

CHAPTER 12: THE TEXAS GROWTH INDEX (TGI)

Introduction

The Texas Growth Index (TGI) is 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 in the state accountability system as the basis for issuance of the Gold Performance Acknowledgments for Comparable Improvement in Reading/ELA and in Mathematics. It is also used to calculate the TAKS Progress Indicator in reading/ELA, mathematics, science, and social studies under the alternative education accountability (AEA) procedures.

Appropriate Use of the Texas Growth Index

The TGI was primarily designed for use in accountability. It is intended for use at the campus level and district level and is not intended for use at the individual student level. The TGI is based on TAKS scale score changes between spring 2003 and spring 2004. The analyses establishing the TGI did not include retesting students. It, therefore, should not be calculated for students retesting on either the exit-level TAKS or TAKS retest administrations at the SSI grades. Finally, the TGI was not designed to compare the growth of different classrooms within a campus and, therefore, is not be used to evaluate teachers.

The following sections describe development work for the TGI, the formulas used in the TGI, steps for calculating the TGI, and an example calculation of the TGI.

Development Work for the Texas Growth Index

In 2005, the Texas Education Agency (TEA) and its test contractor cooperatively developed the Texas Growth Index. In the development of the index, various growth models were evaluated. Research was limited to models that examined growth from one year to the next only, as TAKS was not built using a vertical scale. The model evaluation results were presented to the Texas Technical Advisory Committee (TTAC) at the September 2004 meeting.

Research and development work was initially conducted using data from the 2002–2003 school year. However, since these data were field-test data, analyses from that year were considered preliminary. Additional analyses were conducted using data from the 2003–2004 school year, which served as the baseline year for the establishment of TGI parameters. Follow-up analyses using data from the 2004–2005 school year also were examined. Results from these analyses were used to recommend the model used in the TGI. Expected growth for each model estimated how students who tested with TAKS in spring 2003 performed when tested in spring 2004 in the next grade level of TAKS. These estimates were used to develop expected scores (that is, predicted scores) that were compared to students' actual scores.

The TGI was based on performance only. For example, if two grade 6 students had identical scale scores on the spring 2003 TAKS grade 6 reading test, their expected growth for the spring 2004 TAKS grade 7 reading test had to be identical, regardless of differences in contextual variables such as ethnicity, economic disadvantaged status, and so forth. In other words, growth expectations were the same for similarly performing students. Even though contextual variables were not included in setting student-level growth expectations, comparative evaluations and/or decisions at the campus or district level using the TGI may include contextual variables. The inclusion of the contextual variables at the campus and district levels would allow for an index similar to the comparable improvement measures used by the Division of Performance Reporting of TEA under the TAAS system.

Models

In the TGI development work, four models were evaluated for measuring student growth (where student growth is defined to be the difference between the observed 2004 score and the estimated 2004 score). The four models are described below.

1. Ordinary least squares (OLS) linear regression, which uses students' 2003 test scores to predict their 2004 test scores and accounts for measurement error in the 2004 test scores;
2. Structural equation modeling (SEM), which takes measurement error in both the 2003 and 2004 test scores into account through the inclusion of reliability parameters;
3. Deming's errors-in-variables regression model (Combleet & Gochman, 1979), which also takes measurement error in both the 2003 and 2004 test scores into account by minimizing the perpendicular distances from students' test scores on both years to the regression line; and
4. The mean/sigma (MS) approach (Hambleton, Swaminathan, and Rogers, 1991), which is a linear equating method that produces predicted scores based on the linear conversion parameters for a slope and intercept that set the mean and standard deviation of the pre-test and post-test distributions equal.

Model Evaluation

As part of the evaluation of these four models, the estimated growth parameters were initially calculated and compared for the four models across all grades for mathematics and reading. Several criteria were used to determine which of the models performed best. First, a residual score was calculated for each student and for each method. This residual score was defined as the student's actual score minus his/her predicted score for each of the models evaluated. A positive residual indicated growth above expectations, and a negative residual indicated growth below expectations. The residual scores were compared across districts and campuses for groups of students, where the groups were defined based on students' performance levels over the two years. For example, one group consisted of students who met the standard in 2003 and 2004 and another group consisted of students who met the standard in 2003 and achieved Commended Performance in 2004.

