracy of student assessment data.

When evaluating match rates (or the percent of students with scale scores that can be matched over subjects, campuses, districts, and years) for this proposal, match rates for students within the current year and across subjects will determine the students for whom predictions can be made. These are the students for whom adequate yearly growth can be reported and added to AYP calculations. In addition, the proposed prediction model uses current year reading/English language arts and mathematics scores to predict student performance at the next high stakes grade (i.e., grades used for promotion decisions and for graduation). See page 11 for a summary of the model and Appendix 1 for technical details on the model. Table 1 lists all current grades for which predictions will be made and the prediction grade. In other words, students with valid scores reading/English language arts and mathematics scores in the current grade will have their performance predicted to the prediction grade.

Table 1. Prediction Grades for the Texas Growth Model

Current Grade	Prediction Grade
3	5
4	5
5	8
6	8
7	8
8	11
9	11
10	11
11	N/A

Predictions are calculated for students with valid scores in both reading/English language arts and mathematics in the current year. However, prediction equations are **developed** using data from the prior year and applied to the current year. Each year, prediction equations are updated for use in the following year. See Figure 1 for a graphic illustrating the equation development, application, and testing cycle, or when equations are determined, when they are applied, and when prediction accuracy evaluations begin.

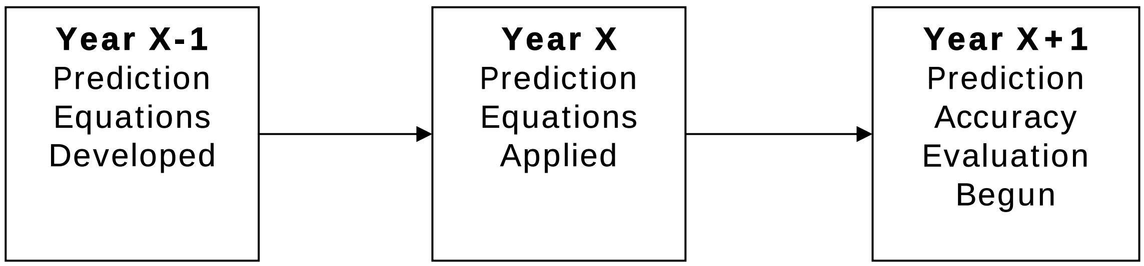

Figure 1. Growth Model Equation Development, Application, and Evaluation Process

Since Texas is proposing to include growth in 2009 AYP calculations, prediction equations will be developed using 2008 data. The equations will be applied in 2009, and prediction accuracy will begin to be evaluated for applicable grades in 2010. Prediction accuracy analyses involve comparing student's actual performance to their predicted performance. These accuracy analyses will be ongoing, such that Texas will be able to evaluate the accuracy of the 2009 predictions to some grades in 2010 (i.e., predictions from 2009 to 2010 in grades 4 to 5, 7 to 8, 10 to 11), to other grades in 2011 (i.e., grades 3 to 5, 6 to 8, and 9 to 11), and to the remaining grades in 2012 (i.e., grades 5 to 8 and 8 to 11).

Given the way that the predictions are developed and applied, match rates ***across subject areas within a year*** are the rates that will determine for which students predictions will be made. In other words, students with these matches will have predictions in 2009. Also important for the development of the equations are match rates ***across years***. For the equations developed using student data from the year to accurately predict students the following year, the equation estimates need to be made using students representative of the students for whom the equations will be applied.

For the ***application*** of the prediction equations and for reporting predictions, the evaluation of matches across subjects within a given year is the focus. Table 2 illustrates the percents of students taking English versions of tests in the Texas growth study, which used 2007 data, for whom sufficient data were matched across subjects for reporting growth. This table is currently being rerun using all students who tested in 2008 (English and Spanish versions), and results will be presented in the final proposal for dissemination to the growth peer reviewers. Preliminary analyses indicate that the percents of students in 2008 with matched data are similar to the percents in the 2007 study. The percent of students with sufficient matched data for growth calculations is indicated for reading/English language arts and mathematics in the column labeled Match. The table includes all students who tested on the English version in spring 2007 and had valid scores in both reading/English language arts and mathematics in 2007. Retained students receive predictions based on current year scores, and are included.

Students excluded from the table are those who did not have a valid score in both the English versions of the assessments in reading/English language arts and mathematics in 2007. These excluded students are those who tested with Spanish and the LAT versions of TAKS and TAKS (Accommodated), students who took the TAKS or TAKS (Accommodated) in one subject in 2007 and took an alternate assessment in the other, and students who were absent on one of the testing days and did not make up the test.

Table 2 results indicate that for the approximately 2.4 million English testers in the 2007 study, 98% of those students had sufficient data to obtain a prediction in reading/English language arts and 98% had sufficient data to obtain a prediction in mathematics. Results for AYP reporting groups indicate that for all ethnic groups, for students identified as economically disadvantaged/non-economically disadvantaged, and for students identified as LEP/non-LEP, cross-subject match rates were high, exceeding 90%. Lower match rates were found for student groups with missing indicator values and students in special education. The number of students with missing indicator values is small relative to the student population (less than 0.2% in all cases), so the lower match rate for these groups does not affect many students. The lower match rate for students in special education is of greater concern. The likely reason the match rate is lower for these students is that admission, review, and dismissal teams for students with disabilities often determine that the most appropriate assessment for a student with disabilities in one subject is TAKS and in another subject is an alternate assessment. In 2007, the alternate assessment available for students with disabilities was the State-Developed Alternative Assessment (SDAA II), which has been replaced by the TAKS (Accommodated), TAKS-M, and TAKS-Alt tests in the 2007-2008 school year. The inclusion of all alternate assessments in the growth calculations in future years will improve match rates for students in special education.

One way to include more students in the growth calculations would be to create prediction equations using only one subject-area score, such as predicting student performance in grade 8 reading using a student's score in grade 5 reading and not grade 5 mathematics. Analyses are underway to evaluate the loss in prediction accuracy due to the removal of one of the current-year scores for students. However, until evidence can be provided that the loss in prediction accuracy justifies the inclusion of students with only one valid score in the growth measure, Texas plans to predict student performance using both current-year scores. Students without both current-year scores will not have a prediction, but they will be included in the accountability system based on their status score in the subjects for which they have data.

Table 2. Match Rates Across Subjects in 2007 for Making Student Predictions

GROUP		READING/ENGLISH LANGUAGE ARTS			MATHEMATICS		
		Number Tested	%	Percent Matched	Number Tested	%	Percent Matched
TOTAL		2,354,561	100	98	2,347,546	100	98
GENDER	MALE	1,175,567	50	98	1,173,287	50	98
	FEMALE	1,177,773	50	98	1,172,884	50	99
	NO INFORMATION PROVIDED	1,221	0	87	1,375	0	78
ETHNICITY	NATIVE AMERICAN	8,447	0	97	8,402	0	97
	ASIAN	82,169	4	99	82,457	4	99
	AFRICAN AMERICAN	336,324	14	97	334,524	14	98
	HISPANIC	1,021,395	44	98	1,019,498	43	98
	WHITE	904,510	39	98	900,651	38	99
	NO INFORMATION PROVIDED	1,716	0	85	2,014	0	72
ECONOMIC DISADVAN.	YES	1,192,278	51	98	1,188,711	51	98
	NO	1,158,287	49	98	1,154,469	49	99
	NO INFORMATION PROVIDED	3,996	0	90	4,366	0	83
LIMITED ENGLISH PROFICIENT	LEP	171,282	7	98	173,752	7	96
	NON-LEP	2,179,599	93	98	2,169,767	92	98
	NO INFORMATION PROVIDED	3,680	0	88	4,027	0	80
SPECIAL EDUCATION	YES	98,602	4	79	90,870	4	86
	NO	2,252,373	96	99	2,252,781	96	99
	NO INFORMATION PROVIDED	3586	0	89	3,895	0	82

As mentioned above, two types of match rates are important in the development and application of the prediction equations. The match rates above are those indicating the numbers of students who will have predictions. The other match rates of importance are those that determine which student data are used in the development of the prediction equations. The equations need to be developed with a sample of students who are representative of the population with whom the equations will be applied. Table 3 illustrates the number and percent of students with sufficient matched data in reading/English language arts and mathematics, respectively, across years to ***develop*** prediction equations for these two subjects. These values are provided for the 2008 data for English TAKS and TAKS (Accommodated), which are the data that will be used to determine the prediction equations for English testers to be applied in 2009 for the Texas growth model. These student numbers and percents were calculated from the 2008 statewide assessment results. The number tested is the number of students in the prediction grade (grades 5, 8, and 11) who tested using the English version of the TAKS or TAKS (Accommodated). The number matched is the number of these 2008 testers who had valid matched history data in both reading/English language arts and mathematics in the grade from which the prediction is made. For example, matches are made for the following situations:

Table 3. Data Needed for Making Predictions

Prediction	Grade in 2008	Valid Scores Needed for Match
Grade 3 to 5	5	Grade 3 Reading and Mathematics
Grade 4 to 5	5	Grade 4 Reading and Mathematics
Grade 5 to 8	8	Grade 5 Reading and Mathematics
Grade 6 to 8	8	Grade 6 Reading and Mathematics
Grade 7 to 8	8	Grade 7 Reading and Mathematics
Grade 8 to 11	11	Grade 8 Reading and Mathematics
Grade 9 to 11	11	Grade 9 Reading and Mathematics
Grade 10 to 11	11	Grade 10 English Language Arts and Mathematics

Table 4 results indicate that the percent of matched students is above 80% for all students and all student ethnicity groups. This percent is high, given matches are made for 1, 2, and 3 years in the past. Analyses examining match rates for predictions across different numbers of years indicate that for all students in reading and mathematics, match rates are 86%, 82%, and 79% for predictions across 1, 2, and 3 years, respectively.

For students identified as economically disadvantaged, the percent of matched data is just under 80%. Since this match rate is slightly lower than the match rate for all students, Texas will carefully monitor match rates in this group. The two AYP reporting student groups with expectedly low match rates in the English versions of the TAKS and TAKS (Accommodated) are students with limited English proficiency and students in special education. In the case of LEP students, match rates are impacted by several factors. First, many current-year LEP students who took the test in English were not in the United States in prior years or did not have test data since they were excluded from testing requirements for first year LEP testers. Second, many LEP students testing in English in the current year tested in Spanish in prior years. Third, analyses indicate that many LEP students without matched data took a different assessment in prior years (e.g., SDAA II). Texas will likely continue to observe lower match rates for LEP students taking English TAKS, given that LEP students will continue to enter Texas as first year immigrants and most LEP students by design will transition from one assessment to another (e.g., Spanish to English).

Match rates for special education students taking TAKS and TAKS (Accommodated) in English were also expectedly low. The match rates for these students are affected by the change in special education testing requirements to meet federal requirements over the last several years. That is, the change from the State Developed Alternate Assessment II (SDAA II) as last administered in spring 2007 to the full implementation of the TAKS Accommodated, TAKS-M, and TAKS-Alt tests in the 2007-2008 school year. For example, based on participation rates for special education students in 2006 and 2007, approximately 62% of the students served in special education programs took the former alternate assessment, the SDAA II, while approximately 33% took TAKS during those years. Therefore, the match rate of special education students taking TAKS in 2008 to a prior year TAKS assessment will not be comparable to the match rate of the other student groups. Based on participation rates for special education students in 2008, these match rates will significantly increase with the transition from the SDAA II assessments to the new assessments for students with disabilities. In 2008, over 60% of the students served in special education programs were tested on either TAKS or TAKS (Accommodated), which is substantially higher than the approximately 33% tested on TAKS in 2006 and 2007.

Though match rates for these two student groups are expectedly lower than for other groups, evidence from the prediction accuracy study shown in Appendix 1 suggests that prediction accuracy from one year to the next using prediction equations with match rates as shown above were as accurate as predictions based on equations developed using more students. Furthermore, match rates used to develop prediction equations will be closely monitored to ensure that a sufficient number of students are available to develop valid and reliable predictions equations.

Texas will submit these same data for the Spanish versions of the TAKS and TAKS (Accommodated) in the final proposal submission in November.

Table 4. Match Rates Across Years for Prediction Equation Development

	READING/ENGLISH LANGUAGE ARTS			MATHEMATICS		
	Number Tested	Number Matched	Percent Matched	Number Tested	Number Matched	Percent Matched
All Students	2,339,378	1,937,065	82.8%	2,332,274	1,928,666	82.7%
African American	329,807	264,986	80.3%	327,009	263,426	80.6%
Hispanic	1,008,893	815,002	80.8%	1,009,502	811,016	80.3%
White Students	906,605	781,082	86.2%	901,276	778,278	86.4%
Economically Disadvantaged	1,129,682	898,204	79.5%	1,127,133	893,082	79.2%
Limited English Proficiency	158,329	71,964	45.5%	163,210	71,346	43.7%
Special Education	164,895	53,188	32.3%	156,806	51,505	32.8%

Summary of the Texas Prediction Model and Planned Models for the Alternate Assessments

Texas will implement a prediction model in 2009 AYP calculations for students taking the general TAKS and TAKS (Accommodated) assessments. Once sufficient data are available for the TAKS-M alternate assessment, Texas will implement prediction equations like those used with the general assessment. For TAKS-Alt, the assessment for students with severe cognitive disabilities, Texas will implement a transition table approach to growth. A summary of the prediction model for students taking TAKS and TAKS (Accommodated) is described below. In addition, the plans for the growth models for the alternate assessments are also described. For technical details on the prediction model approach proposed for TAKS and planned for TAKS-M, see Appendix 1.

Prediction Model for the General Assessment

The proposed Texas growth model for TAKS, TAKS (Accommodated), and Linguistically Accommodated Test (LAT) is a multi-level regression-based prediction model. See Appendix 1 for more details. The model predicts student performance separately in reading/English language arts and in mathematics in the next high-stakes grade (defined by Texas legislation as grades 5, 8, and 11) using current year scale scores in both reading/English language arts and mathematics and campus-level mean scores in these subjects. Prediction equations are developed the year before they are applied, so that the formulas can be published and shared across the state before they are used in AYP calculations. For example, prediction equations developed in 2008 will be applied in 2009 to predict student performance. A student predicted to be at or above proficiency in the prediction grade is determined to have met adequate yearly growth. Predictions will be made for all students each year who have valid scores in reading/English language arts and mathematics. The prediction equations will be updated each year after operational testing and will be published before their use the following spring.

The decision to use only current year reading/English language arts and mathematics scores in the prediction equations was made to balance transparency and validity. By using current year reading/English language arts and mathematics scores and mean campus-level scores in the prediction equations, Texas is able to publish prediction equations months before they are applied, making the growth model fully transparent to decision makers. Furthermore, results from the prediction accuracy study using two cohorts in the 2007 Texas growth study in which student predictions were compared with a more statistically complex, multivariate longitudinal projection model (much like the growth models approved by USDE for Tennessee and Ohio) indicated that similar prediction accuracy values were found for the prediction model using only current year reading/English language arts and mathematics scores and campus-level mean scores for these subjects as for the projection model using all student scores in all subjects from 2004-2007. In Table 5, results of the comparison are summarized. Specifically, for the Texas prediction model using only 2007 reading/English language arts and mathematics scores predicting 2008, 94% of the approximately 271,000 students included in the study were both predicted to meet the standard in 2008 and actually met the standard in 2008, and 2% of the students were accurately predicted to fail to meet the standard. Therefore, 96% of students were accurately predicted using this model. For the more complex model using all students' reading/English language arts, mathematics, science, social studies, and writing scores in 2004-2007 to project to 2008, the percent correctly classified was the same. For the other grades and subjects, the percents of students accurately classified were the same or within one percent across the models. For more details on this study, see Appendix 1.