To compare residuals across districts and campuses, the mean residuals for all students in each district and campus were first calculated. For each grade level, there were approximately 1100 districts and 2000 campuses. Summary statistics (for example, ranges and medians) for those mean district and campus residuals at each grade level were calculated. A comparison of the medians and ranges of the mean residuals across the four methods was conducted as a second way to evaluate the four models. These summary statistics for all districts and campuses and for only those districts and campuses with at least 10 students were also examined.

After examining residuals, the third evaluation method involved comparing the models with respect to the proportions of students showing growth above expectations for the student groups defined by their performance levels. In these analyses, special consideration was given to the model demonstrating the highest occurrence of student growth especially at the lower performance levels (that is, the model that was the least disadvantaging to students, campuses, and districts).

A fourth evaluation approach involved an equipercentile method that evaluated the models at selected percentiles of score distributions (the percentiles chosen were the 10th to the 90th by intervals of 10). For example, the mathematics scale score on grade X in 2003 that corresponded to the 10th percentile was assumed to be equal to the mathematics scale score on grade $X+1$ in 2004 that also corresponded to the 10th percentile. Then, that 2003 scale score was inserted into the prediction equation for each of the four models. At each percentile selected for study, the model that produced the predicted score closest to the 2004 score was deemed to be the most accurate method. Root mean squared errors (RMSEs) were also calculated across the nine studied percentiles for each method as an overall indication of model accuracy.

Finally, in choosing among the competing models, ease of interpretability of the model to parents, campuses, and districts was considered. Specifically, OLS regression was considered to be more explainable than structural equation modeling. In addition, the mean/sigma model was considered to be more explainable than the errors-in-variables regression model. Between models with similar results, it was thought that the simpler model should be chosen.

Results

The results from the analyses indicated several notable trends. First, results were similar for OLS regression and structural equation modeling and for the mean/sigma and errors-in-variables regression models. Furthermore, less growth was expected from students at the lower ends of the distributions and more growth was expected from students at the higher ends of the distributions if the mean/sigma or errors-in-variables models were used rather than OLS regression or structural equation modeling. Graphs that show the prediction lines for each model in three sets of grades are reported in the "Texas Growth Index Report" in the 2007 Texas Education Agency Technical Report Series which can be found at <http://www.tea.state.tx.us/student.assessment/resources/techdig07/index.html>.

Second, residuals were estimated; a summary of the residuals across performance levels and at the district and campus levels is presented in the "Texas Growth Index Report". The

mean/sigma or errors-in-variables models yielded the smallest mean residuals for students who did not meet the standard both years (the B-B group in the "Texas Growth Index Report"). The four models performed similarly for the middle achievement levels or those students who met the standard both years (M-M in the "Texas Growth Index Report"). For the upper grades, the OLS regression or structural equation modeling approaches yielded slightly smaller mean residuals for students who met the standard both years.

Residual comparisons at the district and campus levels were made based on summary statistics for the district and campus mean residuals. The medians for the district and campus mean residuals were compared across the four models. The median district and campus mean residuals for the mean/sigma approach were smallest for most of the reading comparisons. For comparisons of median district and campus mean residuals in mathematics, OLS regression tended to yield the smallest medians, though medians for the mean/sigma method were close to those from OLS regression.

Third, the mean/sigma and errors-in-variables models tended to yield slightly more students with growth above expectations overall and for lower performing students compared with OLS regression and structural equation modeling. For the higher performing students at campuses and districts, the mean/sigma and errors-in-variables models yielded fewer students with growth above expectations compared with OLS regression and structural equation modeling. For more information, see the "Texas Growth Index Report" in the 2007 Texas Education Agency Technical Report Series which can be found at <http://www.tea.state.tx.us/student.assessment/resources/techdig07/index.html>.

Fourth, results from the equipercentile evaluation analyses underscored the findings from other analyses. As seen in the "Texas Growth Index Report", the root mean square errors (RMSE) were typically smallest for the mean/sigma and errors-in-variables models.