Table 5. Prediction Accuracy for Texas Regression-Based Model and More Complex Regression-Based Model

PREDICTION YEAR, GRADE, AND SUBJECT	TEXAS MODEL			MORE COMPLEX MODEL		
	N	Perfect Agreement Met Standard	Perfect Agreement Did Not Meet Standard	N	Perfect Agreement Met Standard	Perfect Agreement Did Not Meet Standard
2008 Grade 8 Reading	270,700	94	2	269,015	94	2
2008 Grade 8 Mathematics	269,675	73	13	267,540	73	14
2008 Grade 11 English Language Arts	224,341	79	10	228,110	78	11
2008 Grade 11 Mathematics	222,603	93	1	225,923	92	3

Prediction Model for the TAKS–M Alternate Assessment

Texas proposes to initiate the same type of multi-level regression-based prediction model as described above starting in 2010 for students taking the TAKS-M alternate assessment (the 2% assessment). However, to develop the prediction equations for this assessment, TAKS-M data for students in both the current and prediction grades need to be available. In other words, to predict students' TAKS-M performance in grade 8 from grade 5, data on at least one cohort of students who took TAKS-M in grades 5 and 8 are needed. Because TAKS-M was administered for the first time in 2008, the ability to apply the prediction model to this assessment will need to be phased in as the data become available.

Before implementing the prediction equations for TAKS-M in 2010, Texas will complete the second operational administration of this assessment in 2009, conduct an empirical analysis of the match rates for this assessment, and evaluate the stability of the prediction equations with the population of TAKS-M testers in 2009. Once these activities are completed, Texas will apply the 2009 prediction equations to the 2010 TAKS-M scores in grades 4, 7, and 10. Then, once TAKS-M data are available for 2008, 2009, and 2010, prediction equations for

TAKS-M for grades 4, 7, and 10 will be updated. These new prediction equations from 2010 will then be applied in 2011 in grades 3, 6, and 9. This process will continue until prediction equations for all TAKS-M grades are available. Table 6 summarizes the phase-in for the TAKS-M prediction equations.

Table 6. TAKS-M Assessments Prediction Equation Timeline

Current Grade	Prediction Grade	Year Data Available on First Cohort	First Year Equations Applied
3	5	2010	2011
4	5	2009	2010
5	8	2011	2012
6	8	2010	2011
7	8	2009	2010
8	11	2011	2012
9	11	2010	2011
10	11	2009	2010
11	N/A	N/A	N/A

Growth Model for the TAKS-Alt Alternate Assessment

For students participating in the TAKS-Alt assessment (i.e., the 1% assessment), Texas will implement a transition table growth model similar to the growth model Michigan was approved to use in AYP calculations. The TAKS-Alt progress model will require that Texas subdivide the three proficiency levels (Did Not Meet Standard, Met Standard, and Commended Performance) into three sublevels (low, middle, and high). Once the performance levels are subdivided, Texas will develop a descriptive transition table that describes students' progress relative to their progress expectations. This descriptive table will describe progress for all students, those in all three proficiency levels. Finally, Texas will set progress targets that require students below proficiency to reach proficiency by the next high-stakes grade (i.e., grade 5, 8, and 11). The growth model for TAKS-Alt will be implemented for the first time in 2010. Since this type of progress model does not require prediction equations, this model will be implemented for all grades in reading/English language arts and mathematics in 2010. Due to the difficulty of gathering evidence to support that the TAKS-Alt scales are interval-level, linear, and measure only one type of achievement, Texas proposes a different type of growth model for this assessment, one that does not rely on the same psychometric properties as a prediction model.

Table 7 below illustrates a progress target table showing transitions that TAKS-Alt students who did not meet the standard would be required to make in order to meet progress targets each year. These progress requirements would result in students' meeting the standard by the next high stakes grade. The table shows the progress students in different sublevels in the Did Not Meet Standard category would need to make each year. Depending on the grade in which students are in, students are expected to meet the standard in one (e.g., grade 7 to 8), two (e.g., grade 6 to 8), or three years (e.g., grade 5 to 8). For a student who has one year to meet the standard, that student must make all transitions to meet the standard by the next grade. For example, a student who is in the lowest sublevel of the Did Not Meet the Standard level in grade 4 would be expected to make three transitions by grade 5, where the transitions would include low to middle, middle to high, and high to Met Standard. A student who is in the low sublevel in a grade for which the student has two years to meet the standard would have different progress targets. For example, a student in the low sublevel would need to make three transitions in two years to meet the standard.

Since that student has two years to make the transitions, the student would be expected to progress one level in one of the two years and two levels in one of the years. A student who moved two levels in the first year and one level in the second year would meet progress expectations. Similarly, a student who moved one level in the first year and two levels in the second year would also meet progress expectations.

Table 7. Example Progress Target Table

Number of Years from Current Grade to Prediction Grade	Previous Performance Level		Number of Sub-Levels Improvement Needed to Achieve Proficiency	Number of Years to Achieve Proficiency	Progress Target
	Level	Sublevel			
1 year	Did Not Meet Standard	Low	3	1	Students must increase 3 sub-levels
		Middle	2	1	Students must increase 2 sub-levels
		High	1	1	Students must increase 1 sub-levels
2 years	Did Not Meet Standard	Low	3	2	Students must increase 2 sub-levels one year and 1 sub-level the other.
		Middle	2	2	Students must increase 1 sub-level each year.
		High	1	2	Students must increase 1 sub-level either year.
3 years	Did Not Meet Standard	Low	3	3	Students must increase 1 sub-level each year.
		Middle	2	3	Students must increase 1 sub-level in 2 of the three years.
		High	1	3	Students must increase 1 sub-level in one of the three years.

Application of Growth/Prediction Model for Students Changing Assessments

As described above, Texas plans to implement a prediction model for all assessments except TAKS-Alt. The same prediction approach will be used for English versions of the TAKS and TAKS (Accommodated) and LAT, testers taking the Spanish versions of TAKS and TAKS (Accommodated) and the TAKS-M tests. However, the prediction equations will be unique to these student populations. The need for the different prediction equations is the lack of sufficient numbers of students who take different assessments in different grades to develop stable prediction equations. Furthermore, the approach for evaluating growth for TAKS-Alt students is fundamentally different from the prediction equation approach. Given that students may take different assessments in reading/English language arts and mathematics across years, the growth/prediction models should account for students changing assessments. Decisions about Texas' proposed growth models were made considering these types of students. For the prediction equations, because scores in reading/English language arts and mathematics in the current year are used for making the predictions, students who change assessments across years will still receive a prediction as long as they have valid reading/English language arts and mathematics scores in the current year. Furthermore, students who switch testing language versions will also receive predictions, where the prediction equations in the tested language will be applied. The following set of general inclusion/exclusion rules will apply to students who do not take the same assessments within one year or across years:

- Students who take different assessments for reading/English language arts and mathematics in the current year
  - Prediction equations are not planned to be applied if students do not have both a valid reading/English language arts and mathematics score. Analyses are underway to examine the prediction accuracy of equations using only one score, but until evidence supporting the reliability and validity of making predictions with only one score is obtained, Texas will not report predictions for these students.
- Students taking Spanish versions of TAKS or TAKS (Accommodated)
  - Students taking Spanish versions of TAKS or TAKS (Accommodated) in grades 3-5 will have predictions starting in 2009. Students in grades 3 and 4 will be predicted to grade 5.
  - Students taking grade 5 Spanish version tests will be predicted to grade 8 English version tests pending analyses to confirm that a sufficient number of the approximately 5,000 students who take the Spanish version TAKS or TAKS (Accommodated) in grade 5 also take the English version TAKS or TAKS (Accommodated) in grade 8. The number of students will need to be great enough to produce stable parameters for the prediction equations. The evaluation of what constitutes a sufficient number of students and stable parameters will involve recalculating the prediction equations for other cohorts, such as the grade 5 cohort of English testers who are predicted to grade 8, after student data are removed systematically to reflect the students with matches in grade 5 Spanish and grade 8 English. These analyses are in progress and will be completed by November 2008.
  - Students taking the Spanish version of TAKS or TAKS (Accommodated) in grade 6 will not be included in the growth calculations. One reason is that the number of students taking the Spanish version of tests in grade 6 is just under 1,000 students. The development of the prediction equations for these students will require that a sufficient number of students in grade 8 English have grade 6 Spanish version test scores to produce stable prediction equations year to year. The number of students that meet this requirement will not likely be great enough, though analyses are underway to verify this assumption.
- Students who take TAKS or TAKS (Accommodated) in different languages
  - If students switch the language version in which they test across years, prediction equations can still be applied as long as students have a valid reading/English language arts and mathematics score in the same language in the current year. For example, if a student takes TAKS reading and mathematics Spanish versions in 2008 and English versions in 2009, that student will receive predictions both years. For 2008, the student's predictions will come from the application of the Spanish TAKS prediction equations. For 2009, the student's predictions will come from the English TAKS prediction equations.
  - For students taking TAKS or TAKS (Accommodated) in different languages across subject areas in the current year, no predictions will be made. The reason is that the numbers of students who would take the assessments in different languages would not support the development of stable prediction equations.

Application of Growth/Prediction Model for Students Under Various Circumstances

The way in which the prediction model will be applied and scores will be interpreted for students under various circumstances is explained below:

- A student who changes schools—the student who changes schools will receive a prediction, and the prediction will be included in AYP calculations as long as the student has a valid score in reading/English language arts and mathematics in the current grade.
- A student who changes LEAs—the student who changes LEAs will receive a prediction, and the prediction will be included in AYP calculations as long as the student has a valid score in reading/English language arts and mathematics in the current grade.
- A student who is retained—a retained student will be predicted using current year scores. For example, a student in grade 5 in 2009 will be predicted to grade 8 using prediction equations from 2008. If that student is retained in grade 5 in 2010 and that student has a valid score in both subjects in 2010 in grade 5, that student will be predicted to grade 8 using prediction equations from 2009.
- A student who is new to Texas—if a new student obtains a valid score in reading/English language arts and mathematics in the current grade, that student will receive a prediction.
- A student who makes progress but then falls behind—this student will be identified using the prediction model, and this information will be reported to campuses and districts so that interventions can be planned and implemented. When the student is making progress, the student's prediction will indicate that progress. When the student starts to fall behind, the prediction will indicate this as well.
- A student who is proficient at all times but steadily declining—the declining performance will result in declining predictions that will be reported to campuses and districts. It is anticipated that annual student growth reports will include a graph of student scale scores over grades. The decline in performance will be evident on these graphs.
- Calculations for growth on all students, not just non-proficient students—Texas' prediction model and growth model for the TAKS-Alt will provide growth information on all students, not just students who are in the Did Not Meet Standard category. As noted in the section below, ***Focusing Interventions Using Growth Information,*** four groups of students will be clearly identified in reports to campuses and districts starting in 2009 including (1) students who meet standard and meet growth, (2) students who meet standard but not growth, (3) students who meet growth but not standard, and (4) students who neither meet standard or growth. By distinguishing these four categories for all students and for students in AYP reporting groups, Texas will promote the use of growth information for all students and not just for those who Did Not Meet the Standard.

Fit of the Proposed Prediction Model in Texas Assessment System

Texas chose the prediction growth model because this model closely fits the context of the state's existing assessment system and the assessment and data systems support the prediction model. The reasons for this fit include:

1. **The prediction model predicts student performance in grades 5, 8, and 11, grades that are already part of the current high stakes structure in the Texas assessment system.** The proposed prediction model uses students' current year reading/English language arts and mathematics scores to predict performance to the next high-stakes grade, that is, the next Student Success Initiative (SSI) grade or exit level. The SSI provides a system of academic support to help ensure achievement on grade level in reading and mathematics so that every student can

succeed throughout his or her school career. The SSI incorporates a grade advancement component adopted by the Texas Legislature in 1999. The law ties promotion to performance on state-mandated assessments in reading at grade 3, reading and mathematics at grade 5, reading and mathematics at grade 8, and graduation to exit-level performance at grade 11.

2. **The prediction model balances accuracy and transparency.** Texas has a history of using transparent calculations in the state accountability system. Texas districts frequently perform their own accountability calculations at the campus and district level based on transparent methods. By using prior-year equations and publishing them in advance of their application, Texas will maintain the use of transparent calculations for high stakes accountability.
3. **The prediction model is built to reflect many of the features of the Dallas regression-based model.** The Dallas Independent School District (DISD) has been implementing a prediction model using current year scores to predict subsequent year scores since 1992. Texas' decision to use a regression-based model similar to the one that Dallas has been implementing was based in part on efficiency, that is, the ability to take advantage of lessons learned by DISD through long term implementation of a regression-based prediction model using state required assessments.

Proposed Use of Growth Data in AYP Calculations

Texas determines AYP for all districts and campuses and for all students and each student group (African American, Hispanic, white, economically disadvantaged, special education, and limited English proficient) meeting minimum size requirements, where minimum size requirements at the student group level are 50 total students in the grades tested (summed across grades 3-8 and 10) and the student group comprises at least ten percent of all test takers in the subject area. There are no minimum size requirements for the All Students group. Steps in determining AYP are listed below. The change to the process introduced by adding the prediction measure is in bold font. Note that for a student without a prediction, the student's actual score is used to determine that student's status in AYP. For a school or district to make AYP, all students and each student group that meets minimum size criteria must

- meet or exceed the AYP targets (shown below) on the assessment measure, **where the percent compared with the AYP target is the percent meeting the standard or predicted to be proficient** or the performance gains provisions under safe harbor, and

	AYP Targets								
	Target 2002-03 2003-04	Target 2004-05 2005-06	Target 2006-07 2007-08	Target 2008-09	Target 2009-10	Target 2010-11	Target 2011-12	Target 2012-13	Target 2013-14
Reading/English Language Arts	47%	53%	60%	67%	73%	80%	87%	93%	100%
Mathematics	33%	42%	50%	58%	67%	75%	83%	92%	100%

- have at least a 95 percent participation rate in the state assessments, and
- meet the state requirements for performance or performance gains on one other academic indicator.

For districts and campuses with a student group that does not meet minimum size requirements in the grades tested (summed across grades 3-8 and 10) for either reading/English language arts or mathematics, or no students in the grades tested, one or a combination of the following methods will be used for AYP calculations. The addition of the prediction model does not change these methods except that the percent of students compared with the AYP targets is the percent of students meeting the standard plus the percent predicted to pass according to the prediction model (or those on track to pass in the transition table approach planned for students participating in TAKS-Alt). The ways in which the prediction model fits in with these methods are noted in bold font.

- Use the pairing relationships established for the state accountability system for campuses with no students in grades tested.
- Evaluate districts and campuses on test results for all students if none of the student groups meet minimum size requirements.
- Incorporate confidence intervals using the standard error of the proportion **(this proportion includes the number of students predicted to pass or are on track to pass according to the transition table approach for TAKS-Alt)** for student groups that do not meet minimum size requirements.
- Aggregate two or more years of assessment **and prediction/growth** results.
- Assign the district AYP status to schools with too few students to evaluate.

Given that growth will be calculated and reported each year for each student and given the way that growth will be incorporated into AYP calculations, it will not be possible for the proficiency of high-performing students to compensate for the performance of lower performing students.

Focusing Interventions Using Growth Information

Texas will use results from the growth model to focus district and campus school intervention efforts. By reporting information from the growth model in a way that clearly distinguishes four groups of students at each campus and district, Texas will provide a richer source of information for campus and district intervention strategies than currently provided under the status model. The four groups of students at a campus and district level that will be reported overall and for each AYP reporting group include:

1. students who meet the standard in the current grade AND who are predicted to meet the standard in the next high-stakes grade
2. students who meet the standard in the current grade but who are NOT predicted to meet the standard in the next high-stakes grade
3. students who do NOT meet the standard but who are predicted to meet the standard in the next high-stakes grade
4. students who neither meet the standard nor who are predicted to meet the standard in the next high-stakes grade

By providing student-level status and growth data, the report allows campuses and educators to easily identify individual students in need of intervention and to implement strategies for improvement. By providing campus-level and district-level rosters of students in these four categories and disaggregated by AYP reporting groups, TEA will equip campuses and districts with early information that can focus resources on students most in need.