Summary and Recommendation

Based on the evaluation criteria and the results summarized above, the mean/sigma model was recommended for use in TEA's accountability system. This model seemed to be less disadvantaging to lower performing students, campuses, and districts (meaning that those not meeting the performance standard are more likely to demonstrate growth) and it appeared to be fairly accurate across the performance-level combinations compared with the other three methods. At the September 2004 meeting, the Texas Technical Advisory Committee supported this recommendation.

However, the results of the analyses should be interpreted with caution. The TAKS tests are designed to provide the most information about students near the middle of the distribution, around the cut scores that define the student achievement levels. The TAKS tests are not designed to provide as much information for students who perform at the upper and lower ends of the proficiency distribution. Therefore, growth estimates for students at the upper and lower ends of the proficiency distribution will not be as accurate as growth estimates in the middle of the distribution. The differential information provided for students across the proficiency distribution is one reason why the TGI is not designed for use at the student level. Another reason the results should be interpreted with caution is that determining the best

method for evaluating the models depends on the anticipated application of the model. The evaluation methods were selected assuming the growth model would be applied as part of the accountability system. However, if the growth model were to be applied in different ways, it is possible that other evaluation criteria could be used.

The TGI Formulas

The TGI formulas are initially applied to scores from individual students prior to aggregation of TGI values to form the TGI for the campus, district, or other larger unit. Student-level TGI values have low reliability and, as such, the TGI is not designed to be used at the student level. Only TGI scores aggregated at the campus level or at levels larger in size, such as districts, should be used.

The TGI uses a formula to describe the growth expected for students from one year (YR_1) to the next (YR_2) in a subject. The formula uses students' scores on YR_1 to define students' expected YR_2 scores. Students' actual growth (or score changes over the two growth years) is then compared to the expected growth based on the TGI formula. Finally the TGI is adjusted so the resulting numbers are comparable across grades and subjects.

Data from 2003 to 2004 for all students who had test scores in a particular subject for both years were used to estimate the formula for predicting a student's growth in that subject. The intercept and slope from the formula were calculated using a linear equating method. In this method, the z-scores (or standardized scores) from the two consecutive grades were set equal to each other. The intercept and slope were estimated with the following statistical approach:

$$\begin{aligned}\hat{Y} &= MX + B \\ \text{where} \\ M &= s_y/s_x \\ \text{and} \\ B &= \bar{Y} - \bar{X}(s_y/s_x),\end{aligned}$$

where \hat{Y} is the YR_2 predicted score for a student, M and B are the slope and intercept parameters, X is the student's scale score in YR_1 , s_x and s_y represent the standard deviations for students' scale scores in YR_1 and YR_2 , and \bar{X} and \bar{Y} are the mean scale scores for YR_1 and YR_2 .

A student's growth is defined as the student's YR_2 observed score minus the student's YR_2 expected score (taking into account the student's score in the first growth year). The number from that subtraction is divided by an adjustment, where the adjustment is the standard deviation of the residuals, or the differences between students' actual and predicted YR_2 scores. If the student's adjusted TGI is positive, the student is considered to have grown, and the magnitude of the adjusted TGI indicates the number of standard deviation units of growth. If the student's adjusted TGI is below the expected score, the student is considered to have regressed. For example, a -1.00 adjusted TGI means that a student performed one standard deviation below what was expected. Similarly, a 1.50 adjusted TGI means that a student performed 1.5 standard deviations higher than expected.

The resulting formula is one for a straight line, such that it includes an intercept and a growth rate (or slope). The intercept is the score we expect for a student who had a zero score in YR_1 . Since a zero is not an obtainable score on the TAKS, this intercept is not particularly meaningful beyond its use in the formula. The slope is the number of scale score points expected to increase in YR_2 for every scale score point increase in YR_1 .

As an example, growth from grades 3 to 4 in mathematics is predicted from the following formula:

$$\text{grade 4 Mathematics Scale Score} = -3.38 + 1.006 (\text{grade 3 Mathematics Scale Score})$$

In this formula, the -3.38 is the intercept. It means that for a student with a zero in grade 3 mathematics, the expectation would be for that student to have a -3.38 scale score in grade 4 mathematics. It is important to note that the scale scores in grades 3 and 4 are not on the same scale and that a zero is not an obtainable TAKS scale score. It is, therefore, necessary to consider also the grade 3 scale score in the equation. The slope, or 1.006 , is multiplied by the student's scale score in grade 3. The 1.006 slope value indicates that for every grade 3 scale score point increase, we expect the grade 4 scale score points to increase 1.006 .