As an example, consider two campuses with 100 students in each (campus A and campus B) where 80 out of 100 students met the standard in the current year. Under the AYP status model, these campuses appear to be performing in the same way and intervention resources might be allocated similarly for the students in these campuses. However, with the addition of growth, it might be that for campus A the number who are proficient but not predicted to meet the standard in the next high-stakes grade is low and, in campus B the number who are proficient but predicted to meet the standard in the next high-stakes grade is high. This is illustrated in Tables 8 and 9 below.

Table 8. Status Model View

CAMPUS A		CAMPUS B	
Met Standard	**Did Not Meet Standard**	**Met Standard**	**Did Not Meet Standard**
80	20	80	20

Table 9. Growth Model View

CAMPUS A				CAMPUS B			
Met Standard		**Did Not Meet Standard**		**Met Standard**		**Did Not Meet Standard**	
80		20		80		20	
Met Growth	**Did Not Meet Growth**	**Met Growth**	**Did Not Meet Growth**	**Met Growth**	**Did Not Meet Growth**	**Met Growth**	**Did Not Meet Growth**
20	60	5	15	60	20	5	15

The added information from the growth model will help these two campuses direct their attentions and resources so that interventions can be applied differently and effectively for each campus. For example, Campus A will be able to use the growth information to identify and intervene with students who score high enough in the content area to meet the current-year standard but who are not predicted to pass in the future.

Full Academic Year (FAY) Definition Used for AYP Calculations

The Texas definition of “full academic year” for Adequate Yearly Progress (AYP) is linked to the state fall enrollment snapshot date - the last Friday in October.

> Districts: Test results for students enrolled in the district on the Public Education Information Management System (PEIMS) fall enrollment snapshot date are included in the district-level performance measure.
>
> Campuses: Test results for students enrolled on the campus on the PEIMS fall enrollment snapshot date are included in the campus-level performance measure.

There are approximately 100 instructional days between the last Friday in October (PEIMS snapshot date) and the primary administration testing dates for the TAKS in April. This represents just over half (54%) of the instructional days in the 185-day school year. Compared to most other states, Texas has one of the shortest full academic year definitions in the nation because the state enrollment snapshot date is at the end of October, instead of enrollment at the same campus or district from the testing period of the previous school year. The definition of FAY in Texas does not impact match rates related to the growth measure.

Impact Data—Growth in AYP for 2008

Table 10 presents impact data compared with AYP targets showing what would happen if the percent of students in each district, campus, or student group included students predicted to be proficient. The table shows impact data for all students in the 2007 Texas growth study, for students by gender, and for students in each AYP reporting group. Analyses evaluating the impact for including student predictions in 2008 are in progress. In these 2008 analyses, data for all students in the state are included in the analyses. Furthermore, results will indicate the number and percent of campuses and districts that would meet AYP due to the inclusion of the prediction measure. The reason these analyses are not in this version of the report is that Texas has an agreement with USDE to report AYP in October 2008 so that results from TAKS-M (the 2% assessment) can be included. Results from the 2008 analyses will be reported in the final Texas growth proposal that will be ready in November 2008, in time to provide to the growth pilot study peer reviewers for their December 2008 review.

For all students in the study (TOTAL), by gender group, and for each student group, several sources of data are reported in Table 10:

- N=total number of students in the 2007 Texas growth study
- %=percent of the total number of students in each student group. For each student group, there were a small number of students who did not have the student group indicator reported. Data for these students is noted in the rows labeled "No Information." In no student group was the number of students that had no information great enough to round to 1%.
- Match=the percent of students in that group (N) who had sufficient data for the prediction.
- G=the percent of students in that group (N) who were predicted to be at or above proficiency in the prediction year.
- MS=the percent of students in that group (N) who met the standard in the current year, in 2007 for this table.
- G/M=the percent of students in that group (N) who either met the standard or were predicted to be proficient.

The impact of adding the prediction equation to AYP determinations is best evaluated by subtracting the percent MS from the percent G/M, since this difference indicates the percent of students who would be added to the numerator of the percent compared with the AYP targets due to the inclusion of the prediction measure. For example, in the first row of the table, the percent of the 2,354,561 students in the 2007 growth study who met the standard in reading/English language arts was 87%. The percent who were predicted to be at or above proficiency in the prediction grade was 93%. The percent of students who either met the standard or were predicted to be at or above proficiency was 95%. Therefore, the addition of the prediction increased the percent of the approximately 2.4 million students who had not met the standard in 2007 but were expected to meet the standard in the next high stakes grade by 8%. For AYP calculations, this additional 8% of students would be counted in the numerator for comparison to AYP targets for all students. For mathematics, 10% of the growth study students who did not meet the standard in 2007 were predicted to be at or above proficiency; therefore, an evaluation of AYP for these students would include 10% more students in the numerator of the percent compared with the AYP target.

Table 10. Impact of Implementing the Texas Prediction Model Using 2007 Growth Study Data

		Reading/English Language Arts						Mathematics					
		N	**%**	**Match**	**G**	**MS**	**G/M**	**N**	**%**	**Match**	**G**	**MS**	**G/M**
TOTAL		2354561	100	98	93	87	95	2347546	100	98	84	75	85
GENDER	**MALE**	1175567	50	98	93	85	94	1173287	50	98	83	76	84
	FEMALE	1177773	50	98	94	89	96	1172884	50	99	84	75	85
	NO INFORMATION	1221	0	87	76	70	84	1375	0	78	51	48	58
ETHNICITY	**NATIVE AMERICAN**	8447	0	97	94	89	96	8402	0	97	86	77	87
	ASIAN	82169	4	99	98	95	99	82457	4	99	95	92	96
	AFRICAN AMERICAN	336324	14	97	90	81	92	334524	14	98	73	62	74
	HISPANIC	1021395	44	98	92	82	93	1019498	43	98	79	69	80
	WHITE	904510	39	98	97	94	98	900651	38	99	92	86	93
	NO INFORMATION	1716	0	85	73	69	83	2014	0	72	46	45	53
ECONOMIC	**YES**	1192278	51	98	91	81	93	1188711	51	98	77	68	78
DISADVAN.	**NO**	1158287	49	98	96	93	98	1154469	49	99	90	83	91
	NO INFORMATION	3996	0	90	83	77	89	4366	0	83	60	55	65
LIMITED	**LEP**	171282	7	98	80	58	81	173752	7	96	62	55	64
ENGLISH	**NON-LEP**	2179599	93	98	95	89	96	2169767	92	98	85	77	86
PROFICIENT	**NO INFORMATION**	3680	0	88	77	71	84	4027	0	80	51	47	56
SPECIAL	**YES**	98602	4	79	71	71	85	90870	4	86	61	56	67
EDUCATION	**NO**	2252373	96	99	95	88	96	2252781	96	99	85	76	85
	NO INFORMATION	3586	0	89	78	71	85	3895	0	82	52	48	57

Some of the main findings from this table are that approximately 8-10% of students in each group were predicted to meet the standard but failed to do so in the current grade. If predictions were included in AYP determinations in 2007, the groups that would be impacted the most would be limited English proficient and special education students in reading/English language arts. Specifically, of the 171,282 limited English proficient students in the 2007 growth study, 23% who did not meet the standard in 2007 in reading/English language arts were predicted to meet the standard in the future. Of the 98,602 special education students in the growth study, 14% who did not meet the standard in 2007 were predicted to meet the standard in the future.

Alignment with the Core Principles

Texas' proposal aligns with all of the seven core principles for the growth model pilot as demonstrated below:

1. Goal of All Students Proficient by 2013-2014; Closing the Achievement Gap

Texas' proposal aligns with this principle, since all students are expected to become proficient by 2013-2014 or are expected to be predicted by the regression model to reach proficiency within no more than three years by 2013-2014.