Steps for Calculating TGI

The parameters used to determine the TGI (shown in tables below) were developed using the empirical data from the base comparison years — spring 2003 to spring 2004. The following steps are used to determine student-level TGI, which is then aggregated at the campus or district level. Student growth is estimated as a line with an intercept (or starting point) and slope (or increase).

Listed below are the steps taken for calculating TGI growth at the student level:

1. Find the starting point for an individual student in the row of the table below that matches that student's grade and subject.
2. Take the student's scale score in 2006.
3. Find the increase for that student in the row of the table below that matches that student's grade and subject.
4. Multiply the student's scale score from 2006 by the increase.
5. Add the amount from Step 1 and the total from Step 4. This is the expected student scale score for 2007.
6. Take the student's actual scale score from 2007 and subtract the expected student score from it. This number is the difference in expectation.
7. Calculate Adjusted TGI by dividing the result from Step 6 by the Adjustment factor shown on the tables below. Round to the second decimal place.

8. If the difference in expectation is positive, that student's performance grew more than expected. If the difference in expectation is negative, that student's performance grew less than expected.

The TGI for a campus or district is the mean of the student-level TGI scores for all students in the campus or district. A TGI of zero means that the year-to-year change in average scale score is equal to the average predicted change as calculated in the 2003–2004 base comparison years. A positive TGI means the group demonstrated growth that is larger than the expected growth for that group. A negative TGI indicates the group grew less than expected.

The TGI growth equation parameters for the state are in the tables below. These parameters were calculated over the 2003–2004 base comparison years and are applied in measuring growth across subsequent years.

Table 7. TGI Growth Equation Parameters—Mathematics and Science

Growth Grades	Subject	Starting Point	Increase	Adjustment
3–4	Mathematics	–3.38	1.01	138.07
3–4 (Spanish)	Mathematics	–903.49	1.44	190.11
4–5	Mathematics	–530.83	1.26	160.01
4–5 (Spanish)	Mathematics	–32.22	1.03	160.29
5–6	Mathematics	–167.96	1.09	152.94
5–6 (Spanish)	Mathematics	–11.10	1.04	173.12
6–7	Mathematics	612.26	0.71	95.40
7–8	Mathematics	–544.89	1.27	118.89
8–9	Mathematics	–775.75	1.38	136.19
9–10	Mathematics	480.79	0.77	95.47
10–11	Mathematics	–138.43	1.09	104.38
10–11	Science	410.23	0.83	75.94

Table 8. TGI Growth Equation Parameters—Reading, ELA, and Social Studies

Growth Grades	Subject	Starting Point	Increase	Adjustment
3–4	Reading	-12.89	0.99	135.97
3–4 (Spanish)	Reading	-158.07	1.03	158.44
4–5	Reading	-520.23	1.24	149.93
4–5 (Spanish)	Reading	-480.94	1.24	159.13
5–6	Reading	-66.29	1.07	151.85
5–6 (Spanish)	Reading	109.69	0.99	143.36
6–7	Reading	372.28	0.83	126.53
7–8	Reading	-87.53	1.07	128.61
8–9	Reading	712.12	0.66	101.31
9–10	Reading/ELA	535.21	0.76	91.11
10–11	ELA	128.38	0.96	96.41
10–11	Social Studies	464.43	0.81	93.98

Sample TGI Calculation

The following figure provides an example of how the steps above would be applied. Suppose one intends to examine a student’s mathematics growth from grade 10 to grade 11. Suppose the student had a scale score of 2188 in grade 10 and a scale score of 2161 in grade 11.

Figure 8. Sample TGI Calculation

	Steps	Example Values
Step 1	Find the starting point for the student in the row of the table that matches the student’s grade and subject.	-138.428
Step 2	Take the student’s scale score in the first year.	2188
Step 3	Find the increase for the student in the row of the table that matches the student’s grade and subject.	1.092
Step 4	Multiply the student’s scale score from the first year by the increase.	$2188 \times 1.092 = 2389.296$
Step 5	Add the amount from Step 1 and the total from Step 4. This is the expected student scale score for the second year.	$-138.428 + 2389.296 = 2250.868$
Step 6	Take the student’s scale score from the second year and subtract the expected student score from it. This number is the difference in expectation.	$2161 - 2250.868 = -89.868$
Step 7	Calculate Adjusted TGI by dividing the result from Step 6 by the Adjustment factor shown in the above tables. Round to the second decimal place.	$-89.868/104.38 = -0.86$
Step 8	If the difference in expectation is positive, the student’s performance grew more than expected. If the difference in expectation is negative, the student’s performance grew less than expected.	Since -0.86 is negative, the growth in student performance was less than expected.

How TGI is Used in Determining Comparable Improvement

Comparable Improvement (CI) is calculated separately for TAKS reading/ELA and TAKS mathematics. The student-level TGI values for each subject are aggregated to the campus level to create an average TGI for each campus.

Who is Included

Students included in a campus's Comparable Improvement are those who

- took the spring 2007 TAKS reading/ELA and/or mathematics tests, in grades 4–11;
- are part of the 2007 *Accountability Subset*;
- can be matched to the spring 2006 TAKS administration to find their prior year TAKS performance for reading/ELA and/or mathematics; and
- have been promoted to one higher grade than in 2006.

Calculating Average TGI

$$\text{average TGI (reading/ELA)} = \frac{\text{sum of individual student TGI values for reading/ELA}}{\text{total number of students with TGI in reading/ELA}}$$

$$\text{average TGI (mathematics)} = \frac{\text{sum of individual student TGI values for mathematics}}{\text{total number of students with TGI in mathematics}}$$

Once the average TGI of a campus is determined, it is listed with the other average TGIs of all 40 campuses in that campus's comparison group. The campuses are arranged from highest to lowest average TGI. If the target campus falls in the top quartile and all other eligibility criteria are met, it is awarded a Gold Performance Acknowledgment for Comparable Improvement. For more information about Comparable Improvement and the determination of a campus's comparison group, refer to the following page at the TEA web site: <http://www.tea.state.tx.us/perfreport/ci/2006/index.html>

Other Information

- *Retesters.* The analyses establishing the TGI did not include the retest administrations; therefore, it is calculated from the first administration for grade 11 exit-level students and for the first administration in the SSI grades—grade 3 reading and grade 5 reading and mathematics.
- *Quartile Size.* Because there are 40 campuses in a comparison group, there are usually 10 campuses in each quartile (with the target campus being the 11th campus in its quartile). Exceptions to this occur when groups have tied average TGI values at the border between quartiles, or when a campus in a group has too few matched students, and is, therefore, not assigned an average TGI value or a quartile. This will cause the number of campuses in each quartile to vary.
- *Quartile Position Across Subjects.* A campus's quartile position can vary by subject. For instance, a campus may be Q1 in reading, but it may be Q2 in mathematics. Quartile position is relative to the performance of the other campuses in the comparison group.

- *Quartile Position Across Groups.* A campus may be Q1 for its own group and Q4 as a member of another campus's group. However, the quartile value evaluated for a particular campus is the one determined for the campus's own group.
- *Minimum Size.* Any campus with fewer than 10 matched students for a subject will not have average TGI values calculated and will not be assigned a quartile position.
- *Number of Matched Students.* The number of matched students for reading may differ from the number of matched students for mathematics.
- *TGI Uses.* The TGI is not intended for use with individual students, nor is it intended for comparing the growth of different classrooms within a campus to evaluate teachers.
- *Negative TGI Values.* The TGI is a statistic with a mean of zero; negative values for students indicate the growth is less than expected. A negative TGI does not mean that performance of students declined from the prior year. Campuses with negative TGI values are not prohibited from earning CI acknowledgements.

For a more detailed explanation of Gold Performance Acknowledgment, see Chapter 5: Gold Performance Acknowledgments of the *2007 Accountability Manual*. This manual can be found online at <http://www.tea.state.tx.us/perfreport/account/2007/manual/>.

How TGI is Used in Determining the TAKS Progress Measure

The TAKS progress measure is used in evaluating registered alternative education campuses (ACEs). For an explanation of how TGI is used in the progress measure, see Chapter 10: AEA Base Indicators in the *2007 Accountability Manual*.