2. Growth Expectations Not Moderated Based on Group or School Characteristics

Texas' proposal meets this expectation, since the prediction equations do not include group or school characteristics and the targets for performance level changes planned for the TAKS-Alt growth will not be affected by group or school characteristics.

3. Separate Accountability Determinations Based on Reading/English Language Arts and Mathematics

   The Texas proposal meets this expectation in that separate prediction equations and performance level change targets (for TAKS-Alt) are calculated and evaluated for reading/English language arts and mathematics.

4. Inclusion of All Students, Schools and Districts; Accountability for Group Performance

   The Texas proposal meets this expectation since predictions and level changes (in the case of TAKS-Alt) are calculated at the individual student level. In addition, almost all Texas students will be evaluated for adequate yearly growth. As described above, Texas will not predict student performance only in the rare instances when students change assessments and data are insufficient in these cases to develop prediction equations.

5. Two Years of Annual Assessments (Peer-Approved) in Reading/English Language Arts and Mathematics in Grades 3-8

   The Texas proposal meets or will meet this element for all of its assessments by 2009. Texas has been assessing general education students in these subjects and grades since 2003, and the TAKS, TAKS (Accommodated), and LAT reading/English language arts and mathematics assessments have been fully approved in the Standards and Assessment peer review process. The second operational administration for TAKS-M will occur in 2009, with full approval of this optional assessment expected in 2009. For the TAKS-Alt assessment for students with severe cognitive disabilities, Texas has entered a Memorandum of Agreement (MOA) with USDE in which Texas is expected to have full approval in 2009. Texas is currently fully compliant with the MOA for this assessment.

6. Texas' Data System and Proposed Growth Model will Track Individual Students

   The Texas data system that will be used in the proposed growth models meets this element since this system has a history of successfully tracking individual students and using student data to evaluate student growth.

7. Student Participation Rates and Achievement on an Additional Academic Indicator

   Texas requires schools and districts to meet the participation requirements related to all students in the tested grades. Texas will continue to use the other academic indicators of attendance rates for elementary and middle schools, graduation rates for high schools, and performance gains as required elements of AYP.

Alignment with Additional Guiding Principles

Texas' proposal also aligns with the additional guidance provided by USDE over the past few years. The guidance noted below comes from the May 17, 2006, cross-cutting themes document published after the first round of growth pilot peer review and USDE presentations and communications with states.

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

Texas' growth model uses only current year reading/English language arts and mathematics scores to predict student performance. Though USDE has recommended using data from multiple years, Texas uses only current year scores because it:

- makes the prediction process transparent. By developing formulas in one year and applying the formulas the next year, Texas makes the growth analyses transparent. Texas has a long history of using transparent formulas in the state accountability system, so this approach to the prediction equations matches well with the broader Texas assessment and accountability systems.
- facilitates the inclusion of as many students as possible. By using only reading/English language arts and mathematics scores, Texas includes almost all students (98% in study) and does not lose prediction accuracy (see bullet below).
- allows Texas to report student growth on confidential student reports during the regular reporting timeframe. By using prior-year formulas, this approach provides prediction feedback to students, parents, educators, campuses, districts, and the state as quickly as possible, so that score users can begin planning instructional interventions before the school year ends.
- capitalizes on the successful history of using current year scores in the Dallas growth model, which has been used since 1992.
- is supported by the prediction accuracy study presented in Appendix 1 in which predictions using the Texas model with current year scores and prior year formulas were as accurate as projections using all student scores and current year formulas.

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

By predicting to the next high-stakes grade and by using current-year scores in the prediction equations, the Texas proposal meets this guidance. Most schools in Texas are configured such that students in grades 3-5, 6-8, and 9-12 are in the same school. See Appendix B for counts of Texas campuses by different grade configurations. By predicting to grades 5, 8, and 11, campuses and districts are held accountable for the growth of students during the years the students are on the campuses. Furthermore, by using current year scores in the prediction equations, students who move from one campus to the next are not typically excluded from the growth calculations.

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

The Texas prediction method is used to predict performance for all students, regardless of the performance level of the student.

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

Texas will hold campuses accountable for the same groups as they did under the status model.

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

Texas does not intend to use confidence intervals in the predictions for TAKS, TAKS (Accommodated) or TAKS-M or in the growth expectations for students participating in TAKS-Alt.

6. States should not reset growth targets each year.

Texas does not propose resetting students' growth targets each year. The growth target for all students each year is proficiency in the next high-stakes grade. Students are expected to be on track to meeting proficiency by grades 5, 8, and 11. Texas proposes to calculate predictions each year for all students with sufficient data. This practice is similar to the practice approved by USDE in Tennessee and Ohio.

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

Texas does not propose to average scores for proficient and non-proficient students. Predictions and growth evaluations (for TAKS-Alt students) are made for each student.

8. States should not implement a growth model in addition to an index system.

Texas does not implement an index system, so this guidance does not apply to Texas.

9. States should not dilute accountability by adding growth to the accountability system.

By adding predictions to Texas' state and federal accountability systems, Texas will not dilute accountability. The purpose of including growth in Texas' state and federal accountability systems is to offer alternative approaches to demonstrating achievement that meet both state and NCLB goals. The intention is not to lower student performance expectations but to hold all students, student groups, public schools, local education agencies, and the state to the same high expectations that have been set.

10. States should ensure that all students are included in its growth model proposal to the extent possible, particularly students with disabilities who take an alternate assessment based on alternate academic achievement standards or an alternate assessment based on modified academic achievement standards.

Texas proposes to report growth on as many students as possible. By predicting proficiency for all students except those with the most severe cognitive disabilities and by applying a transition table approach to those students with the most severe cognitive disabilities, Texas will maximize the number of students for whom growth is reported.

11. Growth model information should be freely available for scientific scrutiny to enhance and validate the model.

Texas' decision to use prior-year formulas and make predictions using current year reading/English language arts and mathematics scores allows the prediction equations to be published and made publicly available. This practice of making calculations available to those who will use and be held accountable for the results is one Texas has embraced for many years.

12. The growth model should demonstrate maximum transparency *and* validity.

Texas' approach to measuring student growth was determined so that it maximized transparency without sacrificing validity. The prediction accuracy study illustrates this feature; however, the study only evaluated prediction accuracy on two grades and subjects. Texas will continue to document student predictions, follow cohorts of students, and conduct annual prediction accuracy studies as part of its continuous improvement efforts. In these studies, prediction accuracy for students in all grades and subjects will be evaluated and reported.

13. Appropriate statistical model for the assessment scales.

The statistical model and reliability and validity evidence supporting TAKS and TAKS (Accommodated) reading/English language arts and mathematics assessment scales in Texas have been reviewed and approved through the Standards and Assessment Peer Review. Although not yet approved through the peer review process, (submitted for peer review to USDE on October 3, 2008), the same rigorous statistical model and psychometric support will apply to the TAKS-M assessments. Due to the difficulty of gathering evidence to support that the TAKS-Alt scales are interval-level, linear, and measure only one type of achievement, Texas proposes a different type of growth model for this assessment, one that does not rely on the same psychometric properties as a prediction model.

Conclusion

Texas proposes to incorporate a measure of adequate yearly growth in Texas' AYP calculations starting in 2009. The proposed approaches to measuring student growth meet with the alignment and foundational elements of No Child Left Behind and with state statute, fit well with and enhance the Texas assessment and accountability systems, and include almost all Texas students. By including a measure of student growth in AYP calculations, Texas will promote a richer evaluation of student progress than offered currently with the status model and will focus the state's attention and intervention efforts on the progress each student is making.

Appendix 1
Development and Testing of the Texas Prediction Model

Background

Texas recently completed a pilot study to evaluate two possible approaches to measuring annual student improvement to satisfy House Bill 1 and Senate Bill 1031. Texas is proposing to use a measure of student growth as a criterion for campuses to meet adequate yearly progress (AYP) for federal reporting purposes in 2009. In addition, Texas will include student growth in the state accountability system for evaluating campuses and districts. The two approaches evaluated in the pilot study included a growth to proficiency model and a complex regression-based model much like the models approved for the USDE growth pilot program in Tennessee and Ohio. The complex regression-based models in the pilot study were the SAS® EVAAS® mixed-model longitudinal methods computed by Dr. Bill Sanders. These two approaches were chosen for the pilot study because they were well matched to the data conditions in Texas, offered the flexibility to potentially satisfy more than one requirement for growth measures, and could be adapted when end-of-course assessments are initiated in the 2011-2012 school year.

While the pilot study was being conducted, the Texas Select Committee on Public School Accountability convened two groups of district representatives at its April 2008 and August 2008 meetings. These representatives shared information about ways in which their districts have developed and used student growth measures. After hearing district testimony at the April meeting, the Select Committee expressed interest in using a regression-based model at the state level and questioned whether the state could implement a model like the one developed by the Dallas Independent School District (DISD), one that had been in use since 1992. In summer 2008, the Texas Education Agency (TEA) evaluated a model much like the Dallas model, using a few cohorts from the pilot study. A modified regression model similar to that used by Dallas ISD is the one currently proposed for the USDE pilot program.

The purpose of this document is to describe the development and testing of a student prediction model much like the Dallas ISD model using a procedure published by Lissitz, et. al. (2006). In particular, this document will describe (1) the procedures used to develop the model formulas in mathematics and reading/English language arts using 2006 data predicting 2007 scores, (2) how the formulas were applied in 2007 to predict 2008 scores, (3) the prediction accuracy of these models with these cohorts, and (4) how the prediction accuracy of these models in 2008 compares with the projection accuracy of the more complex regression-based models for this cohort in 2008.

Methods

The procedure used to fit Texas' regression-based prediction model was one recommended by the Texas Technical Advisory Committee at their July 14-15, 2008, meeting. It consists of two steps and is based on a method described in a paper by Lissitz, et al. (2006). The first step in the process is an ordinary least squares (OLS) multiple regression, which serves to identify variables that potentially significantly affect achievement. The second step is an analysis of the variability that is due to student clustering within schools, which determines whether or not multilevel modeling is justified.

Two sets of analyses were conducted. The first set used the 2007 8th-grade TAKS mathematics (*TAKS_M07*) score as the outcome variable for all analyses. The second set of analyses used the TAKS reading (*TAKS_R07*) scores as the outcome variable. The following section describes the procedures used with the first set of analyses, with the mathematics score as the outcome variable. The procedures were repeated for this cohort using the reading score as the outcome variable and for the grade 10 cohort using the mathematics and English language arts scores as the outcomes, respectively.

Procedures

With the 2007 8th-grade TAKS mathematics (*TAKS_M07*) score as the outcome variable, the initial group of student-level predictors entered into the OLS multiple regression included a 2006 7th-grade TAKS reading score (*TAKS_R06*) and a 2006 7th-grade mathematics score (*TAKS_M06*). These variables were aggregated at the campus level and included as predictors in the model as well (*MEAN_TAKS_R06*, *MEAN_TAKS_M06*).

The results from the initial OLS regression model indicated that the predictor variables accounted for 62.4% of the variance in the dependent variable, *TAKS_M07*. All predictor variables were significant at an alpha level of 0.05. Thus, all initial predictors were eligible to be included in a multilevel model. The second step of the process involved analyzing the variability of campus-level TAKS scores and the intra-class correlation in order to determine whether or not a multilevel model was justified. This analysis is conducted using what is known as an unconditional multilevel model. The unconditional model at level 1 can be defined as

$$TAKS\_M07_{ij} = \beta_{0j} + r_{ij} \qquad (1)$$

where $TAKS\_M07_{ij}$ represents the 2007 8th-grade mathematics score for individual *i* within school *j*, β_{0j} represents the mean *TAKS_M07* score for school *j*, and r_{ij} represents the residual for individual *i* within school *j*. The variance of $r_{ij} = \sigma^2$.

Level 2 of the unconditional multilevel model can be defined as

$$\beta_{0j} = \gamma_{00} + u_{0j} \qquad (2)$$

where γ_{00} is the grand mean of the *TAKS_M07* scores and u_{0j} is the residual for school *j* (i.e., the deviation of school *j* from the grand mean). The variance of $u_{0j} = \tau_{00}$. For the model under consideration, this variance was statistically significant (Z = 26.54, p < .001), meaning there was significant variability in the mean *TAKS_M07* scores among schools. The intra-class correlation is calculated as $\rho = \tau_{00} / \left(\tau_{00} + \sigma^2\right) = 5612/(5612 + 30758) = 0.15$, meaning 15% of the variance of *TAKS_M07* scores is attributable to the effects of students being clustered within schools.

Since the unconditional model indicated variability at the school level, a model with school- and student-level predictors was run. The model may be run twice: once with all the student and school-level predictors indicated by the OLS model, and, if necessary, again with variables that were not statistically significant in the multilevel model omitted. The predictor variables in the model under consideration were all statistically significant, resulting in a final level 1 model,

$$TAKS\_M07_{ij} = \beta_{0j} + \beta_{1j}\left(TAKS\_M06_j\right) + \beta_{2j}\left(TAKS\_R06_j\right) + r_{ij}.$$

In multilevel modeling, the level 1 regression coefficients (i.e., the βs) are tested for variability at level 2. If variability at level 2 is indicated, then level 2 predictor variables can be added to the model. After testing the coefficients for variability, and school-level predictor variables for statistical significance, the final level 2 model for the model under consideration was,

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(MEAN\_TAKS\_M06) + \gamma_{02}(MEAN\_TAKS\_R06) + u_{0j} \quad (3)$$

$$\beta_{1j} = \gamma_{10} \quad (4)$$

$$\beta_{2j} = \gamma_{20} \quad (5)$$

Equation 3 above illustrates that there is variability among the schools with respect to mean *TAKS_M06* scores, and some of that variability is accounted for by the aggregation of the *TAKS_M06* and *TAKS_R06* at the school level. Another way of stating this is that the achievement level of the school influences student achievement above and beyond what is predicted from the students' individual scores. Equations 4 and 5 indicate that there is no significant variation among schools with respect to the slopes associated with *TAKS_M06* and *TAKS_R06*. Substituting the level 2 equations into the level 1 equations results in the final multilevel prediction equation,

$$TAKS\_M07_{ij} = \gamma_{00} + \gamma_{01}(MEAN\_TAKS\_M06) + \gamma_{02}(MEAN\_TAKS\_R06) + \gamma_{10}(TAKS\_M06) + \gamma_{20}(TAKS\_R06) + (u_{0j} + r_{ij})$$

Inserting the regression coefficients from the solution gives the final prediction equation,

$$TAKS\_M07_{ij} = 572.68 + .02264(MEAN\_TAKS\_M06) + .07050(MEAN\_TAKS\_R06) + .5006(TAKS\_M06) + .1255(TAKS\_R06).$$

This equation developed using 2006 scores as predictors was then used to predict 2008 8th-grade mathematics scores for the 2007 7th-grade cohort. If a student's predicted score was 2100 (the Met Standard score) or above, that student was classified as meeting growth targets in 2007. If the predicted score was below 2100, that student was classified as not meeting growth targets. Finally, the accuracy of the growth classifications based on predicted scores was assessed by comparing them to the observed 2008 8th-grade results.

Results

The procedures described above were repeated three times, once to predict grade 8 reading for the 7th-grade cohort, once to predict grade 11 mathematics for the 10th-grade cohort, and once to predict grade 11 English language arts for the 10th-grade cohort. The percent of variance accounted for by the predictors and the intra-class correlation coefficients are presented in Table A1 for the cohorts. The unstandardized regression coefficients and p-values from the multilevel model equations for the two cohorts in both subjects are presented in Table A2. The projection accuracy results for all cohorts are contrasted with projection accuracy results from the more complex regression-based EVAAS projection model and presented in summary form in Table A3 and in more detail in Tables A4 through A11.

Table A1. *Texas Regression Model Results*

Prediction Grade and Subject	Year Formulas Developed	Year Formulas Applied	Year Prediction Accuracy Evaluated	Percent Variance Accounted for by Predictors	Intra-class Correlation Coefficient
Grade 8 Reading	2006	2007	2008	53.4%	0.11
Grade 8 Mathematics	2006	2007	2008	62.4%	0.15
Grade 11 English Language Arts	2006	2007	2008	56.1%	0.16
Grade 11 Mathematics	2006	2007	2008	70.2%	0.18

Table A2. *Texas Regression Model Unstandardized Regression Coefficients and p-values*

Indicators	7th-Grade Reading		10th-Grade Mathematics		10th-Grade English Language Arts	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Constant	121.05	< .0001	100.05	0.0007	64.7245	0.0176
Student-level variables						
TAKS_R06	0.5867	< .0001	0.1225	< .0001	0.5630	< .0001
TAKS_M06	0.2567	< .0001	0.6962	< .0001	0.2302	< .0001
School-level variables						
MEAN_TAKS_R06	0.2306	< .0001	0.03396	0.1375	0.1403	< .0001
MEAN_TAKS_M06	-0.08639	< .0001	0.1211	< .0001	0.06696	< .0001

Table A3. *Prediction Accuracy for Texas Regression-Based Model and More Complex Regression-Based Model*

PREDICTION YEAR, GRADE, AND SUBJECT	TEXAS MODEL			MORE COMPLEX MODEL		
	N	Perfect Agreement Met Standard	Perfect Agreement DNM	N	Perfect Agreement Met Standard	Perfect Agreement DNM
2008 Grade 8 Reading	270,700	94	2	269,015	94	2
2008 Grade 8 Mathematics	269,675	73	13	267,540	73	14
2008 Grade 11 English Language Arts	222,603	93	1	225,923	92	3
2008 Grade 11 Mathematics	224,341	79	10	228,110	78	11

Table A4. *Texas Regression Model 7th-grade Cohort: 8th-grade Reading*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Predicted Did Not Meet Growth	4474 (1.65)	3192 (1.18)	7666 (2.83)
Predicted Met Growth	7822 (2.89)	255212 (94.28)	263034 (97.17)
Total	12296 (4.54)	258404 (95.46)	270700 (100.00)

Table A5. *More Complex Model 7th-grade Cohort: 8th-grade Reading*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Projected Did Not Meet Growth	5097 (1.89)	3461 (1.29)	8558 (3.18)
Projected Met Growth	7235 (2.69)	253222 (94.13)	260457 (96.82)
Total	12332 (4.58)	256683 (95.42)	269015 (100.00)

Table A6. *Texas Regression Model 7th-grade Cohort: 8th-grade Mathematics*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Predicted Did Not Meet Growth	34632 (12.84)	18853 (6.99)	53485 (19.83)
Predicted Met Growth	18393 (6.82)	197797 (73.35)	211345 (80.17)
Total	53025 (19.66)	216650 (80.34)	269675 (100.00)

Table A7. *More Complex Model 7th-grade Cohort: 8th-grade Mathematics*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Projected Did Not Meet Growth	37007 (13.83)	19188 (7.17)	56195 (21.00)
Projected Met Growth	15882 (5.94)	195463 (73.06)	211345 (79.00)
Total	52889 (19.77)	214651 (80.23)	267540 (100.00)

Table A8. *Texas Regression Model* 10*th-grade Cohort: 11th-grade English Language Arts*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Predicted Did Not Meet Growth	3056 (1.36)	1037 (0.46)	4093 (1.82)
Predicted Met Growth	10889 (4.85)	209359 (93.32)	220248 (98.18)
Total	13945 (6.22)	210396 (93.78)	224341 (100.00)

Table A9. *More Complex Model* 10*th-grade Cohort: 11th-grade English Language Arts*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Projected Did Not Meet Growth	6555 (2.87)	3877 (1.70)	10432 (4.57)
Projected Met Growth	8520 (3.74)	209158 (91.69)	217678 (95.43)
Total	15075 (6.61)	213035 (93.39)	228110 (100.00)

Table A10. *Texas Regression Model 10th-grade Cohort: 11th-grade Mathematics*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Predicted Did Not Meet Growth	22138 (9.95)	9419 (4.23)	31557 (14.18)
Predicted Met Growth	16177 (7.27)	174869 (78.56)	191046 (85.82)
Total	38315 (17.21)	184288 (82.79)	222603 (100.00)

Table A11. More Complex Model 10th-grade Cohort: 11th-grade Mathematics

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Projected Did Not Meet Growth	25484 (11.28)	9550 (4.23)	35034 (15.51)
Projected Met Growth	14976 (6.63)	175913 (77.86)	190889 (84.49)
Total	40460 (17.91)	185463 (82.09)	225923 (100.00)

Discussion

Results from this study indicated that prediction accuracy for the Texas regression model was similar to projection accuracy with the more complex models for these cohorts. Thus, the Texas model would appear to share some of the advantages of projection models in general and the EVAAS model in particular. These advantages include that the Texas prediction models have evidence supporting their accuracy and reliability, offer the flexibility to be adapted when end-of-course assessments are initiated in the 2011-2012 school year, and would likely be beneficial in the calculation of Adequate Yearly Progress for campuses, districts, and the state. Though the Texas model is similar and produces similar results to the EVAAS model, it is simpler and uses formulas from the prior year, so that the process for predicting student performance as an indicator of student growth is transparent to the state and can be reported on the Confidential Student Report during Texas' regularly-scheduled reporting timeframe. State law requires that schools receive results from the first administration of TAKS within ten working days of receipt of the test materials by the testing contractor.

Because the Texas model is simpler than the EVAAS model, it lacks some of the flexibility of the more complex model in handling missing data. Students must have valid scores in both reading/English language arts and mathematics to be predicted in the Texas model. Whereas the simplicity of the model is a disadvantage with regard to missing data, it nevertheless has some advantages over the more complex model. For example, the Texas prediction model is easy to implement using standard statistical software, so the turnaround time between test completion and prediction calculation would be relatively short, and student prediction results could be reported on the student reports and used in instructional planning as early as possible. Furthermore, the regression coefficients could be made publicly available so that school and district personnel would be able to calculate predicted scores relatively easily. Though the intricacies of the development of the multilevel regression equations may be difficult for stakeholders to understand, the basic idea of using students' current-year test scores to predict future performance is fairly straightforward.

A potential disadvantage of any regression-based model is explaining the methodology to stakeholders who do not have a statistical background. In addition, the accuracy and reliability of regression-based models are likely to decrease the closer a student is to the classification cut score, where small errors can mean the difference between being classified correctly and incorrectly. It is also expected that prediction accuracy decreases as the time between testing and final growth target increases. Analyses to evaluate prediction accuracy for all grades and subjects are planned as annual analyses, so that prediction accuracy for students at all score points and those being predicted one, two, or three years in the future can be documented and monitored over time.

References

Lissitz, R. W., Fan, W., Alban, T., Hislop, B., Strader, D., Wood, C., et al. (March, 2006). *The prediction of performance on the Maryland high school graduation exam: Magnitude, modeling and reliability of results*. A paper presented at the National Council on Measurement in Education, San Francisco.

Appendix 2
Grade Configuration of Texas Campuses

2008 School Types
Counts of Schools for Each Low and High Grade Combination

High Grade →

Low Grade ↓

	EE	PK	K	1	2	3	4	5	6	7	8	9	10	11	12
EE	9	56	49	43	68	57	174	1001	159	3	13	0	0	2	29
PK		25	15	15	20	38	143	896	227	7	66	0	2	3	121
K			0	5	8	23	109	510	171	6	36	8	4	1	54
1				1	16	19	17	47	21	1	1	1	2	5	10
2					3	22	15	22	4	0	1	0	2	4	6
3						1	15	84	9	1	3	0	0	4	6
4							3	64	51	1	3	2	0	1	15
5								12	137	3	78	3	3	10	16
6									41	7	986	15	12	42	113
7										4	294	27	18	36	157
8											13	9	8	11	46
9												51	37	31	1187
10													25	7	39
11														12	25
12															21

Elementary | Junior High | Middle School | Elementary/Secondary | High School